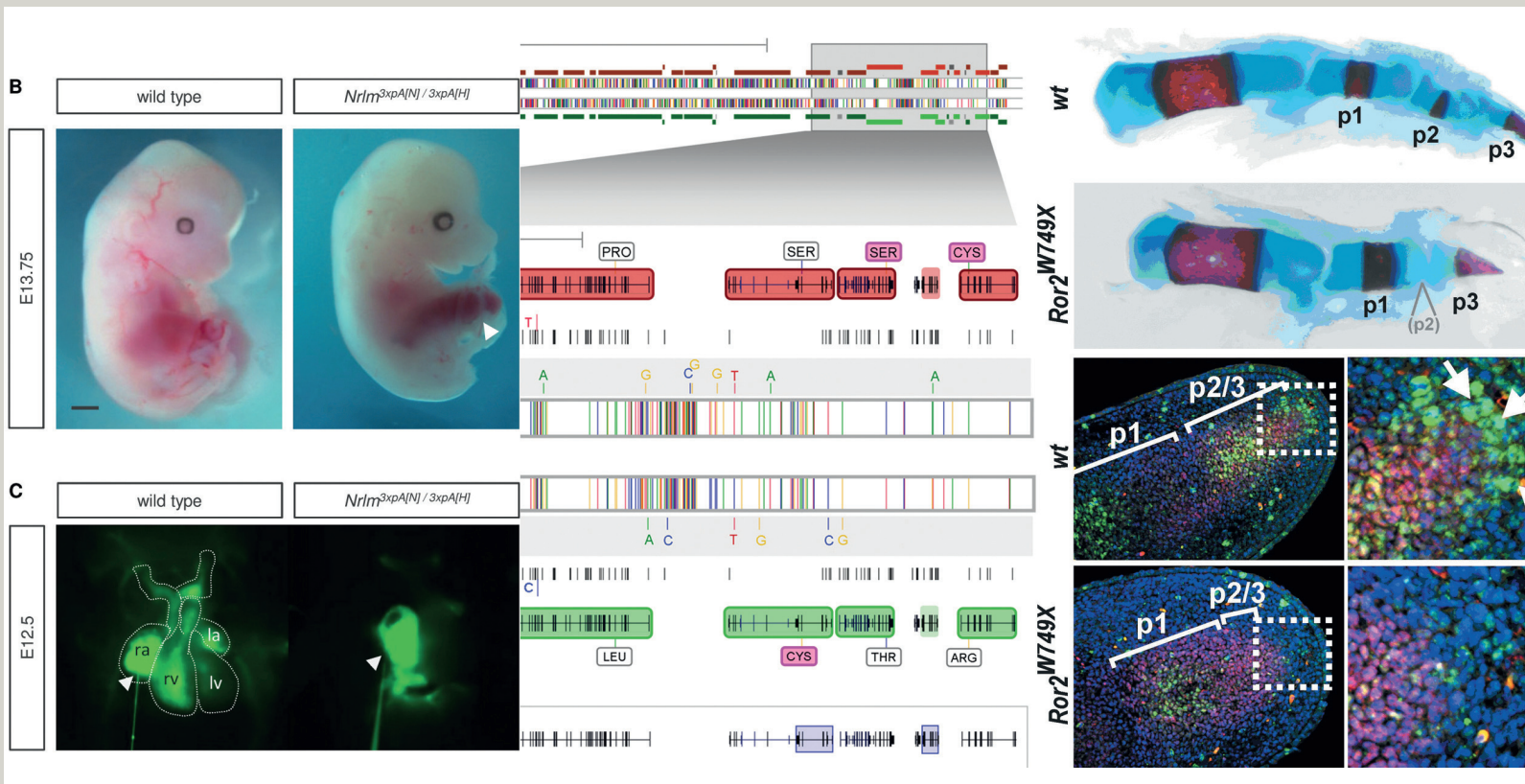
MPIMG

MAX-PLANCK-GESELLSCHAFT

Research Report 2012

Max Planck Institute for Molecular Genetics, Berlin



Imprint | Research Report 2012

Published by the Max Planck Institute for Molecular Genetics (MPIMG), Berlin, Germany, August 2012

Editorial Board:	Bernhard G. Herrmann, Hans Lehrach, H.-Hilger Ropers, Martin Vingron
Coordination:	Patricia Marquardt
Photography:	Katrin Ullrich, David Ausserhofer (p 14), Norbert Michalke (p 211)
Scientific Illustrations:	MPIMG
Production:	Thomas Didier, Meta Druck
Copies:	1,000

Contact:	Max Planck Institute for Molecular Genetics Ihnestr. 63 – 73 14195 Berlin Germany
Phone:	+49 (0)30 8413-0
Fax:	+49 (0)30 8413-1207
Email:	info@molgen.mpg.de

For further information about the MPIMG, see http://www.molgen.mpg.de

Research Report 2012

Max Planck Institute for Molecular Genetics

1

Berlin, August 2012

2

New perspectives: roof light in the entrance hall of the new building (tower 3)

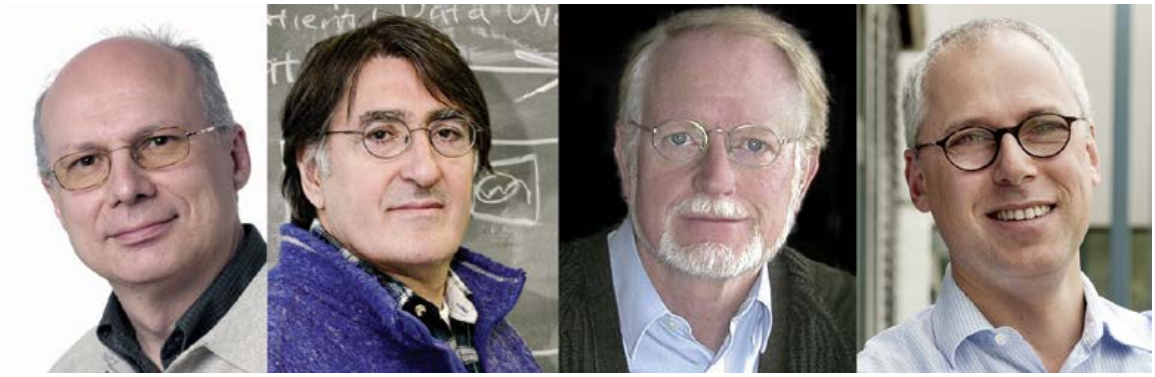
Table of contents

Please note:
In the publication lists of the group reports, group members are underlined. In the publication lists of the departments, department members are underlined.

Foreword

It is our pleasure to present the 2012 Research Report of the Max Planck Institute for Molecular Genetics (MPIMG) in Berlin. This report covers the period from 2009 to mid 2012. Founded in 1964, the MPIMG already houses its second generation of scientific members, who have firmly established the Institute as a centre for research at the interface of genomics and genetics. The MPIMG concentrates on genome analysis of humans and other organisms to elucidate cellular processes and genetic diseases. It is the overall goal of all our groups to gain new insights into the mechanisms of diseases on a molecular level, thus contributing to the development of cause-related new medical treatments.
Our institute will go through a major transition period due to the upcoming retirement of two of its directors, Hans Lehrach and H.-Hilger Ropers, in 2014. Thus, this report will both give a broad summary of the scientific work of the MPIMG during the last 3.5 years and describe the ongoing changes. We hope that it will provide a clear impression of the institute and the work we are doing here.

Bernhard G. Herrmann
Hans Lehrach
H.-Hilger Ropers
Martin Vingron



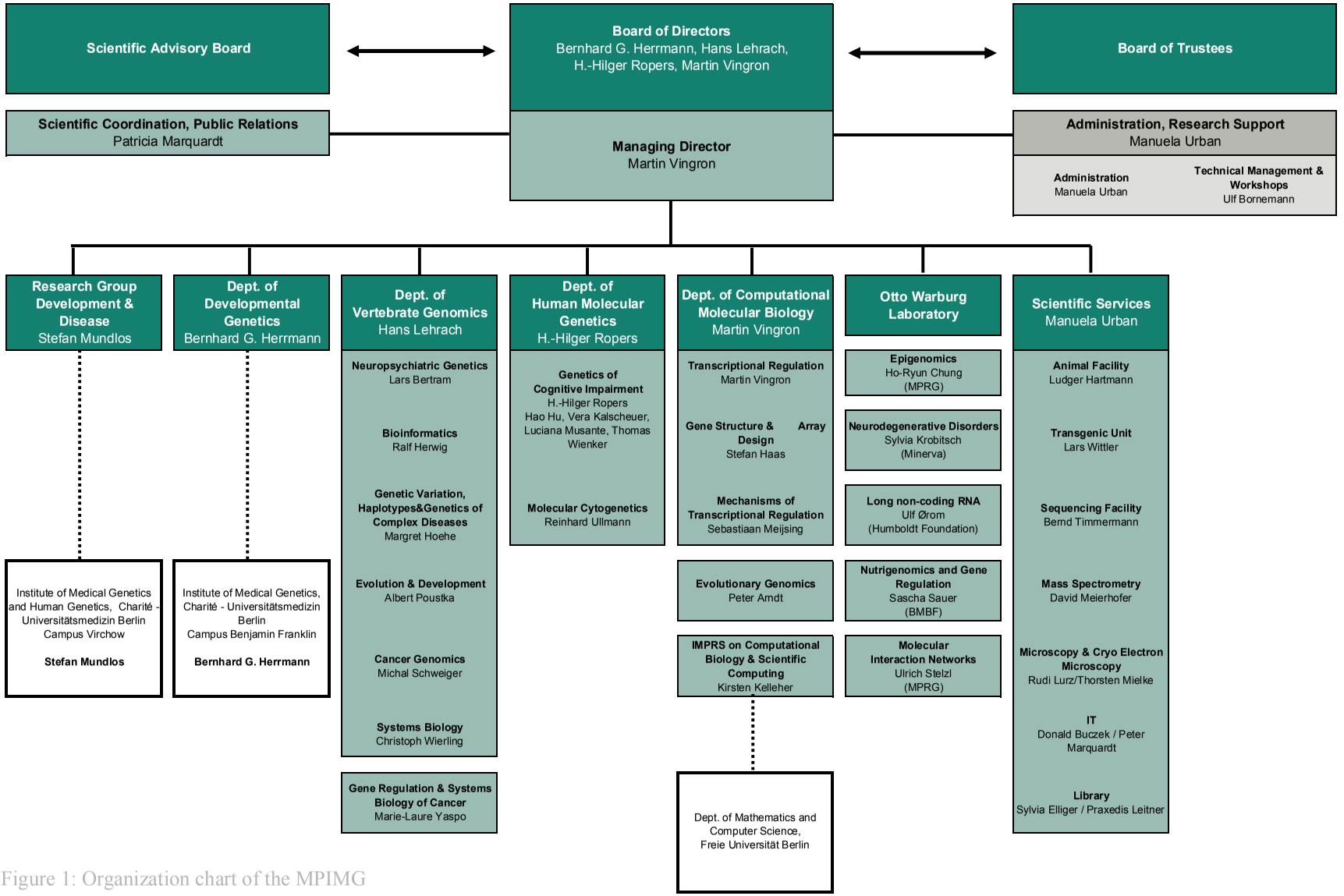


Figure 1: Organization chart of the MPIMG

The Max Planck Institute for Molecular Genetics

Overview

Scientific concept and institute structure

The general goal of molecular genetics lies in the study of how life processes function as a consequence of the genetic make-up of an organism. When applied to humans, this is particularly important for the understanding of disease processes. Genome research, the systematic study of genes and genomes, has changed the way, in which research in molecular genetics is pursued. Genomic technologies allow posing scientific questions broadly in order to determine all parts of a genome that influence the process under study. The Max Planck Institute for Molecular Genetics (MPIMG) works at the interface of genome research and genetics, concentrating on genome analysis of humans and other organisms to elucidate cellular processes and genetic diseases. It is the overall goal of the combined efforts of all MPIMG groups to gain new insights into the development of diseases on a molecular level, thus contributing to the development of cause-related new medical treatments. Research at the institute combines genome research, genetics, and computational biology.

The MPIMG is one of the larger institutes of the Max Planck Society. It consists of four independent departments, each headed by its own director. The individual departments concentrate on Developmental Genetics (Bernhard Herrmann), Vertebrate Genomics (Hans Lehrach), Human Molecular Genetics (H.-Hilger Ropers) and Computational Molecular Biology (Martin Vingron). They are complemented by an independent research group Development & Disease of Stefan Mundlos, who also heads the Institute of Medical Genetics and Human Genetics at the Charité – Universitätsmedizin Berlin, and the Otto Warburg Laboratory (OWL) that offers excellent junior scientists to work on their own scientific programs with an independent research group over a longer, but limited period of time (usually between five and nine years). Current members of the OWL are Ho-Ryun Chung (Epigenomics), Sylvia Krobitsch (Neurodegenerative Disorders), Ulf Ørom (Long non-coding RNA), Sascha Sauer (Nutrigenomics/Gene Regulation), and Ulrich Stelzl (Molecular Interaction Networks). The scientific groups are supported by a number of scientific service groups that maintain a range of core technologies, and the general administration/research support.

Scientific highlights

The main results of the scientific work of the MPIMG during the last three years are described in detail in the research reports of the single departments. At this point, we wish to present some of the most important and interesting results to give a general impression of the research performed at the institute.

One of the highlights is a work of Hilger Ropers and his co-workers, who describe the identification of 50 hitherto unknown genetic causes of intellectual disability by using new next generation sequencing technologies. The findings, which have been published in Nature in 2011, demonstrate the enormous genetic diversity of intellectual disabilities and subdivide them into different monogenetic defects. This will not only help families to obtain a reliable diagnosis, but also provide a model for the explanation of related disorders, such as autism, schizophrenia and epilepsy.

In 2008, Hans Lehrach has been invited to join the 1000 Genomes project, an international project that aims at finding most genetic variants with frequencies of at least 1% in the populations studied. Initial results of the project have been published in Nature, Nature Genetics and Science in 2010 and 2011. This is complemented by research on two projects of the International Cancer Genome Consortium as well as the "Virtual Patient" models, forming the basis to a new approach of truly personalized cancer treatment.

Margret Höhe from the Department of Vertebrate Genomics has been the first and second to comprehensively decode both sets of parental chromosomes of a human genome separately. This "diploid genomics" is essential for gaining a deeper understanding of human biology, the analysis of disease risks and, accordingly, the development of new and more individualised strategies for the prevention and treatment of diseases.

Using quantitative mass spectrometry coupled to high-pressure and/or nano-flow liquid chromatography, Markus Ralser and his team have been very successful in performing targeted quantitative and qualitative analysis of small molecules and peptides to gain insights into the regulatory function of metabolic networks in aging processes and cancer. Ralser, who has already done his PhD work at the MPIMG in the group of Sylvia Krobitsch, received a prestigious ERC Starting Grant from the European Research Council and changed to the University of Cambridge, UK, in January 2012.

The analysis of gene regulatory networks belongs to the central themes of many MPIMG groups. For functional analysis of multiple genes, Bernhard Herrmann developed an integrated vector system for inducible gene silencing. The system allows the dissection of gene function at unprecedented detail and speed, and provides tight control of the genetic background minimizing intrinsic variation.

Martin Vingron and his co-workers have developed the TRAP method for Transcription Factor Affinity Prediction that calculates the affinity of transcription factors for DNA sequences on the basis of a biophysical model. The method and its derivatives are free and easy accessible *via* a web portal and widely applied within the community. In the context of a long-standing cooperation, they had been used to identify an interferon regulatory factor 7 (IRF7)-driven inflammatory network (IDIN), with genes from this network being involved in the pathogenesis of type I diabetes in humans. This work has been published in Nature in 2010.

Another project of the Department of Computational Molecular Biology that will now be continued by Ho-Ryun Chung in his own Max Planck Research Group concerns the information encoded by histone modifications. In a paper published in Proc Natl Acad Sci USA in 2010, Chung and Vingron could show a strong correlation between histone modification levels and gene expression. Moreover, they could show that only a small number of histone modifications are necessary to accurately predict gene expression and that the connections between histone modifications and gene expression seem to be general feature for all types of cells.

Stefan Mundlos and his team are interested in the identification and characterization of skeletal defects on a genetic, molecular and developmental level. One focus is on the identification of genetic factors involved in monogenic diseases. Using whole-exome capture and SOLiD sequencing in combination with an HMM algorithm, they have been able to identify *PIGV* mutations in patients with Hyperphosphatasia mental retardation (HPMR) syndrome, also known as Mabry syndrome. Their results, published in Nature Genetics in 2010, open up the way for a streamlined gene discovery in future exome sequencing projects.

In addition to the research performed by the departments and independent research groups, many service groups of the MPIMG pursue their own research projects and have their own long-standing cooperation with national and international groups. Thorsten Mielke, e.g., who heads the institute's cryo-EM group, is very well established within the Berlin-Brandenburg scientific community and performs the entire cryo-EM work for the groups of Christian Spahn, Charité, and Roland Beckmann, LMU Munich. Amongst others, these successful cooperations resulted in five publications in Science and Nature between 2009 and 07/2012.

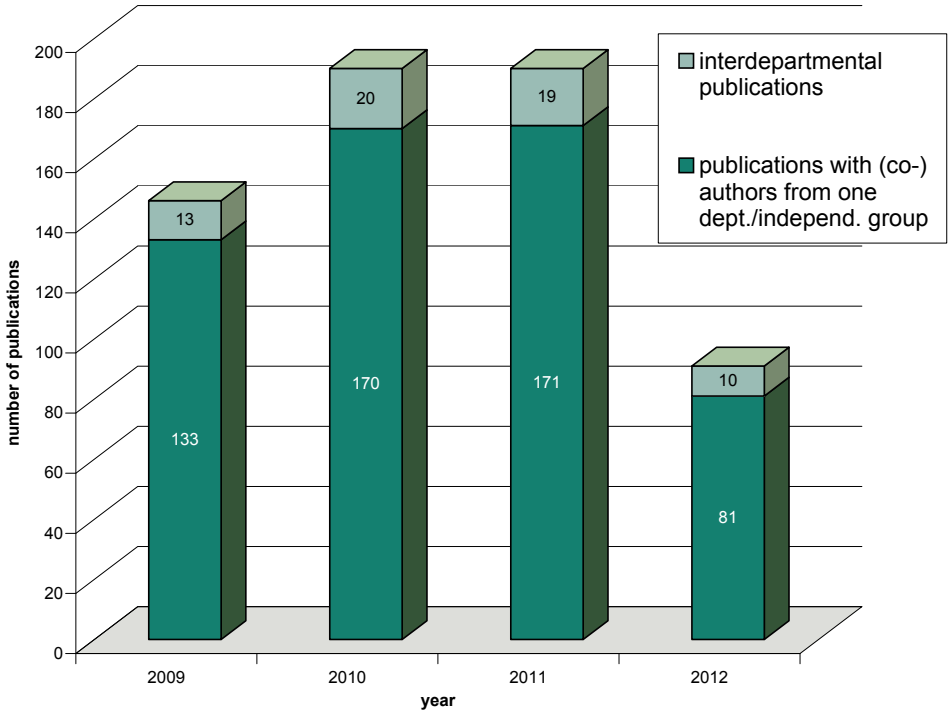


Table 1: Publications of the MPIMG. Publications with contributions from more than one MPIMG department / independent research group are shown in light green.

In all departments, the project groups interact very actively and many publications of each department include contributions from several researchers/ project groups. In addition, a number of interdepartmental interactions have also been established in recent years, resulting in 62 published papers, where at least one co-author was from another MPIMG department or independent group (see Table 1).
To summarize the publication activities, MPIMG researchers have (co-)authored 617 publications during the reporting period (2009-07/2012). 163 of these had been in journals with impact factors between 5 and 10, including Bioinformatics, Nucl Acids Res, PLoS Genetics and Proc Natl Acad Sci USA. 73 publications have been published in journals with impact factors between 10 and 30, including Am J Hum Genet, Genome Research, Mol Cell and Nature Biotechnol. 28 manuscripts have been published by Science, Cell, Nature Genetics and Nature (see Table 2).

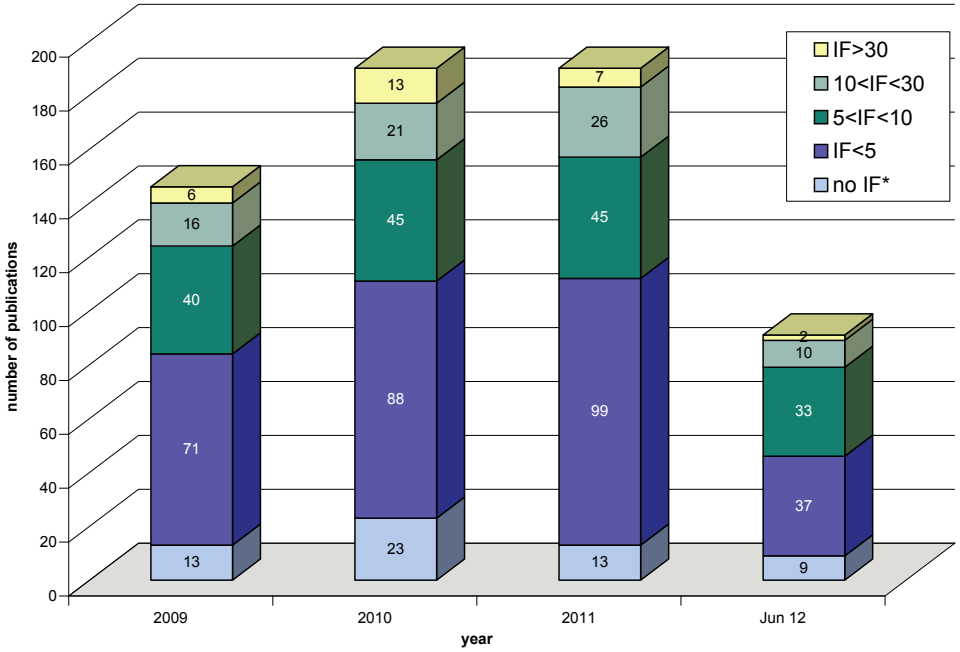


Table 2: MPIMG publications according Journal Citation Index (Thompson Reuters).
* Book contributions and publications in new journals (peer reviewed, but (still) without impact factor).

Scientific awards and appointments

In appreciation of their scientific achievements, members of the MPIMG have been awarded with several prizes. Among others,

- *Ulf Ørom* obtained a Sofja Kovelevskaya Award by the Alexander von Humboldt Foundation (2012);
- *Martin Vingron* has been selected as an ISCB Fellow in the Fellows Class of 2012;
- *Markus Ralser* received an ERC European Starting grant (2011), a Wellcome-Beit Prize of the Wellcome Trust, UK (2011) and an EMBO Young Investigator Award (2012);
- *Nana-Maria Grüning* has been awarded with the Nachwuchswissenschaftlerinnen-Preis 2012 of the Forschungsverbund Berlin;
- *Irina Czogiel* won the Gustav-Adolf-Lienert prize of the Biometrical Society (German Region) in 2012;
- *Marcel Holger Schulz* has been awarded an Otto Hahn Medal of the Max Planck Society (2011)
- *Rosa Karlic* has been honoured with the L'Oreal Adria-UNESCO National Fellowship "For Women in Science" (2010);
- *Lars Bertram* obtained a *S*pecial Award of the Hans-und-Ilse-Breuer Foundation for Research in Alzheimer's in 2010 and has been recipient of the 2009 Independent Investigator Award, NARSAD;
- *Eva Klopocki* recieved a Finalist Trainee Award of the American Society of Human Genetics, an ESHG Young Scientist Award of the European Society of Human Genetics, and the Vortragspreis of the Deutsche Gesellschaft für Humangenetik (German Society for Human Genetics), all in 2009.

Several people have left the MPIMG and took up positions in other institutions or universities in Germany and abroad. Some of the most important appointments have been

- *Silke Sperling* on a Heisenberg professorship (W3) for Cardiovascular Genetics at the Charité - Universitätsmedizin Berlin (2011)
- *James Adjaye* on a W3 professorship for Stem Cell Research and Regenerative Medicine at Heinrich Heine University, Düsseldorf (2012)
- *Tim Hucho* on a W2 professorship on Anaesthesiology and Pain Research at the University Hospital of Cologne (2012)
- *Andreas Kuss* on a W2 professorship on Molecular Human Genetics at the University of Greifswald (2010)
- *Ho-Ryun Chung* as head of the Max Planck Research Group "Computational Epigenomics" at the MPIMG (2011)
- *Markus Ralser* as principal investigator at the Cambridge Systems Biology Centre at the University of Cambridge, UK (2011)

Due to the downsizing of two departments, only a few new group leaders have been recruited at the MPIMG. In 2012, Ho-Ryun Chung and Ulf Ørom joined the Otto Warburg Laboratory, while David Meierhofer has been recruited as head of the new mass spectrometry service group.

External funding

A substantial part of the institute's research budget is obtained as grant money from external sources, such as the German Ministry of Education and Research, the European Union, the German Research Foundation (DFG) and others. Due to the upcoming retirement of two directors, the amount has decreased from about 80% in 2010 to 60% in 2012. About 40% of the money derives from the Federal Ministry of Education and Research as support e.g. for the International 1,000 Genome Project, a MHC Haplotype sequencing project and several other projects in the context of the National Genome Research Network plus. The second major source of grant money is the European Commission with the institute participating in several projects of the 7th Framework Health Program like IT Future of Medicine, GENCODYS, OncoTrack, BLUEPRINT and others.
Other prominent projects include the institute's participation in DFG Collaborative Research Centres ("Sonderforschungsbereiche") and grants from several foundations. With its involvement in national and international research projects and by virtue of the research output of the institute, the MPIMG is perceived internationally as a stronghold of genome and genetics research in Germany.

Material resources, equipment and spatial arrangements

As already mentioned before, the genomic approach to biology is a unifying feature for all MPIMG's research groups. The institute therefore houses a range of large-scale equipment like sequencing systems, mass spectrometers, transmission electron or laser scanning microscopes, and a large IT infrastructure. Most of it is maintained by the institute's service groups who operate the equipment and provide high-level support for all in house-scientists and many external collaborators. Details can be found in the reports of the single service groups. Some equipment, for example a FACS Cell Sorting System and some sequencing systems, is also maintained by single departments with special need for it. In addition to their own requirements, department members usually provide access for scientists from other groups and support them in handling the respective equipment.

Past, present, future

The MPIMG was founded in 1964 with the appointments of Heinz-Günther Wittmann and Heinz Schuster as heads of department, followed by the appointment of Thomas Trautner in 1965. At this time, the research of the institute was focussing on DNA replication and gene regulation in bacteria, bacterial phage and fungi (departments Schuster and Trautner) and on the structure, function and evolution of ribosomes, which were central to the work of H.-G. Wittmann. In 1970, the three departments, as well as four independent junior research groups (the future Otto Warburg Laboratory) moved into the new premises of the institute situated in Ihnestraße in Berlin-Dahlem. After the sudden death of H.-G. Wittmann in 1990 and the retirement of H. Schuster in 1995, Max Planck Society decided to reorient the institute towards the new field of genome research and its impact on human genetics. This has been realised by the appointments of Hans Lehrach (1994, Dept. of Vertebrate Genomics) and Hans-Hilger Ropers (Dept. of Human Molecular Genetics, full-time since 1997), followed by Martin Vingron (2000, Computational Molecular Biology), and Bernhard Herrmann (2003, Developmental Genetics). Stefan Mundlos, Professor of Genetics at Charité - Universitätsmedizin Berlin was jointly appointed as head of an independent research unit at the MPIMG (2000, Development & Disease).

Two of the directors, Hilger Ropers and Hans Lehrach, will retire in 2014. Thus, those two departments are undergoing a process of downsizing, and identification of new directors is underway. At the same time, the ensemble of buildings has been extended and now comprises the soon to be finished “tower 3”, which houses office space, seminar rooms and a large server room. With the subsequent renovation of the oldest parts of the institute, i.e. towers 1 and 2, we will soon be equipped to accommodate two new departments to replace the outgoing ones. With the upcoming personnel changes, a certain reorientation of the institute’s research landscape will take place. The remaining institute directors together with

Figure 2: Front detail of the new building (tower 3) of the MPIMG.

the pertinent committees of Max Planck Society have drafted an update to the research agenda. While genome research is continuously increasing its impact on the way genetics is done, the sequencing of complete genomes and the determination of the genes has become largely a routine task. The emphasis has shifted towards the interactions among genes, embodied in the regulatory networks, which genes form in order to accomplish tasks like the differentiation of cells or the reaction of a cell to a stimulus. Thus, in the future, the scientific focus of the MPIMG shall shift towards the study of gene regulatory networks and their role in development and in disease processes. In this effort, the genomic approach to biology shall remain a unifying feature for the activities of all institute researchers.

At the same time, the institute is reshaping its structure in terms of research groups and scientific service units. In the past, many aspects of scientific service provision had been taken care of within departments. In addition, Richard Reinhardt headed a large scientific service that provided access not only to DNA sequencing technology, but also to numerous other scientific services. As Reinhardt left the MPIMG in 2010 to build up a new sequencing facility at the Max Planck Institute for Plant Breeding Research in Cologne, the MPIMG reorganized the scientific service. Several new units have been created outside the departments and with a mission to offer service to the entire institute. The institute has newly established a sequencing unit (head: Bernd Timmermann) and a mass spectrometry unit (head: David Meierhofer). The IT group is now jointly headed by Donald Buczek and Peter Marquardt and the leadership of the imaging unit is currently being handed over from Rudi Lurz to Thorsten Mielke. In close collaboration with the animal house, headed by Ludger Hartmann, a transgenic facility has been built up, run by Lars Wittler. This restructuring has also been done in preparation for the arrival of the new departments. At the same time, each service unit is under continuous scrutiny and as technologies change, this design may also change in the future.

PhD students at the MPIMG

In summer 2012, approximately 80 doctoral students pursue their research at MPIMG. All students are integrated into a department or independent research group. In addition, the institute offers a range of cross-departmental activities to support the individual development of each PhD student. In November 2008, a structured PhD program has been established, consisting of four one-week subject courses (genome research, developmental genetics, human genetics, and bioinformatics), which are taught regularly, so as to have each PhD student take all four courses. The scientific courses are complemented by a range of soft skill courses open to all students and dealing with matters like scientific writing, presentation techniques, or rhetoric. Until July 2012, 13 scientific one-week courses have been offered with altogether 219 participants (though, of course, most students took part in more than one course). The response of the students is very positive; with many of them pointing out that they are getting not only an overview of the main scientific fields of the MPIMG, but also becoming acquainted with each other and thus building personal links between the different departments.

In addition to the support given by their supervisors at the MPIMG, many students wish for closer links to the local universities. To help the students establish such contacts early during their thesis work, the Board of Directors has asked all heads of research groups in 2010 to establish individual PhD committees for their students. The PhD committees consist not only of scientists from the MPIMG,

but also of colleagues from University; they supervise and support the student's efforts on a regular basis.
In 2006, the International Max Planck Research School for Computational Biology and Scientific Computing (IMPRS-CBSC) was founded together with the Bioinformatics Program of the Freie Universität Berlin. Faculty members of the IMPRS on the part of MPIMG are Ho-Ryun Chung, Ralf Herwig, Peter Arndt, and Martin Vingron. The IMPRS hosts around 22 students including the ones at the university. The institute is further involved in the Berlin School for Regenerative Therapies, housed at Humboldt University, and in teaching of medical students in several curricula.
Since April 2001, the Student Association (STA), formed by the students themselves, acts for the interests of all students at the MPIMG by addressing the directors, group leaders, or the staff association. In addition, it provides a platform for scientific and social interactions like sports or parties. The STA also organizes an annual PhD retreat, funded by the institute, to foster interdepartmental scientific discussion and exchange of technical and scientific knowledge between the students.

Figure 3: International students at the MPIMG.

Equal opportunities - work-life balance

The MPIMG takes an active role in supporting the institute's members who have children. The institute cooperates closely with a day-care centre in the neighbourhood and grants financial support to keep spare place for the children of MPIMG staff as permitted by the funding provisions. To assist parents who need to take their children to work awhile, the institute offers a room dedicated to children for some years. This will be continued and optimized in the newly built tower 3. Flexible working hours and teleworking is supported. According to the German law on temporary employment contracts in the scientific field (Wissenschaftszeitvertragsgesetz), temporary contracts during the so-called qualification period (Qualifizierungsphase) are extended automatically to make up for times of parental leave. Furthermore, the institute is making efforts to offer this to externally funded employees, too, whenever possible. Since 2010, the funding bodies

Figure 4: MPIMG staff at the annual summer party.

also permit supporting fellowship holders by several means such as additional payment for children, prolongation of fellowships and part-time fellowships. The "Besser betreut" service, provided by the Max Planck Society to all institutes, offers a broad range of services from child- and elderly care to housekeeping. The Christiane Nüsslein Volhard Foundation supports talented young female scientists with financial grants to pay for assistance in household chores and for additional childcare. In 2009, the Max Planck Society was awarded the certificate "Beruf und Familie" for the second time for its family-friendly human resources policy. The re-awarding is being sought for 2012.

Public relations work

Genome research, especially with a clear focus on human genetics and molecular mechanisms of human diseases, is not only for scientists, but also a matter of general interest. The MPIMG feels obliged to make the institute's research transparent and to discuss its implications – the positive as well as the negative ones - with the public. For these reasons, a set of different communication tools has been established:

- The institute discloses press releases about different themes to the press and publishes them at the MPIMG's website, too. Since January 2009, 50 press releases on scientific findings of various research groups, awards, and events have been published, most of them in close collaboration with the press office of the Max Planck Headquarter in Munich.
- The MPIMG has established a visiting program for school children. Under this program, the institute invites classes to visit a lab, discuss with the scientists and perform simple experiments like DNA isolation, cell staining, or microscopy on their own. Since 2009, 39 school classes with altogether 941 school children and teachers have visited the MPIMG.
- Since 2002, the MPIMG participates in the "Lange Nacht der Wissenschaften" (Long Night of Sciences). At this event, about 70-80 universities and research

institutions all over Berlin and Potsdam open their doors for one night and invite the general public to visit their labs, learn about the work that is done here and discuss it with the scientists. During the reporting period, the MPIMG participated in the Long Nights 2009, 2010, and 2012, with about 100-120 staff members being involved and between 800 to 1,000 visitors at the institute each year.

- Also once a year, the institute participates in the Girls' Day – Future Prospects for Girls, a large nationwide campaign, in which a wide range of professions and activities is presented to girls of 10 years upwards. Each year, a selected group of girls is invited to visit the institute for a whole day and trying out themselves at areas that are presently not typically female. Since 2006, between four to eight girls per annum have joined the electro mechanics workshops and the IT group of the institute and gained own experiences by handling CNC (computerized numerical control) machines, soldering irons, or computer hard discs.

In addition to these regular activities, MPIMG scientists give interviews to the press and are engaged in several public events like public lectures, exhibitions etc.

Figure 5: MPIMG member with school children at the exhibition "WeltWissen" at the Martin-Gropius-Bau in Berlin (09/10-01/11).

Department of Developmental Genetics

(Established: 11/2003)

Head
Prof. Dr. Bernhard G. Herrmann
Phone: +49 (0)30 8413-1409
Fax: +49 (0)30 8413-1229
Email: herrmann@molgen.mpg.de

Secretary
Christel Siegers (since 02/09, part time)
Phone: +49 (0)30 8413-1344
Fax: +49 (0)30 8413-1130
Email: siegers@molgen.mpg.de

Scientific assistant
Marta Caparros Rodriguez (since 02/06, part time)
Phone: +49 (0)30 8413-1344
Fax: +49 (0)30 8413-1130
Email: caparros@molgen.mpg.de

Senior scientists
Lars Wittler* (since 04/09)
Phillip Grote* (since 11/08)
Yves Charron* (since 05/06)
Heinrich Schrewe* (since 09/04)
Hermann Bauer (since 10/03)
Markus Morkel* (06/05-10/11)
Martin Werber* (01/05-04/11)

Scientists
Frederic Koch* (since 11/11)
Eun-ha Shin (since 11/11)
Smita Sudheer (since 12/10)
Tracie Pennimpede* (since 09/10)
Pamela Riemer* (since 02/09)
Alexandra Farrall (06/09-10/12)
Benedikt Schwartz (11/11-06/12)
Mikhail Sukchev (03/10-02/12)
Michaela Mayer (02/10-05/11)
Marc Leushacke* (12/10-02/11)
Vladimir Mazurov* (02/09-12/10)
Arnold Schröder (11/09-12/10)
Nathalie Véron (01/09-12/09)
Ralf Spörle* (07/04-09/09)

PhD students
Lisette Lange (since 10/10)
Matthias Marks (since 10/10)
Arica Beisaw (since 06/10)
Karina Schöfisch (since 09/09)
Constanze Nandy (since 02/09)
Sabrina Schindler (since 08/08)
Daniele Sunaga* (10/10-03/12, part of PhD work)
Benedikt Schwartz (01/07-10/11)
Eun-ha Shin (10/06-10/11)
Joana Vidigal* (02/07-12/10)
Pedro Rocha* (01/07-12/10)
Marc Leushacke* (11/05-11/10)
Arnold Schröder (10/03-10/09)
Sebastiano Battaglia (07/08-04/09, part of PhD work)

* externally funded

Master students
Stefanie Wolter (since 05/12)
Anna Dorothea Senft (since 02/12)
Amulya Sreekumar (06/11-06/12)
Sören Reinke (04 11-03/12)
Marcel Gätjen* (11/10-10/11)
Judith Proske (11/09-08/10)

Technicians
Sandra Währisch (since 03/09)
Gaby Bläß* (since 02/09)
Barica Kusecek (since 09/08)
Karol Macura* (engineer, since 12/04)
Andrea König (since 09/04)
Manuela Scholze* (since 09/04)
Jürgen Willert (since 11/03)
Antje Brouwer-Lehmitz* (08/08-08/11)
Silvia Kietzmann* (01/05-05/11)
Larissa Mosch* (01/10-05/11)
Ursula Schulz* (10/08-02/11)
Barbara Kosiol* (09/07-12/10, part time)
Gerhild Lüder (07/08-06/09, part time)
Uta Marchfelder (10/05-03/09)

Introduction

The question of how a human being can develop from just a single cell has intrigued scientists for many decades. Research in the field of developmental biology has clarified many of the principle mechanisms of pattern formation, morphogenesis and organ development, and has elucidated the function of particular cell groups in, for instance, body axis or head formation. Through careful analyses following genetic and tissue manipulations the role of diverse genes in various embryonic processes has been revealed. However, there are still many gaps in our knowledge. In particular, we still do not fully understand how various inputs acting on a cell, either from neighboring cells, the extracellular matrix, or from somewhere in the organism are “computed” and converted into a phenotypic change. We do not know in detail, what changes occur in particular cells along their differentiation path from a stem cell to a specialized cell. And furthermore, we do not understand how the program for building a tissue or organ consisting of several cell types is imprinted within the stem cells giving rise to them. There is a need to tackle these questions, as understanding the regulatory

* externally funded

circuits of cells, and particularly stem cells, is crucial not only for understanding embryonic development, but also tissue homeostasis and repair. A fundamental understanding of these processes is therefore indispensable for furthering regenerative medicine, for achieving a longer healthy lifespan, and for the treatment of numerous diseases.

The major research interest of the Department of Developmental Genetics is focused on understanding the gene regulatory networks controlling trunk development from axial stem cells. We investigate genetic and epigenetic control mechanisms involved in trunk development by utilizing state-of-the-art technology, such as massively parallel sequencing. In addition, we investigate molecular mechanisms of tumourigenesis and tumour progression, with a strong focus on intestinal adenoma formation. The latter is characterized by an aberrant tissue regeneration process, which has been shown in humans to be caused predominantly by the constitutive up-regulation of WNT signaling, resulting in an excess of proliferating cells.

As a third line of research we investigate the molecular basis of "transmission ratio distortion" (TRD), a particular phenomenon of non-Mendelian inheritance discovered in the mouse. It allows a "selfish" genetic element to be transmitted at a high ratio from generation to generation, and is based on the complex interplay of several QTLs involved in the control of sperm motility.

The Department of Developmental Genetics is structured along the aforementioned three major lines of research. The scientists cooperate on various projects and report directly to the department head. Senior scientists may also act as principal investigators in various projects involving pre- and post-doctoral scientists. During the reporting period, Bernhard Herrmann acted as coordinator of the NGFNplus (National Genome Research Network plus) Consortium *Modifiers of Intestinal Tumour Formation and Progression*, funded by the German Ministry for Education and Research (BMBF), involving two subprojects led by Markus Morkel and by Bernhard Herrmann. We also contributed with one subproject led by Markus Morkel to the NGFNplus Consortium *Mutanom*. Bernhard Herrmann served as speaker of the national consortium *CURE-Net: Network for Congenital Uro-rectal Malformations,* funded by the BMBF in the framework of the "Rare Diseases" Program, involving a subproject led by Lars Wittler. The department also contributed to the *TREAT20 Consortium "Tumour REsearch And Treatment - 20 Patient Pilot",* funded by the BMBF. Heinrich Schrewe was subproject leader in the Marie Curie Research Training Network Consortium *NucSys (Systems biology of nuclear receptors: A nutrigenomic approach to aging-related diseases)*. In addition, Bernhard Herrmann received funding from the German Research Society (Deutsche Forschungsgemeinschaft) for the long-term project *Genetic and molecular dissection of signaling pathways causing non-mendelian inheritance in the mouse*.

Scientific methods and findings

Genetic and epigenetic control of trunk development in the mouse

The differentiation of specialized cell types from stem cells is brought about by differential gene expression, where transcription factors controlling lineage decisions and differentiation are expressed in a tissue-specific manner. Therefore, knowledge of the transcriptomes of stem cells and their descendants is key to unraveling the gene regulatory networks (GRN) controlling lineage restriction and differentiation. The identification of key regulators requires the functional analysis of candidate genes *in vivo* using gene targeting or RNAi technology, and is guided by the specificity of expression of the candidate control factor. For GRN reconstruction the target genes of essential regulators need to be determined. This requires comparisons between wild-type and loss-of-function mutants on the transcriptome level, as well as knowledge of the genomic binding sites of the regulatory protein. To obtain a more global view of the molecular mechanisms controlling lineage restriction and differentiation processes, the epigenetic landscape needs to be included in the analysis. In recent years, we have invested in the development of tools required for GRN reconstruction, and have made important steps towards these goals.
Herein I would like to mention some recent achievements. We have started a comprehensive expression analysis for mid-gestation mouse embryos and identified key players of trunk development, among them long non-coding RNA genes. We show for one lncRNA that it acts as an essential regulatory RNA in the lateral mesoderm; it is the first of its kind known in the mouse. We have also developed a vector system for conditional knock-down of genes *in vivo* using shRNAmir-mediated RNAi. This system was used to generate a mouse model resembling the caudal regression syndrome in man. We have also shown that MED12 is an essential transcriptional co-factor for the WNT signaling pathway during trunk development.

Detailed analysis of gene expression in mid-gestational mouse embryos

The identity of a cell is determined by its transcriptome and proteome, and cell differentiation is directly related to differential gene expression. Therefore, the knowledge of differential gene expression in a developing organism is fundamental in reconstructing GRNs controlling cell differentiation. Since we aim at understanding the GRNs controlling trunk development, we have focused our attention to the transcriptomes of mid-gestational embryos (roughly embryonic stages E8-E11).
We first used whole-mount *in situ* hybridization (WISH) techniques to screen thousands of cDNA clones derived from genes expressed at these stages to identify genes showing a restricted expression pattern in the embryo. From approximately 8,000 genes analyzed we identified close to 2,000 genes, which displayed differential expression. We described the expression patterns according to the EMAP TS15 ontology of anatomical structures, and stored images, sequence information and pattern descriptions in a public database (Molecular Anatomy of the Mouse Embryo Project (MAMEP) database: http://mamep.molgen.mpg.de/). This data complements other public expression data, provides an excellent collection of differentially expressed genes in mid-gestational mouse embryos,

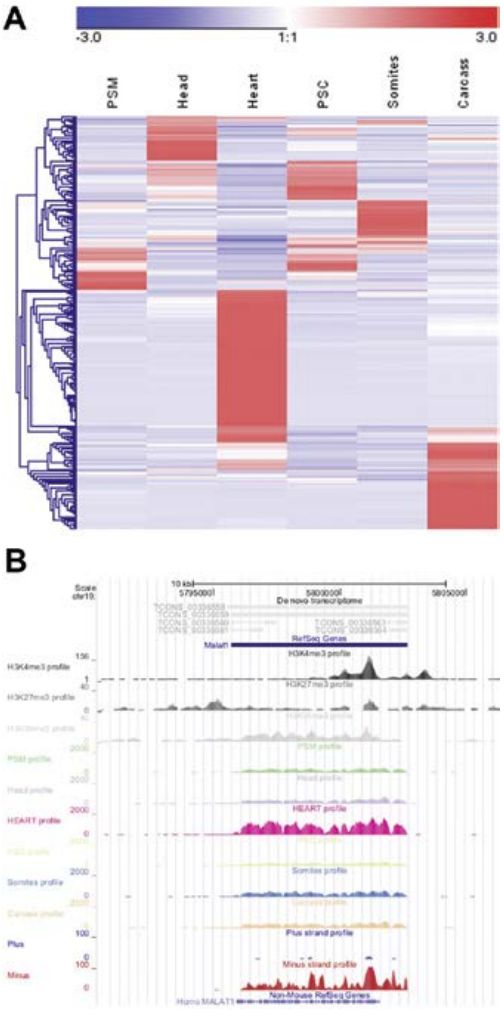


Figure 1: RNA-seq analysis of six tissues derived from TS12 mouse embryos reveals a large number of differentially expressed noncoding RNAs. A) Hierarchical clustering of 296 differentially expressed noncoding RNAs. B) Genome browser screenshot of a noncoding RNA gene, Malat1, showing enhanced expression in the heart tissue. Histone modification data shows enrichment at transcription start site (H3K4me3 and H3K27me3) and along the transcribed region (H3K36me3).

and identified a large number of important regulators of differentiation processes in the embryo.

To gain more depth in the expression analysis, that is to say to obtain a nearly complete picture of all transcripts, coding and non-coding, expressed in mid-gestational embryos, we have utilized massively parallel sequencing of cDNA derived from E8.25 (TS12) embryos. The embryos were dissected in six parts (head, heart, prospective spinal cord, caudal mesoderm, somites, and carcass), providing data about differential expression in these subregions. Several hundred million paired end reads (2 x 100b) were analyzed providing very deep insight into the transcriptome, splice variants, antisense transcripts, non-coding genes, or coding/non-coding gene pairs, and above all their differential expression in the embryo. This thorough analysis provides a reference dataset for the analysis of differentiation processes and GRNs in trunk development. One particularly interesting set of genes identified in this work is the class of long non-coding RNAs (lncRNA), in particular differentially expressed lncRNAs, which are involved in the epigenetic control of gene expression (Figure 1). This class of genes is just starting to be analyzed for their role in embryogenesis, and so far none has been functionally investigated in mouse, except for the one presented in the next section.

The role of the lncRNA Nrlm in mouse embryogenesis

In the transcriptome analysis of E8.25 mouse embryos we have identified a large number of differentially expressed lncRNAs. We isolated several by cDNA cloning and identified their expression pattern in mid-gestation embryos by WISH analysis. One of them showed a restricted expression in the nascent lateral mesoderm, another one was confined to the embryonic heart. We started a functional analysis of the first, which we finally named *Nrlm* (Non-coding regulatory RNA in lateral mesoderm). We targeted both alleles of *Nrlm* in ES cells and analyzed the outcome of the *Nrlm* loss-of function mutation. The result was quite striking: embryos died at around 13.5 days of development from heart failure and showed severe omphalocele (protrusion of the liver and umbilical vessels through the ventral body wall) (Figure 2). Both the heart and the body wall derive from the

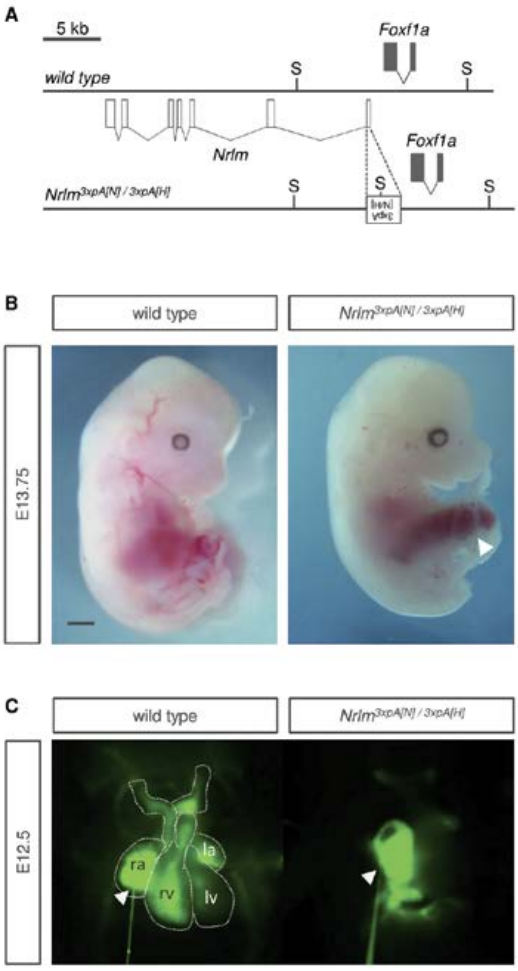


Figure 2: The lncRNA Nrlm is essential for proper heart and body wall development. A) Schematic showing the wild-type (upper) and mutant (lower) Nrlm-Foxf1a genomic region; the first exon of Nrlm was replaced with a 3xpA stop cassette by gene targeting in order to prevent transcription of the gene. B) E13.75 embryos derived by tetraploid complementation from wild-type ES cells or homozygous Nrlm mutants. In mutant embryos the ventral body wall is ruptured by umbilical vessels and parts of the liver (white arrowhead); the tail and limbs were removed. Scale bar: 1 mm. C) Fluorescent imaging of E12.5 embryo hearts. FluoSpheres® were injected into the right atrium and their dispersion through the heart and vasculature was visualized. The image shown is at 75 seconds following injection. In the wild type embryo, FluoSpheres® dispersed throughout the heart and arteries (n=3), while in the mutant the microspheres accumulated in the right atrium (n=5) indicating heart dysfunction (ra= right atrium, rv= right ventricle, lv= left ventricle, la= left atrium). White arrowheads indicate injection needles. The figure was adjusted from Grote et al. (under review).

lateral mesoderm, in which *Nrlm* is transiently expressed. Thus, the *Nrlm* transcript is essential for proper differentiation of lateral mesoderm. We followed up the molecular basis for the phenotype and were able to show that *Nrlm* interacts with the Polycomb Repressive Complex 2 (PRC2) and the Trithorax Group/MLL complex (TrxG/MLL), important protein complexes involved in epigenetic gene control, which set repressive or activating chromatin marks, respectively, on regulatory regions such as promoters or enhancers of genes. We found that *Nrlm* promotes the occupancy of PRC2 on the promoters of three transcription factors playing important roles in lateral mesoderm differentiation, *Foxf1a*, *Irx3* and *Pitx2*. In the wild-type embryo, *Nrlm* promotes the repressive histone H3 Lysine 27 trimethylation mark, causing reduced expression of the three regulators as compared to embryos lacking *Nrlm*. Apparently, this lower expression level is essential for proper lateral mesoderm differentiation. In the *Nrlm* mutant, several more transcription factors involved in lateral mesoderm, heart development or heart function are up-regulated as well, probably in a PRC2 independent manner. In summary, loss of *Nrlm* appears to cause disequilibrium of the GRN controlling lateral mesoderm differentiation resulting in embryonic death.

A vector system for conditional knock-down of genes in vivo

The analysis of GRNs requires the functional analysis of many regulatory factors. However, in general gene targeting in the mouse is a slow process. In contrast, RNA interference (RNAi) offered a faster gene analysis tool. However, though this approach has been utilized quite successfully in cell culture systems, *in vivo* RNAi has been hampered by the lack of appropriate vector systems. Consequently, only few successful gene knock-downs in mouse embryos have been reported. In order to overcome previous limitations we have developed an inducible RNA

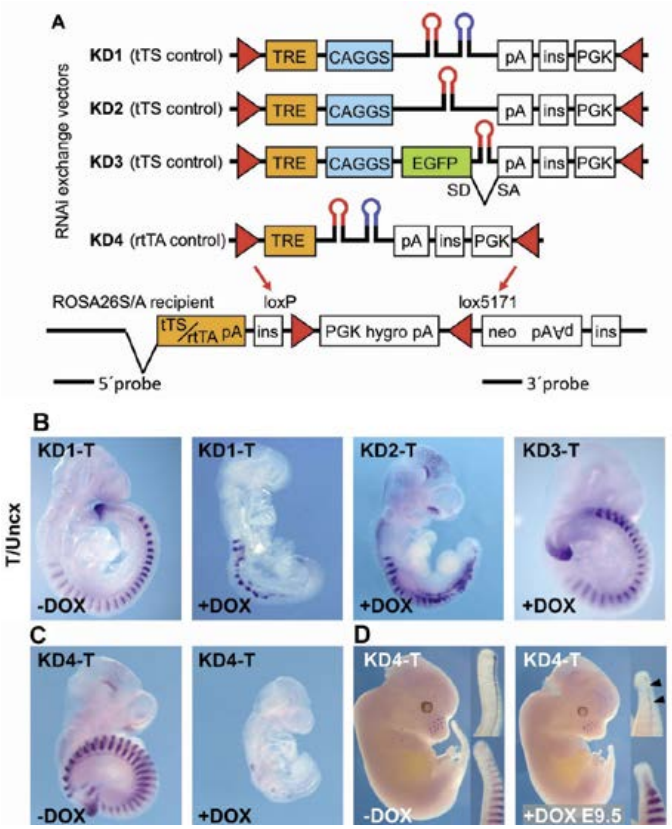


Figure 3: Vector system for inducible RNAi in mouse embryos. A) Schematic representation of ROSA26A/ROSA26S recipient loci and KD1 to KD4 exchange vectors used for RNAi. The recipient locus contains either the tTS (for KD1- KD3) or the rtTA (for KD4) gene under control of the ROSA26 promoter; RNAi exchange vectors are integrated directionally by RMCE using Cre recombinase. Different shRNAmir hairpins are indicated in red and blue. Abbreviations: TRE: tTS/rtTA responsive element; SD, SA: splice donor and acceptor sites; ins: chicken globin insulator. B – D) Analysis of embryonic phenotypes obtained by RNAi. RNAi constructs are indicated at the top, induction schemes at the bottom of each panel. B) C) E9.5-10.0 transgenic control embryos (-Dox) or embryos induced for shRNAmir expression (+Dox) targeting T (KD1-T to KD3-T) gene transcripts were hybridized *in situ* with T plus Uncx specific probes. D) E12.5 control (left) and mutant embryo (right) induced for RNAi at E9.5 using the rtTA controlled KD4-T construct. Inlays show Shh (upper) or Uncx (lower) expression detected by *in situ* hybridization. Arrowheads mark discontinuous Shh expression. The figure was adjusted from Vidigal et al. (2010).

interference system, which is suitable for the functional dissection of mouse embryogenesis (Figure 3A). The key features are: 1) temporal control; allows to overcome early lethality and to assess the genetic integrity of the transgenic ES cells used in the experiments; 2) single copy integration of the transgene construct into a defined recipient locus; avoids variability caused by the integration site; 3) a series of RNAi vectors allows generation of mutants of different strength (hypomorphs; loss-of-function alleles); 4) the use of PolII promoters avoids massive over-expression typical for previously utilized vectors, thus reducing or eliminating off-target effects.

We generated four RNAi vectors and used the system to knock-down *Brachyury* (*T*); three mutant phenotypes varying in strength from a complete loss-of-function to a mild hypomorphic phenotype were obtained (Figure 3B). In addition, we knocked down *Foxa2* and *Noto* in order to prove general applicability of the system. We found no hints for off-target effects in the *T* knock-down alleles. Thus the system is suitable for investigating the function of genes playing multiple roles in development in more detail than was previously possible with knock-out alleles. In conclusion, we have generated an excellent tool for quick assessments and detailed analysis of gene function in mouse embryogenesis.

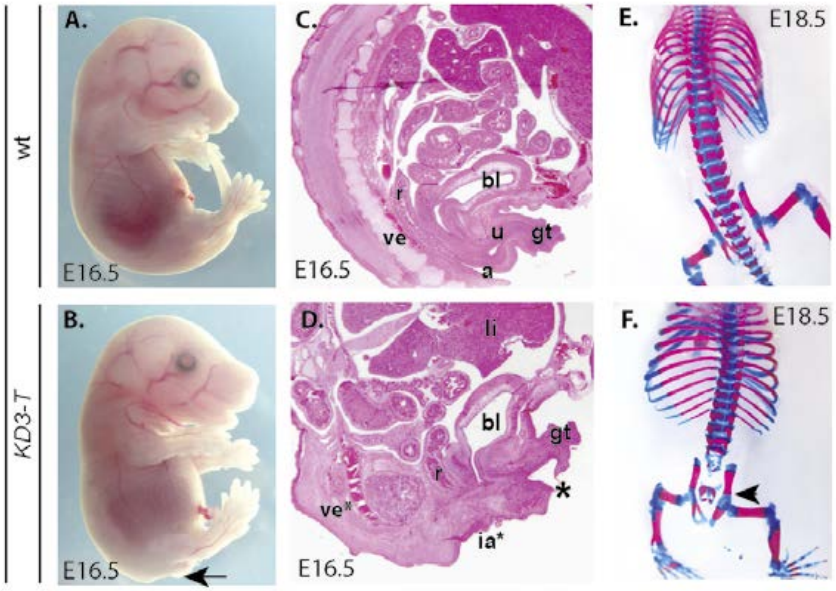


Figure 4: Phenotype of KD3-T embryos. A), B) Gross morphology of KD3-T embryos at E16.5 shows normal outer morphology with the exception of a caudal truncation (arrow) accompanied by spina bifida occulta (not shown). C), D) Mid-saggital sections through the caudal region of the embryos shown in A and B reveal urorectal malformations in KD3-T embryos, including hypospadias (asterisk) and imperforate anus (ia*), along with absence of caudal vertebrae (ve*). E), F) Skeletal analysis at E18.5 shows loss of vertebrae below the lumbar/sacral level. Taken together, the combination of urorectal malformations and caudal defects in KD3-T resemble phenotypes seen in caudal regression syndrome (OMIM #600145; Duesterhoeft et al. (2007) Am J Med Genet Part A 143A:3175–3184). The figure was adjusted from Pennimpede et al. (under revision)

A knock-down mutant of Brachyury causing axial skeletal defects accompanied by uro-rectal malformations is a model for the caudal regression syndrome in man

We have used the shRNAmir-based knock-down system to generate a mild hypomorphic *Brachyury* phenotype characterized by frequent survival of the embryos to birth, accompanied by tail loss and uro-rectal malformations (*KD3-T*, see Figure 3A). Morphological and histological examination of E18.5 embryos revealed varying degrees of axial skeletal defects primarily in the lumbar and sacral region (Figure 4). These were accompanied by malformations of the urorectal region including imperforate anus and anal stenosis, and also rare cases of hypospadias. The defects were traced back specifically to loss of the notochord in the lumbosacral region, which caused apoptotic degradation of neighboring somites and vertebral malformations. This led to disorganization of the mesodermal structures forming the urorectal system, ultimately resulting in a uro-recto-caudal phenotype which resembles the human caudal regression syndrome. The knock-down had no strong effect on two other tissues, which have been previously shown to be dependent on *Brachyury* function: the tailbud and the allantois. Thus, the RNAi vector system allowed for an examination of the role of *T* selectively in one particular embryonic structure. This has not been possible previously with loss-of-function alleles. In conclusion, we have utilized *in vivo* RNAi to generate a hypomorphic *Brachyury* phenotype, which serves as model for a human disease phenotype.

Med12 is essential for canonical Wnt and Wnt/PCP signaling

The Mediator complex is commonly seen as a molecular bridge that connects DNA-bound transcription factors to the RNA polymerase II (Pol II) machinery. It is a large complex of 30 subunits that is present in all eukaryotes. The MED12 subunit has been implicated not only in the regulation of Pol II activity, but also in the binding of transcription factors to the bulk of the Mediator complex. We targeted *Med12* in mouse embryonic stem cells to investigate the *in vivo* function of this

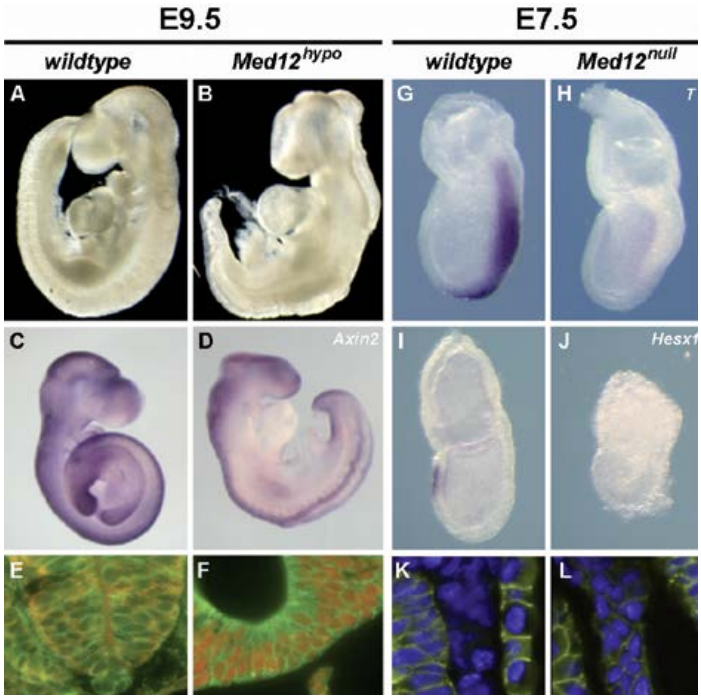


Figure 5: Med12 is critical for early mouse development. A), B) Med12 hypomorphic embryos have neural tube closure, somitogenesis and axis elongation defect, and die at E10.5 from cardiovascular defects. C), D) Marker gene expression analyses indicate that target genes of Wnt/β-catenin signalling, e.g. Axin2, are downregulated in Med12hypo embryos. Low levels of Med12 lead to mislocalization of Prickle-1, a core PCP component, in the neural tube (E, F) resulting in severe closure defects. Med12 deficient embryos die at early gastrulation and fail to establish an anterior-posterior axis, as shown for specific markers like T (G, H) and Hesx1 (I, J). Loss of Med12 disturbs WNT-signaling, and mesoderm formation is severely impaired. Epiblast cells are still able to ingress through the primitive streak, but fail to become mesenchymal, caused by maintained E-cadherin expression (K, L; E-cadherin in green, DAPI in blue). The figure was adjusted from Rocha et al. (2010).

subunit. *Med12* hypomorphic mutants, which have a drastic reduction in MED12 protein levels, failed to develop beyond embryonic day 10 and had severe defects in neural tube closure, axis elongation, somitogenesis and heart formation (Figure 5). We showed that in *Med12* hypomorphic embryos, the Wnt/planar cell polarity pathway is disrupted and that canonical Wnt/β-catenin signaling is impaired. In agreement with this, embryos that are incapable of *Med12* expression failed to establish the anterior visceral endoderm or activate *Brachyury* expression, and did not complete gastrulation.

Planned developments

In general, we are refining our tools and approaches to unraveling the gene regulatory networks controlling trunk development.

We are utilizing BAC recombineering technology to derive reporter transgenes for marking specific cell types in the embryonic trunk and for purifying them for molecular analysis. This provides tissue-specific resolution for the developmental processes we are analyzing.

We use lineage-tracing techniques to follow the fate of defined cell groups and their descendant. This allows back-tracking of the various differentiated cell types to their origins in the embryo anlage. We aim to identify and characterize the axial stem cells and all their early descendants in the trunk.

We characterize purified cell groups on the transcriptomic and epigenomic level using massively parallel sequencing.

We continue to genetically distort the GRNs controlling trunk development and to compare the molecular and phenotypic outcome between wild-type and mutant embryos. Complementary to this approach we try to establish tools for determining the genomic binding sites of regulatory factors *in vivo*. This will allow the derivation of regulator-target relationships, which are the core of GRNs.

Tools for elucidating the epigenetic control mechanisms guiding cells through induction, lineage selection and commitment are now available. Therefore, epigenetic processes receive stronger attention, and this includes the role and function of lncRNAs in these processes.

Mechanisms of intestinal tumour formation

Tissue homeostasis is controlled by developmental mechanisms, in which adult stem cells give rise to various differentiated cell types replacing dead cells. This process is balanced by signals acting on stem cells, such as the intestinal stem cells (ISC). Tumourigenesis is triggered by constitutive activation of signaling cascades leading to an imbalance between stem cell proliferation and cell differentiation producing an excess of proliferating cells. In intestinal tumours, this is caused primarily by constitutive activation of the WNT signaling pathway. We have taken several approaches to investigate the molecular mechanisms involved in intestinal tumourigenesis and tumour progression.

WNT and BMP signals control intestinal adenoma cell fates

Cellular hierarchies and signals that govern stemness and differentiation of intestinal adenoma cells are not well defined. We used organotypic culture to investigate the impact of β-catenin and BMP signals in cells that form intestinal adenoma in the mouse (Figure 6). We found that activation of β-catenin signaling by loss of APC or transgenic induction of oncogenic mutant β-catenin (Ctnnb1mut) initiates the conversion of untransformed intestinal cells to tumour cells. These tumour cells display cancer stem cell (CSC) traits such as increased expression of the CSC markers Cd133 and Cd44, a high capacity for self-renewal and unlimited proliferative potential. Subsequent inactivation of transgenic *Ctnnb1mut* results in the reversion of tumour cells to normal intestinal stem cells, which immediately reinstall the cellular hierarchy of the normal intestinal epithelium. We were able to demonstrate that oncogenic activation of β-catenin signaling

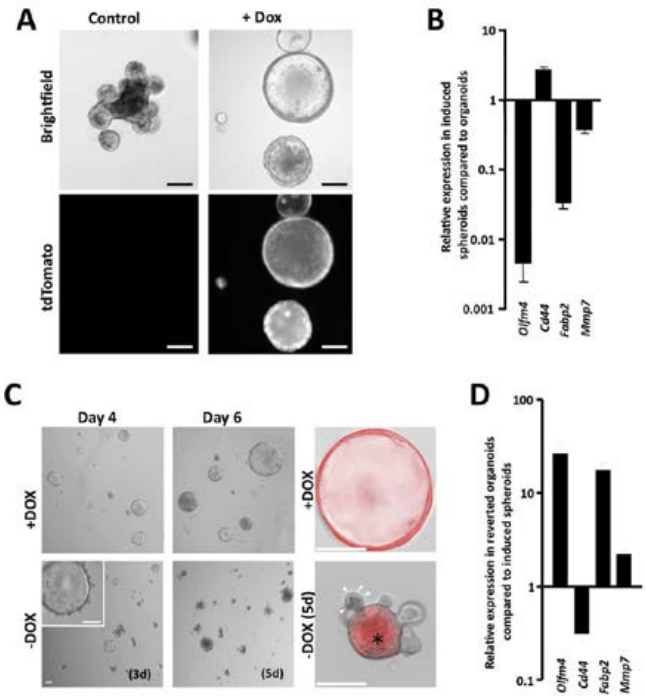


Figure 6: Conditional expression of stabilized β-catenin in intestinal organoids triggers reversible oncogenic transformation. A) Doxycyline-mediated (+Dox) induction of the transgene construct td-Tomato_T2A_Ctnnb1mut in organoid cultures leads to formation of spheroids. B), D) Analysis of characteristic intestinal marker genes in Dox-induced (B) and reversed (D) cultures obtained by qRT-PCR; error bars SEM. C) Reversion of spheroids to organoids upon Dox removal effecting down-regulation of the transgene construct. Arrows indicate Paneth cells of the crypt domain in a rescued organoid. Red fluorescence (tdTomato) observed in the lumen of reversed organoids indicates apoptotic debris derived from formerly induced cells. The figure was adjusted from Farrall et al. (2012).

initiates the early steps of intestinal cellular transformation in the absence of irreversible genetic or epigenetic changes. Interestingly, we found that tumour cells in culture and in adenoma produce BMP4, which counteracts CSC-like traits by initiating irreversible cellular differentiation and loss of self-renewal capacity. We conclude that the opposition of stemness-maintaining oncogenic β-catenin signals and autocrine differentiating BMP signals within the adenoma cell provides a rationale for the formation of cellular hierarchies in intestinal adenoma and may serve to limit adenoma growth.

DNA-methylome analysis of mouse intestinal adenoma identifies a tumour-specific signature that is partly conserved in human colon cancer

Aberrant CpG methylation is a universal epigenetic trait of cancer cell genomes. However, human cancer samples or cell lines preclude the investigation of epigenetic changes occurring early during tumour development. We have used the MeDIP-seq technique to analyze the DNA methylome of APC^{Min} adenoma as a model for intestinal cancer initiation, and identified more than 13,000 recurring differentially methylated regions (DMRs) characterizing intestinal adenoma of the mouse (Figure 7). We show that Polycomb Repressive Complex (PRC) targets are strongly enriched among hypermethylated DMRs, and several PRC2 components and DNA methyltransferases were up-regulated in adenoma. We further showed that the DMR signature arises *de novo* in adenoma cells rather than by expansion of a pre-existing pattern. Moreover, we found no significant correlation between differential promoter or gene body methylation and changes in gene expression upon adenoma formation, and epigenetic silencing of tumour suppressors was rare. Quite strikingly, we identified a core set of DMRs, which is conserved between mouse adenoma and human colon cancer, thus possibly revealing a global panel of epigenetically modified genes for intestinal tumours. Thus we were able to distinguish between conserved epigenetic alterations occurring early in intestinal adenoma and stochastic events promoting the progression of late stages of colon cancer. This data may facilitate the selection of more specific clinical epigenetic biomarkers.

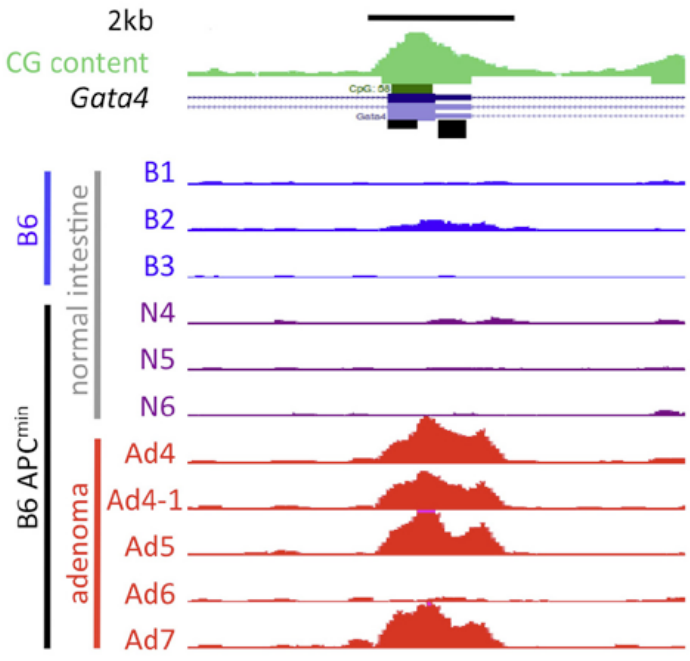


Figure 7: Example of a MeDIP-seq analysis revealing DNA methylation differences between normal intestine and tumor tissue (adenoma). A CpG island within Gata4 is hypermethylated in a majority of mouse intestinal adenomas. Green, CpG density; blue, purple, red: MeDIP-seq tracks of B6 mouse normal intestine, APCMin mouse normal intestine, and APCmin adenoma, respectively. The figure was adjusted from Grimm et al. (under review).

An RNA interference screen identifies a role of FGF signals in colon cancer progression

In tumour cells, stepwise oncogenic deregulation of signaling cascades induces alterations of cellular morphology and promotes the acquisition of malignant traits. We identified a set of 21 genes, including FGF9, as determinants of tumour cell morphology by an RNA interference phenotypic screen in SW480 colon cancer cells. Using a panel of small molecular inhibitors, we subsequently established phenotypic effects, downstream signaling cascades, and associated gene expression signatures of FGF receptor signals. We found that inhibition of FGF signals induces epithelial cell adhesion and loss of motility in colon cancer cells. These effects are mediated *via* the mitogen-activated protein kinase (MAPK) and Rho GTPase cascades. In agreement with these findings, inhibition of the MEK1/2 or JNK cascades, but not of the PI3K-AKT signaling axis also induced epithelial cell morphology. Finally, we found that expression of FGF9 was strong in a subset of advanced colon cancers, and overexpression negatively correlated with patients' survival. Our functional and expression analyses suggest that FGF receptor signals can contribute to colon cancer progression.

Transmission ratio distortion: elucidating the molecular-strategies of a selfish genetic element

The term transmission ratio distortion (TRD) "refers to instances, in which the two alleles carried by a heterozygote are transmitted unequally to the zygote at the time of fertilisation" (Lyon). TRD is caused by the *t*haplotype, a variant form of the centromer-close half of chromosome 17, and occurs only in males. Heterozygotes (*t*/+ males) can transmit the *t*haplotype chromosome almost exclusively to their offspring, quite different from the normal "Mendelian" transmission ratio (50:50). However, the variability in *t*haplotype transmission is very high and largely depends on several genetic variants, quantitative trait loci (QTL), called "distorter", which act additively to enhance the transmission ratio of the *t*haplotype. Thus, (experimentally obtained) partial *t*haplotypes, which express only one or two distorter genes, show a lower transmission ratio than the complete *t*haplotype, which contains 5-6 distorters. The distorters cannot achieve a high transmission ratio by themselves; they promote the transmission of a "selfish" gene called "responder" (*Tcr*). It sits in the middle of the 30-50 Mb *t*haplotype region and is flanked by distorter gene loci. Solely the responder is able to cause nonMendelian inheritance; the distorters only function as "helper". The massive overexpression of distorters, however, triggers a disease phenotype: male sterility.

TRD occurs in male spermatozoa competing for egg cells. TRD is possible because *t*sperm, spermatozoa carrying the *t*haplotype, swim faster and more directly towards the eggs. They use an ingenious trick to outcompete sperm carrying the wild-type chromosome 17 (+sperm). During spermatogenesis the *t*sperm cells express the distorter genes, whose products impair sperm cell motility, and the responder, which protects from this negative effect of the distorters. However, while the distorter products act in all sperm cells the responder acts only in *t*sperm. Later, in the fertilisation process, +sperm are impaired in their swimming behaviour by the action of the distorter gene products, while *t*sperm are rescued by the responder, and thus can swim normally and outcompete the +sperm.

In order to understand the mechanism leading to TRD on a molecular level, we have isolated and analyzed several key components, the responder and the dis-

torters *Tcd1a*, *Tcd2a* and *Tcd2b* (see below). We have shown that the latter code for regulators of Rho small G proteins, known to control cell motility of slow migrating cells such as fibroblasts. These are the GAP (GTPase activating protein) TAGAP1, a negative regulator of RHO, and the GEF (GDP/GTP exchange factor) FGD2, a positive regulator of RHO small G proteins. Both act as QTLs; they increase or reduce TRD of the *t*haplotype in a dosage and/or activity-dependent manner. The involvement of these factors in Rho signaling indicates that at least some important parameters of the motility of spermatozoa, which are specialized as fast moving cells, are controlled by the same signaling pathways as the motility of slow moving cells.
The distorters act in two pathways controlling the activity of the kinase SMOK (sperm motility kinase) in an antagonistic manner. SMOK is the wild-type form of TCR, which acts in a dominant negative manner (Figure 8). We have proposed that SMOK controls the parameters of sperm motility, which are impaired by the distorters. TAGAP1 inhibits a negative regulator of SMOK, while FGD2 activates a positive regulator of this kinase. The *t*alleles of both *Tagap1* and *Fgd2* are up-regulated. As a consequence, SMOK1 is hyperactivated by stronger activation through FGD2 and reduced inhibition through TAGAP1, and this leads to motility defects, which, in the extreme case, can result in male infertility. The dominant negative kinase TCR can counterbalance this negative effect exerted by the distorters and thus restore normal motility in *t*sperm, providing a selective advantage to *t*sperm resulting in TRD.
But how can TCR act exclusively in *t*sperm? Below we show that there is quite a simple explanation for this outstanding behavior, having a stunning effect.

Retention of gene products in syncytial spermatids promotes non-Mendelian inheritance as revealed by the t-complex-responder

The capability of *Tcr* to cause phenotypic differences between *t*- and +sperm derived from *t*/+ males contradicts the concept of phenotypic equivalence proposed for sperm cells, which develop in a syncytium and actively share gene products. By analyzing a *Tcr* minigene in hemizygous transgenic mice we showed that *Tcr* gene products are post-meiotically expressed and retained in the haploid sperm cells. The wild-type allele of *Tcr*, *Smok1*, behaves in the same manner suggesting that *Tcr*/*Smok* reveal a common mechanism prone to evolve non-Mendelian inheritance in mammals.
Below I briefly summarise the identification and characterization of a third distorter, *Tcd2b*, encoding the enzyme NME3.

The nucleoside diphosphate kinase gene Nme3 acts as quantitative trait locus promoting non-Mendelian inheritance

The reduction of the *Nme3* dosage by gene targeting of the wild type allele enhanced the transmission rate of the *t*haplotype and phenocopied distorter function. Genetic and biochemical analysis showed that the *t*allele of *Nme3* harbors a mutation (P89S), which compromises the enzymatic activity of the protein and genetically acts as a hypomorph. Transgenic overexpression of the *Nme3 t*allele reduced *t*haplotype transmission, proving it to be a distorter.
We propose that the NME3 protein interacts with RHO signaling cascades to impair sperm motility through hyperactivation of SMOK, the wild type form of the responder (Figure 8). This deleterious effect of the distorters is counter-balanced by the responder, $SMOK^{Tcr}$, which is exclusively expressed in *t*sperm, thus per-

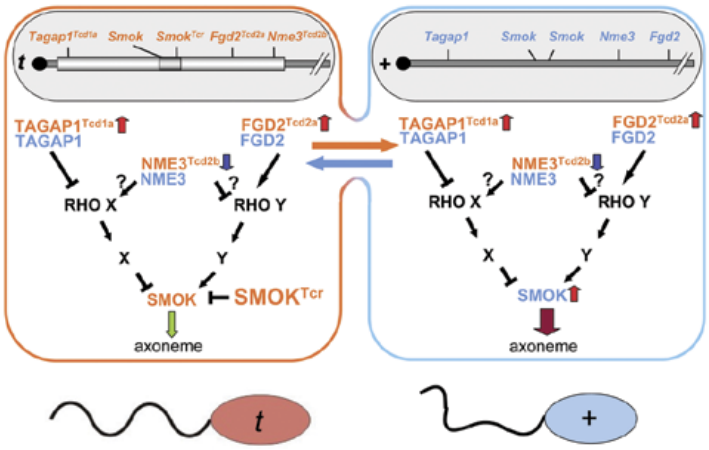


Figure 8: Model of the role of the distorters and the responder in *t*haplotype transmission ratio distortion. TagapTcd1a and Fgd2Tcd2a encode hypermorphic alleles of genes controlling Rho small G proteins involved in sperm motility control, Nme3Tcd2b encodes a hypomorphic allele of Nme3 acting on RHO. Wild-type and *t*alleles are expressed in all sperm derived from a t/+ male. NME3 may be an activator of the inhibitory Rho signaling pathway or an inhibitor of the activating pathway controlling SMOK activity, or both. NME3Tcd2b thus might synergize with TAGAP1Tcd1a to reduce inhibition and/or with FGD2Tcd2a to enhance activation of SMOK in all sperm. The combined activity of all distorters leads to impairment of sperm motility, which is rescued exclusively in *t*sperm by the responder encoding the dominant-negative variant SMOKTCR, resulting in transmission ratio distortion in favor of *t*sperm. Black arrows indicate activation, blocked lines inhibition; red upward pointing arrows indicate up-regulation, blue down-pointing arrows down-regulation; the green down-pointing arrow symbolizes rescued, the dark-red down-pointing arrow impaired flagellar motility. The figure was adjusted from Bauer et al. (2012).

mitting selfish behaviour and preferential transmission of the *t*haplotype. In addition, the previously reported association of NME family members with RHO signaling in somatic cell motility and metastasis, in conjunction with our data involving RHO signaling in sperm motility, suggests a functional conservation between mechanisms for motility control in somatic cells and spermatozoa.
We have isolated a fourth distorter candidate gene; experiments for proving it as distorter are in progress.

Perspectives

We are currently working in four directions:
We search the *Tcr* transcript for control elements, which are responsible for RNA tethering and translational control. Since all data have to be obtained *in vivo* using transgenic approaches, answering these questions is a long term and slow process.
We are in the process of expanding the signaling network involving the distorters and SMOK. This should reveal in more detail, how the antagonistic Rho pathways are interconnected, and what are the interaction partners and hopefully the downstream targets of SMOK. We have used Y-2-Hybrid screening in collaboration with Uli Stelzl, and have identified promising candidate genes.
We construct and test "mini-*t*haplotypes", which are efficiently causing TRD. In particular, we want to insert such constructs into the sex chromosomes in order to achieve sex ratio distortion. This is not trivial since the sex chromosomes are silenced in spermiogenesis, when *Tcr* needs to be activated. Thus, we need to identify a suitable site for transgene integration. The Y-chromosome is particularly tricky, since it has been reported that integration of constructs on this chromosome is extremely difficult.
It is our goal to apply our knowledge to farm animal breeding, in particular to achieve sex ratio distortion in farm animals. Towards this goal, we have started a collaboration with Heiner Niemann, director of the "Institut für Nutztiergenetik" at the Federal Research Institute for Animal Health, Neustadt, in trying to achieve TRD in boars carrying constructs expressing the mouse *t*-complex responder.

Cooperation within the institute

Department of Vertebrate Genomics

- Bernhard Herrmann was coordinator of the NGFNplus Consortium *Modifiers of Intestinal Tumour Formation and Progression,* funded by the BMBF, comprising the following subprojects supervised by scientists from the Department of Vertebrate Genomics:
  - SP1: *Analysis of human colon tumour and normal tissue samples and validation experiments in human cell lines* (supervisor: Michal Schweiger)
  - SP3: *Immunoprecipitation of methylated DNA (MeDIP) and expression profiling with a new high throughput sequencing technique* (supervisor: Christina Grimm)
  - SP6: *Computational analysis and data integration* (supervisor: Ralf Herwig)
- Marie-Laure Yaspo: *TREAT20 Consortium "Tumour REsearch And Treatment - 20 Patient Pilot*", funded by the BMBF
- Florian Mertes/ Heinz Himmelbauer: *Generation of a orofacial cleft 1 candidate 1 (Ofcc1) mouse mutant*
- Markus Ralser: Generation and phenotypical characterization of a triosephosphate isomerase 1 (Tpi1) knock-in mouse mutant

Department of Computational Molecular Biology

- Juliane Perner , Martin Vingron: *Analysis of Polycomb regulation and recruitment during early mesoderm formation*.

Research Group Development and Disease

- Silke Lohan: *Functional analysis of the basic helix-loop-helix family member a9 gene (Bhlha9) in the mouse*
- Sigmar Stricker/Stefan Mundlos: *Generation of an odd-skipped related 1 (Osr1) mouse mutant*

Otto Warburg Laboratory

- Ulrich Stelzl:
  - *Identification of mediator protein 12 (Med12) interactors*
  - *Y-2-H screening for interactors of the responder and distorters involved in TRD*
  - *Search for interactors of METTL2 and Brachyury using Y-2-H screening*

Mass Spectrometry Facility

David Meierhofer: *Identification of mesoderm specific KDM6A/UTX and KDM5B interacting proteins.*

Sequencing Facility

- Bernd Timmermann:
  - Transcriptome sequencing of embryonic tissues derived from E8.25 mouse embryos
  - RNA-seq and DNA methylome sequencing of mouse intestine and adenoma samples
  - RNA-seq analysis of staged mouse testes
  - Genomic sequencing of the *t*-haplotype t^{w5} following region-specific enrichment
  - Region specific enrichment and sequencing of the genomic interval containing the *tufted* locus

Microscopy Group
Rudi Lurz: *Localization of new interaction partners of SMOK in spermatozoa by electron microscopic analysis of the mouse epididymis*

Special facilities and equipment

The Department of Developmental Genetics provides personnel and expertise to the *Transgenic Unit* and the *Stem Cell Unit* of the institute. The former produces transgenic embryos and mice from genetically modified ES cells for various groups at the institute. It is also responsible for cryopreservation of transgenic mouse strains. The latter provides expertise in handling and culturing of ES cells, and gives advice and help in vector construction and generation of knock-out mice and other genetically modified mouse models.
Members of the department have the supervision over the Fluorescence Activated Cell Sorter of the institute. Another important piece of equipment is the recently acquired 2-Photon Laser Scanning Microscope of the institute, which is housed in the department.

General information about the whole Department

Complete list of publications (2009-2012)

2012

Bauer H, Schindler S, Charron Y, Willert J, Kusecek B, Herrmann BG (2012). *The nucleoside diphosphate kinase gene Nme3 acts as quantitative trait locus promoting non-Mendelian inheritance.* PLoS Genetics 8(3):e1002567

Farrall AL, Riemer P, Leushacke M, Sreekumar A, Grimm C, Herrmann BG, Morkel M (2012). *Wnt and BMP signals control intestinal adenoma cell fates.* Int J Cancer 3663, doi:10.1002/ijc.27500

Geffers L, Herrmann B, Eichele G. (2012) *Web-based Digital Gene Expression Atlases for the Mouse*, Mammalian Genome, in press

Herrmann BG, Bauer H (2012). *The mouse t-haplotype: a selfish chromosome – genetics, molecular mechanism and evolution*. In: Evolution of the Mouse, ed. Milos Macholan, Cambridge University Press

Morkel M (2012). *Cell Signal Transduction*. In: Principles of Molecular Diagnostics and Personalized Cancer Medicine, eds. Tan D, Lynch H, Lippincott Williams & Wilkins 2012, in press

Sudheer S, Bhushan R, Fauler B, Lehrach H, Adjaye J (2012). *FGF inhibition directs BMP4-mediated differentiation of Human Embryonic Stem Cells to syncytiotrophoblast*. Stem Cells Dev 2012 Jun 22 (Epub ahead of print

Wittler L, Hilger A, Proske J, Pennimpede T, Draaken M, Ebert AK, Rösch W, Stein R, Nöthen MM, Reutter H, Ludwig M (2012). *Murine expression and mutation analyses of the prostate androgen-regulated mucin-like protein 1 (Parm1) gene, a candidate for human epispadias*. Gene, in press

2011

Leushacke M, Spörle R, Bernemann, Brouwer-Lehmitz A, Fritzmann, Theis M , Buchholz F, Herrmann BG, Morkel M (2011). *An RNA interference phenotypic screen identifies a role for FGF signals in colon cancer progression.* PLoS ONE 6 (8):e23381

Norton LJ, Zhang Q, Saqib KM, Schrewe H, Macura K, Anderson KE, Lindsley CW, Brown HA, Rudge SA, Wakelam MJO (2011). *PLD1, rather than PLD2, regulates phorbol ester-, adhesion dependent-, and Fcγ receptor-stimulated ROS production in neutrophils.* J Cell Sci 124(Pt 12):1973-83

Qi L, Chen K, Hur DJ, Yagnik G, Lakshmanan Y, Kotch LE, Ashrafi GH, Martinez-Murillo F, Kowalski J, Naydenov C, Wittler L, Gearhart JP, Draaken M, Reutter H, Ludwig M, Boyadjiev SA (2011). *Genome-wide expression profiling of urinary bladder implicates desmosomal and cytoskeletal dysregulation in the bladder exstrophy-epispadias complex.* Int J Mol Med 27(6):755-765

2010

Althoff GEM, Wolfer DP, Timmesfeld N, Kanzler B, Schrewe H, Pagenstecher A (2010). *Long-Term Expression of Tissue-Inhibitor of Matrix Metalloproteinase-1 in the Murine Central Nervous System Does Not Alter the Morphological and Behavioral Phenotype but Alleviates the Course of Experimental Allergic Encephalomyelitis.* Am J Pathol 177(2):840-853

Ching BJ, Wittler L, Proske J, Yagnik G, Qi L, Draaken M, Reutter H, Gearhart JP, Ludwig M, Boyadjiev SA (2010). *p63 (TP73L) a key player in embryonic urogenital development with significant dysregulation in human bladder exstrophy tissue*. Int J Mol Med 26(6):861-7

Draaken M, Proske J, Schramm C, Wittler L, Bartels E, Nöthen MM, Reutter H, Ludwig M (2010). *Embryonic expression of the cysteine rich protein 61 (CYR61) gene: A candidate for the development of human epispadias*. Birth Defects Res A Clin Mol Teratol 88(7):546-50.

Herrmann BG (2010). *Embryology meets cancer research*. Public Service Review, Science & Technology 07:242-243

Jürchott K, Kuban R-J, Krech T, Blüthgen N, Stein U, Walther W, Friese C, Kiełbasa SM, Ungethüm U, Lund P, Knösel T, Kemmner W, Morkel M, Fritzmann J, Schlag PM, Birchmeier W, Krueger T, Sperling S, Sers C, Royer HD, Herzel H, Schäfer R (2010). *Identification of Y-box binding protein 1 as a core regulator of MEK/ERK pathway dependent gene signatures in colorectal cancer cells*. PLoS Genetics 6(12):e1001231

Rocha PP, Bleiss W, Schrewe H (2010). *Mosaic Expression of Med12 in Female Mice Leads to Exencephaly, Spina Bifida, and Craniorachischisis.* Birth Defects Research (Part A) 88:626-632

Rocha PP, Scholze M, Bleiß W, Schrewe H (2010). *Med12 is essential for early mouse development and for canonical Wnt and Wnt/PCP signaling*. Development 137:2723-31

Scholz AK, Klebl BM, Morkel M, Lehrach H, Dahl A, Lange BM (2010). *A Flexible Multiwell Format for Immu-*

nofluorescence Screening Microscopy of Small-Molecule Inhibitors. Assay Drug Dev Technol 8(5):571-580

Timmermann B, Kerick M, Roehr C, Fischer A, Isau M, Boerno ST, Wunderlich A, Barmeyer C, Seemann P, Koenig J, Lappe M, Kuss AW, Garshasbi M, Bertram L, Trappe K, Werber M, Herrmann BG, Zatloukal K, Lehrach H, Schweiger MR (2010). *Somatic mutation profiles of MSI and MSS colorectal cancer identified by whole exome next generation sequencing and bioinformatics analysis.* PlosONE 5(12):e15661

Vidigal JA, Morkel M, Wittler L, Brouwer-Lehmitz A, Grote P, Macura K, Herrmann BG (2010). *An inducible RNA interference system for the functional dissection of mouse embryogenesis.* Nucl Acids Res 38:e122

2009

Birtwistle J, Hayden RE, Khanim FL, Green RM, Pearce C, Davies NJ, Wake N, Schrewe H, Ride JP, Chipman JK, Bunce CM (2009). *The aldo-keto reductase AKR1C3 contributes to 7,12-dimethylbenz(a)anthracene-3,4-dihydrodiol mediated oxidative DNA damage in myeloid cells: implications for leukemogenesis.* Mutant Res 662(1-2):67-74

Fritzmann J*, Morkel M*, Besser D, Budczies J, Kosel F, Brembeck FH, Stein U, Fichtner I, Schlag PM, Birchmeier W (2009). *A colorectal cancer expression profile that includes transforming growth factor beta inhibitor BAMBI predicts metastatic potential.* Gastroenterology 137(1):165-75. (* M. Morkel is equally contributing principal author)

Khanim FL, Hayden RE, Birtwistle J, Lodi A, Tiziani S, Davies NJ, Ride JP, Viant MR, Günther U, Mountford JC, Schrewe H, Green RM, Murray JA, Drayson MT, Bunce CM (2009). *Combined bezafibrate and medroxyprogesterone acetate: Potential novel therapy for acute myeloid leukaemia.* PLoS One 4(12):e8147

Verón N, Bauer H, Weisse A, Lüder G, Werber M, Herrmann BG (2009). *Retention of gene products in syncytial spermatids promotes non-Mendelian inheritance as revealed by the tcomplex-responder.* Genes & Develop 23(23):2705-10

Awards and scientific honours

Pedro Rocha: *Best talk at the NucSys Marie Curie Research Training Network*, Wageningen, 2010

PhD theses

Benedikt Schwartz: *Migration of Mesodermal Cells and Axial Elongation in Mouse Embryos Depend on the Serum Response Factor*, Freie Universität Berlin, 01/2012

Eun-ha Shin: *Transcriptome analysis of Bmp4-induced mesoderm formation in vitro*, Technische Universität Berlin, 12/2011

Marc Leushacke: *Functional Characterisation of genes that regulate Intestinal Tumor Progression*, Freie Universität Berlin, 02/2011

Pedro Rocha: *Med12 is an essential coordinator of gene regulation during mouse development*, Freie Universität Berlin, 02/2011

Joana Vidigal: *An inducible RNAi system for the functional dissection of genes in the mouse*, Freie Universität Berlin, 02/2011

Anja Michaela Mayer: *Analysis of the expression pattern and knock-out phenotype of Slit-like 2 (Slitl2) in the mouse*, Freie Universität Berlin, 02/2010

Nathalie Verón: *Untersuchungen zu den molekularen Grundlagen der nicht-mendelschen Vererbung in der Maus*, Freie Universität Berlin, 04/2009

Solveig Müller: *Funktionelle Charakterisierung regulatorischer Gene bei der Bildung der Wirbelsäule der Maus*, Freie Universität Berlin, 02/2009

Student heses

Lisette Lange: *Functional Characterization of the Brachyury Interacting Protein METTL2*, Master Thesis, Humboldt University of Berlin, 10/2010

Mathias Marks: *Characterization of genes involved in somitogenesis*, Master Thesis, Humboldt University of Berlin, 10/2010

Oliver Sedelmeier: *Charakterisierung der Funktion von Mettl-Genen während der frühen Embryonalentwicklung von Mus musculus,* Bachelor Thesis, Freie Universität Berlin, 10/2009

Teaching activities

Bernhard Herrmann is Professor for Genetics and Head of the Institute for Medical Genetics at the Charité – Universitätsmedizin Berlin, Campus Benjamin Franklin, and involved in the education of medical students. All scientists within the department participate in the teaching of students. We teach embryology in a practical course and a seminar to medical students, and provide a 3-week-long practical course with lectures and tutorials to students of the Masters Program "Molecular Medicine" at the Charité – Universitätsmedizin Berlin. In addition, we offer a 4-week-long practical course with lectures and tutorials to students of the Biology and Biochemistry curricula at the Freie Universität Berlin. We also offer a 1-week lab course with lectures to students of the MPIMG. These activities bring both early and later career scientists from the department in contact with students and help them develop their teaching skills. At the same time students of the Life Sciences are introduced to state-of-the-art developmental biology concepts and techniques in the mouse, which is not taught anywhere else in Berlin. In this way we hope to initiate an interest into the fascinating field of stem cell and developmental biology.

Seminar and lectures given by external speakers in the department

2012

Amelie Wegener, Université Pierre et Marie Curie, France, 27.01.2012. *OLIG2 function in oligodendrocyte differentiation*

2011

Denis Duboule, Department of Genetics and Evolution, Faculty of Sciences, University of Geneva, Switzerland, 19.11.2011. *The dynamic architecture of Hox gene clusters*

Tony Bolger, MPI of Molecular Plant Physiology, Golm, 15.11.2011. *Sequencing Solanum Pennellii*

Jiri Foreijt, Institute of Molecular Genetics, Academy of Sciences, Czech Republic, Prague, CZ, 24.10.2011. *Epigenetics and Speciatin: The case oft he Prdm9 Gene*

Robert Philibert, Dept. of Psychiatry, Carver College of Medicine, University of Iowa, USA, 18.07.2011. *The Role of Med12 in Human Health and Behavior*

Andrew Copp, Neural Development Unit, UCL Institute of Child Health, London, UK, 14.03.2011. *Neural tube closure defects – from molecular mechanisms to prevention in humans*

2010

Manuel Leichsenring, Department of Developmental Biology, Albert-Ludwigs University Freiburg, 25.11.2010. *Gene regulatory networks in early zebrafish development*

Yusuke Ohnishi, Mammalian Development Laboratory, Max Planck Institute for Molecular Biomedicine, Münster, 25.10.2010. *Molecular mechanisms underlying the lineage establishment in the early mouse embryo*

Victor Ambros, Program in Molecular Medicine, University of Massachusetts Medical School, Worcester, USA, 11.10.2010. *MicroRNA pathways in animal development and disease*

Robert A. Weinberg, Whitehead Institute for Biomedical Research, Cambridge, USA, 02.06.2010. *Cancer Stem Cells and Malignant Progression*

Jürgen Knoblich, Institute of Molecular Biotechnology, Austrian Academy of Sciences (IMBA), Vienna, Austria, 29.03.2010. *Asymetric Cell Division and Tumorgenesis in Drosophila Neural Stem Cells*

2009

Arica Beisaw, Dartmouth Medical School, Hannover, NH, USA, 30.11.2009. *The Role of the GATA4-FOG Transcription Complex in Cardiac Valve Development*

Mikhail Sukchev, Leibnitz Institute for Age Research Fritz Lipmann Institute, Jena, 21.10.2009. *Role of MCPH1gene in DNA repair and central nervous system development*

Jorge Cham, California Institute of Technology, Pasadena, USA, 14.10.2009. *The Power of Procrastination*

Shalev Itzkovitz, Weizmann Institute of Science, Computer Science and Applied Mathematics, Israel, 13.10.2009. *Microsatellite molecular clocks reveal the lineage relations between oocytes and bone-marrow stem cells*

Pedro Velica, University of Birmingham, UK, 28.08.2009. *Aldo-keto reductases of mice and men towards mouse models for human cancers*

Madelon Maurice, Department of Cell Biology, University Medical Centre Utrecht (UMCU), The Netherlands, 26.06.2009. *Ubiquitin-mediated control of upsteram Wnt signaling*

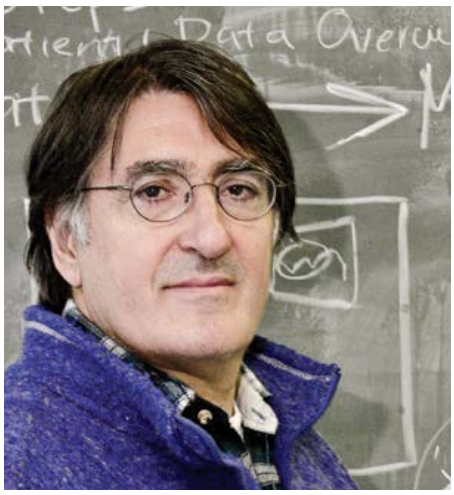
Colin L. Stewart, Institute of Medical Biology Immunos, Singapore, 04.05.2009. *Life at the Edge: or the Functional Architecture of the Nucleus in Development, Aging, and Disease*

Katia Sampieri, Pandolfi Lab, Cancer Genetics Program, Beth Israel Deaconess Cancer Centre, Harvard Medical School, Boston, USA, 24.03.2009. *Constitutive and somatic genetic events in retinoma and retinoblastoma*

Francis Stewart, Biotec TU Dresden, Germany, 26.01.2009. *Fundamental aspects of lambda Red recombineering and applications to high throughput proteomics and functional genomics*

Department of Vertebrate Genomics

(Established: 09/1994)



Head
Prof. Dr. Hans Lehrach
Phone: +49 (0)30 8413-1220
Fax: +49 (0)30 8413-1380
Email: lehrach@molgen.mpg.de

Secretary
Katrin Rettel
Phone: +49 (0)30 8413-1221
Fax: +49 (0)30 8413-1380
Email: rettel@molgen.mpg.de

Scientific Management
Wilfried Nietfeld
Phone: +49 (0)30 8413-1405
Fax: +49 (0)30 8413-1380
Email: nietfeld@molgen.mpg.de

Ralf Sudbrak*
(1000 Genomes, ITFoM)
Phone: +49 (0)30 8413-1612
Fax: +49 (0)30 8413-1380
Email: sudbrak@molgen.mpg.de

Ingrid Stark* (OncoTrack)
Phone: +49 (0)30 8413-1682
Fax: +49 (0)30 8413-1380
Email: stark@molgen.mpg.de

Heads of project groups
Christoph Wierling* (since 01/09)
Lars Bertram* (since 10/08)
Michal Schweiger* (since 01/07)
Albert Poustka* (since02/03)
Margret Hoehe (since 02)
Ralf Herwig (since 02/01)
Marie-LaureYaspo (since 96)
Zoltán Konthur (05/02-09/12)
Alexey Soldatov (06/07-07/12)
James Adyaje (07/01-04/12)
Harald Seitz (05/00-02/12)
Markus Ralser (12/07-12/11)
Bodo Lange (05/03-12/11)
Wilfried Nietfeld (10/98-12/11)
Jörn Glökler (02/08-02/11)
Silke Sperling (2001-02/11)
Georgia Panopoulou (01/04-12/10)
Andreas Dahl (07/07-03/10)
Eckhard Nordhoff (04/05-05/09)
Heinz Himmelbauer (07/95-05/08)

Introduction

Structure and organization of the department

In October 2012, the Department of Vertebrate Genomics consists of one Research Group led by Marie-Laure Yaspo (see group report), and six active project groups (L. Bertram, R. Herwig, M. Hoehe, A. Poustka, M. Schweiger and C. Wierling). In addition, seven project groups, which have either moved their primary affiliation elsewhere, or are currently in the process to do so, are still associated with the department (J. Adjaye (professorship, University of Düsseldorf), H. Himmelbauer (unit leader, CRG, Barcelona), Z. Konthur (DCGMS), B. Lange (CEO, Alacris Theranostics), M. Ralser (ERC group, Cambridge Systems Biology Centre), A. Soldatov (DCGMS), and S. Sperling (Heisenberg professorship, Charité/MDC). The associated groups are in a transit phase to leave the department and most of the project group leaders have already acquired external positions and will leave the department until November 2014. Historically, the

* externally funded

department has been funded to a very large extent by external grants. This has, on one hand, been an absolute prerequisite to allow us to pursue most of the core projects in the department. On the other hand, it has offered additional possibilities to young postdocs/finishing graduate students to develop independent project groups in relevant areas (sometimes outside of, but sometimes complementary to the core research of the department). This environment has given many young scientists a chance to become independent very early in their career, was very successful in the past and has helped to develop new research activities, which would not have happened otherwise.

Research concept

Life is, in a sense, a computational process. The organism "computes" a phenotype from its genome, given a specific environment. Such "computations" are for example taking place in the tumour, resulting in continuous growth, and in the normal tissues of a patient. When we treat a cancer patient with a specific drug, we are trying to modify the "computation" in the tumour to stop its growth, or drive the tumour cells into apoptosis, with minimal effects on the other cell types in the body of the patient. If we get it wrong (and in cancer treatment we get it wrong on the average in almost three quarters of the patients), the patient will end up getting expensive drugs, which have little or no effect on the tumour, but typically have more or less severe side effects on the patient. In oncology, we therefore spend very large sums on expensive drugs, which make the majority of the patients, who receive them, sicker than they might be without treatment. It is therefore essential to be able to predict the outcome of this "computation" taking place in different cell types of the patient in the absence or presence of the different drugs potentially available for therapy.

Much of the research done in the department is to predict the behaviour of complex biological systems, combining the tools of genetics, genomics, bioinformatics, and systems biology, with a special emphasis on systems medicine as the long term goal. The core research is therefore focussed on three main areas: (I) technology development, providing the basis for systematic analyses of biology and medicine on a genome wide level, (II) the analysis of regulatory processes in man and mouse, defining the biological networks disturbed in diseases and (III) the application of this approach to medical problems: systems medicine, with the main focus on oncology.

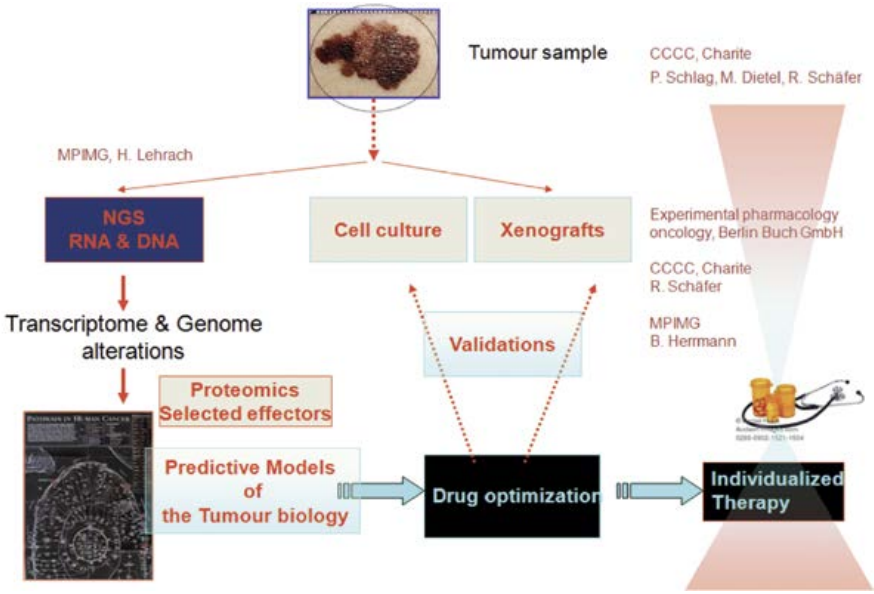


Figure 1: The concept of personalised medicine. Integration of different information sources is the basis for detailed computer models of individual patients. In collaboration with clinicians and basic science researchers, individual information, e.g. genomic and proteomic information from tumours, therapies etc. at different disease stages is integrated into a computer model, which can predict an individualised therapy (Treat20 project).

The concept of data generation and integration for a virtual patient is schematically illustrated in Figure 1. The prediction of the out-

come of the "computation" in the tumour as well as in other tissues of the individual patient is therefore the key to truly personalised medicine. It will enable us to predict effects and side effects of specific therapies on specific patients. Such a personalised medicine on the basis of detailed computer models of individual patients could be a real revolution in medicine. It is, for example, well known, that medical errors are one of the most common causes of death in the US (http://www.cancure.org/medical_errors.htm), a problem, which could be reduced considerably, if therapies could in the future be routinely checked on computer models of the individual patient, before any therapeutic intervention.
Based on the enormous progress in DNA sequencing and other analytical techniques, the wealth of information provided by basic research funded at a high level for decades, and the rapid progress in computing power, we have now the basis for the construction of such virtual patient models, a major focus of our work over more than a decade. It took the research community 10 years, and between one and three billion Dollars to sequence the first human genome. Now we are able to sequence more than one human genome per day in the department. Since the beginning of the war on cancer in the early seventies, roughly a trillion Dollars has been spent on cancer research alone, leading as yet only to moderate improvements in survival for many common forms of cancer. Much of the basic information for modelling the biology of cancer and other diseases should therefore be available. In addition, we have continuous improvements in computing power, with the power of the largest supercomputers roughly increasing by three orders of magnitude every ten years.

Scientific methods and achievements

The scientific work in the department has therefore increasingly focussed on systems genomics/systems medicine with a main focus on oncology. The core research work of the department can be divided into three basic sections:

I) Technology development

(Hans Lehrach, Marie-Laure Yaspo, Alexey Soldatov, Zoltán Konthur, Jörn Glökler, Markus Ralser, Harald Seitz, Wilfried Nietfeld, Florian Mertes)
Historically, the department has had a major role in technology development, ranging from the development of much of the commonly used genome robotics (clone picking, spotting, rearraying, PCR amplification), functional genomics (e.g. protein arrays), protein interaction by automated two hybrid analyses, semi-automated phage display, oligofingerprinting/SBH etc. Many of these developments have been picked up on large scale. Our first high density arrays constructed in 1987 have been the progenitor to the very large field of array-based gene expression analysis. Also, some of the first protein arrays have been developed in the department, which are still in use for the discovery of autoantigenicity patterns in different autoimmune and/or neurodegenerative disorders. Our large scale water bath PCR approach, developed in the late eighties, is a key step in the large scale bead PCR of the Solid sequencing system and is still being used in high-throughput genotyping (http://www.kbioscience.co.uk/). The automated Two Hybrid procedure developed at the department is still being used for the systematic analysis of protein interactions, e.g. in the group of Ulrich Stelzl at the institute. We were the first group to develop strand-specific RNA-Seq procedures,

as well as new approaches in DNA sequencing and DNA-coded library construction. Also new approaches in haplotyping have been developed. Developments in deep sequencing of transcriptomes of single cells and deep genotyping of free DNA in serum are being carried out as part of the ReaDNA and OncoTrack projects.

As work at the department draws to a close, ongoing developments increasingly have to be transferred to new legal entities, the Dahlem Centre for Genome Research and Medical Systems Biology (DCGMS), as well as Alacris Theranostics (Alacris), a start-up company based on technology development in systems genomics/systems medicine at the department over the last decade. This will, for example, include work on the development of new approaches in spatially resolved deep sequencing, a key technology for e.g. cancer pathology or developmental biology, the development of a detailed immune status, as well as new approaches for haplotyping etc.

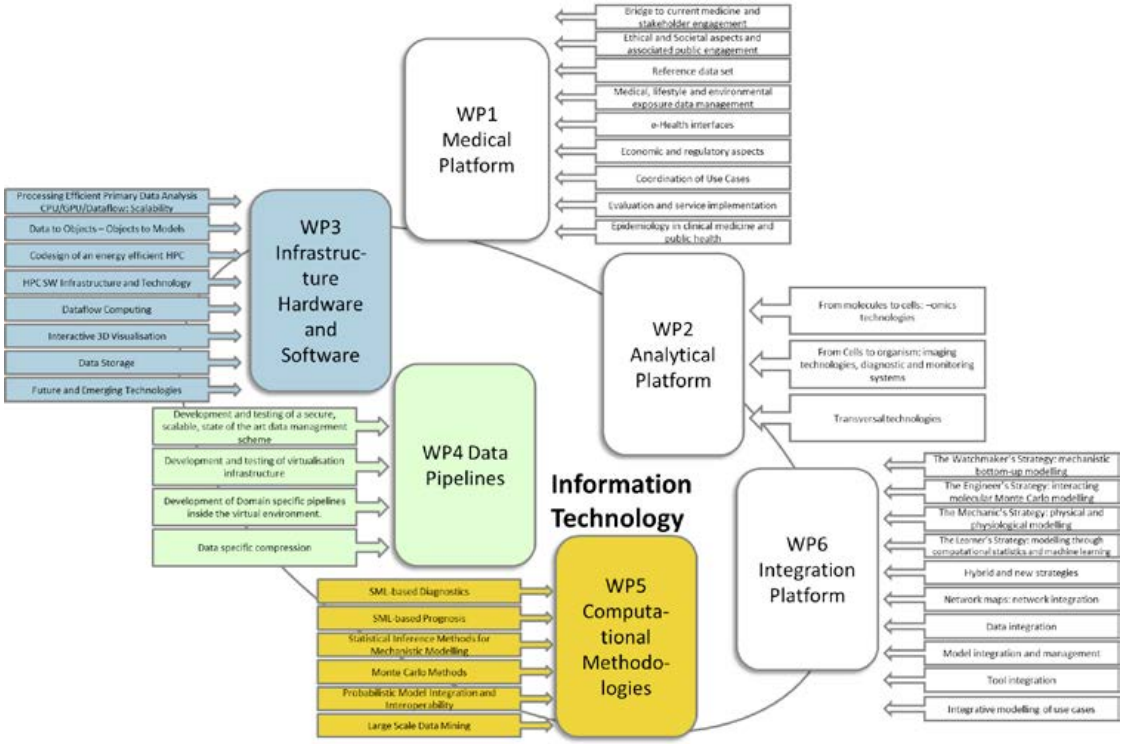


Figure 2: Breakdown of ICT and health care challenges in IT future of medicine (ITFoM). The proposed project consists of 6 platforms (WPs: a medical (WP1), an analytical (WP2), an infrastructure hard- and software (WP3), a data pipeline (WP4), computational methodologies (WP5), and an integration platform (WP6). This information technology concept combines the development of novel computing technologies to make them available for the future personalised medicine.

An important part of the technology development effort has been directed at the development of predictive modelling systems, the basis of our work in systems medicine, increasingly carried out in collaboration with Alacris and the DCGMS. As far as we are aware, the systems we have developed are unique world wide, illustrated by the fact that this system has been one of the major roots in a new EU 'flagship' proposal (IT Future of Medicine, ITFoM, www.ITFoM.eu), an application to the FET/Flagship call of the ICT section of the EU (http://cordis.europa.eu/fp7/ict/programme/fet/flagship/6pilots_en.html), which has the goal of constructing integrated molecular/physiological/anatomical models of every in-

dividual in the health care system, as the basis of a truly personalised medicine/prevention of the future (Figure 2). In a first phase, this proposal has been funded with 1.5 million Euros for one year as one of six pilot projects to develop full proposals, the only project coordinated from Germany, and the only with a focus on improving health care. Of these six proposals, two will be selected for funding at a level of approximately one billion Euros over ten years, funding levels, which would be similar in scale and duration to the human genome project. The project currently has approximately 150 (associated) partners, uniting the best groups in Europe, leading groups in other parts of the world, and a large number of leading companies, ranging from hard- and software development (Intel, Bull, IBM, Microsoft, SAP, Oracle, Xerox, Fujitsu, Siemens, Infinion etc.), over analytical/diagnostic techniques (Illumina, LifeTech, Oxford Nanopore, Bruker, Sciex) to pharma/biomarker companies (Sanofi-Aventis, Bayer, Roche, BioMerieux), subscribing to the vision of the model-driven individualised medicine of the future.

II) The flow of information from the genome to the phenotype

As contribution to a general understanding of the biological networks acting in the healthy organism, an essential base line for modelling disease processes, we are addressing a number of steps in the flow of information from the genome to the phenotype. Examples for this are:

Analysis of genome variation, the 1000 Genomes Project

(Marie-Laure Yaspo, Ralf Herwig, Hans Lehrach, in collaboration with Bernd Timmermann, Sequencing core facility)

As partner in the 1000 Genomes Project, the department contributes to the analysis of the variation in 'normal' (not disease-selected) individuals in multiple populations (http://www.1000genomes.org/), in a project involving mostly the groups of Marie-Laure Yaspo (sequencing, informatics) and Ralf Herwig (informatics). As part of this project, we are sequencing the genomes of about 150 individuals from different populations with a coverage of 4-6 fold. The sequence analysis of 94 genomes is already finished. (BMBF-funding until 12/12, coordinator Richard Durbin, Wellcome Trust Sanger Institute, and David Altshuler, Harvard Medical School)

The information in the haploid genome, the 'Haplome'

(Margret Hoehe)

The functional consequences of variants in the genome depend crucially on the phase they are in, an observation dating back to the original definition of a gene in bacteria as a cistron, defined by a genetic cis/trans test. The group of Margret Hoehe has, as one of the first in the world, independently developed a novel approach to haplotype-resolve whole genomes by next generation sequencing of fosmid pools and has generated the highest number of haplotype-resolved genomes to date. Their approach has been highlighted in the Nature Methods Special 'Method of the Year 2011' as one of the most promising ones. The group has now started a systematic effort to characterize haploid genomes at the DNA sequence and protein level and currently analyses individual variation between a total of 71 haplotype-resolved genomes (NGFN-Plus project 'MHC Haplotype Sequencing: An integrated approach to common disease').

Epigenetic Regulation

(Marie-Laure Yaspo, Michal-Ruth Schweiger, Ralf Herwig, Silke Sperling, Hans Lehrach)

Superimposed on the functional consequences of haplotypes are the effects of epigenetic modifications. We are involved in the blueprint project (http://www.blueprint-epigenome.eu/) aiming to characterize the epigenetic modification of different cell types in blood. Epigenetic analysis is also an important component of the analysis of tumours, and has, for example, provided useful biomarkers in a number of tumour types. Epigenetic mechanisms have also been analysed in detail in healthy and diseased heart, and will be further analysed in samples from the Modifier and Ecrin projects. The bioinformatics group has developed a full pipeline for genome wide methylation experiments which allows us to analyse differential methylation in human tumours in high-throughput and with high accuracy. This development was funded within the NGFN Modifiers project and will be maintained with funding from the BMBF eBIO program in human lung cancers (EPITREAT).

Systems genetics and evolution

(Marie-Laure Yaspo, Alexey Soldatov, Albert Poustka, Hans Lehrach)

A significant effort in the department has been invested in the use of systems genetics/systems genomics approaches to analyse regulatory mechanisms, combining human or mouse genetics with large scale genomic analysis tools. For the functional analysis and identification of cis-regulatory elements of human chromosome 21 genes and other human gene promoters, we applied transfected cell arrays to identify the modular functional architecture responsible for the regulation of promoter activity. In the Euvadis project (http://www.euvadis.eu/), we participate in an international consortium to analyse transcriptomes of cell lines from genetically very well characterized CEPH (www.cephb.fr) families. In addition, we have a long standing effort in genetical analysis of gene regulation in the mouse. For this, we have analysed gene expression patterns in mouse embryos in Eurexpress (http://www.eurexpress.org) and carried out a systematic analysis of gene expression patterns in mouse brain, liver (IMGuS Stetohepatitis project), and intestine (NGFN Modifier project) in chromosome substitution strains. The identification of trans-acting factors in this system has however proven more difficult than originally anticipated, since, on one hand, the relatively small signals caused by trans effects require very high data quality, which was hard to achieve in the early RNA-seq experiments. It has also proven more difficult than expected (both in our hands, but also in a number of other groups working in this area) to map trans-acting factors identified on the level of the chromosome further using back- or intercrosses. It will therefore be particularly interesting to be able to use strains from the collaborative cross, in a project (CCGGENSEQ) to be carried out as part of the ESGI infrastructure grant (in collaboration with Fuad Iriaqi, Ron Shamir, Richard Mott, and Heinz Himmelbauer). The group of Albert Poustka has addressed the question of the evolution at the protostome/deuterostome split, as well as the evolutionary origin of mammalian genes, analysed in more primitive deuterostomes (amphioxus, oikopleura, sea urchin). In an international effort, the group also discovered last year that acoel flatworms are related to the enigmatic worm Xenoturbella and, in contrast to previous studies, are grouped within a new deuterostomes phylum, which was named the Xenacoelomorpha.

III) Systems medicine

Systems medicine of cancer

(Marie-Laure Yaspo, Michal-Ruth Schweiger, Christoph Wierling, Ralf Herwig, Bodo Lange (Alacris), Markus Ralser, Hans Lehrach)

The core effort on systems medicine in the department is very much focussed on cancer, involving detailed characterization and (usually) modelling of different tumour types in a number of projects funded from different sources. The following projects are currently under way or have been concluded:

Pedbrain is an ICGC (International Cancer Genome Consortium) project on childhood brain cancer, with the goal of characterizing the tumour genome and transcriptome and the patient genome of 500 childhood brain tumours (DKG/BMBF, coordinator Peter Lichter, DKFZ).

ICGC Prostate Cancer (http://www.icgc.org/icgc/cgp/70/345/53039) is a project to analyse tumour genomes, epigenome, transcriptomes and patient genomes of 250 early onset prostate cancer patients (BMBF-funding until 10/15, coordinator Holger Sültmann, DKFZ).

Sequencing and bioinformatic analysis of childhood leukemia cases (acute lymphoblastic leukemia (ALL) is a pilot project on genome, exome, transcriptome and epigenome analysis in childhood leukaemia. The research plan is the transcriptome sequencing and analysis of seven B-cell precursor ALLs with a IKZF1 deletion and three B-cell progenitor ALLs with a chromosomal translocation t(17; 19) leading to a gene fusion of hepatic leukemia factor gene with the E2A gene. Both subgroups are highly resistant to therapy. Unfortunately for the very rare t(17; 19)-positive ALL type, there are no reports in the literature available about effective treatment strategies (BfS, until 07/13, coordinator Martin Stanulla & Andre Franke, University Hospital Kiel).

Treat 20 has the ultimate goal to validate our modelling approaches, based on deep sequencing of the genome and transcriptome of the tumour and the genome of 20 patients. We have generated NGS data for nine patients, diagnosed with either cutaneous or uveal melanoma, for which we are scoring somatic sequence variants and RNA expression changes, as well as transcriptome and exome data for a control melanocyte cell line. Our large cancer model shown in Figure 3 was used to generate patient specific tumour models by initialization with the patient's RNA expression profiles and mutation patterns to compute the differential response of normal and cancer cells in a Monte Carlo-based approach. Thus, the model predictions can identify potential key drivers of tumour progression and thereby detect targets for therapeutic approaches.

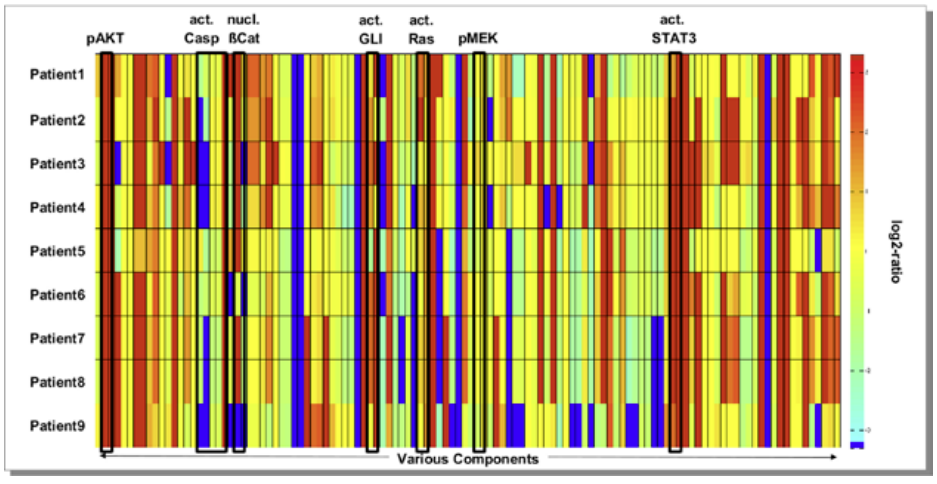


Figure 3: Heatmap representation of log2-ratios of the tumour vs. melanocyte state for four cutaneous melanoma (patient 1-4) and five uveal melanoma (patient 5-9) patients for key components of the model based on the expression profiles and mutation patterns. The model predicts elevated nuclear beta-catenin levels (red) in patient 1, 2, 3, 4, 5, 7 & 8 tumours and high activated STAT3 and pAKT levels in all patients indicating activated Wnt, PI3K or STAT3 signaling pathways, respectively. Predicted down-regulation (blue) of active-caspase levels in patients 2, 3, 4, 7, 8 & 9 indicates impaired apoptosis signaling.

In addition, the transcriptomes of tumour derived stem cells, tumour stem cells and xenografts will be analysed, providing the possibility to validate the predictions of the model by cell line/xenograft treatment (BMBF-funding until 09/12, coordinator Peter Schlag, CCCC, Charité).
The *OncoTrack consortium* (http://www.oncotrack.eu/) includes partners from academic research institutes, major clinical cancer centres, the European Federation of Pharmaceutical Industries and Associations (EFPIA) and SMEs. The consortium generates genomic and epigenomic sequence data from clinically well-defined colon tumours and colon tumour-derived materials. These data will be complemented by a detailed molecular characterization of the tumours using routine diagnostic procedures based on genomic or protein analyzes. A series of xenograft tumour models and cell lines derived from the same set of tumours will be established and characterized. The combined data from all phases of the project will be used to refine the model of the cancer cell, an approach based on the *in silico* "Virtual Patient" modelling system developed at the Max Planck Institute for Molecular Genetics and now being further developed by Alacris, both members of OncoTrack. Candidate biomarkers will be characterized in xenograft models and qualified for clinical use by analysis of an independent set of tumour tissues (Innovative Medicines Initiative (IMI), until 12/15, coordinator Hans Lehrach).
Prostate cancer - the scope of this research project is the genome and transcriptome analysis and modelling of prostate cancer in patients (IMGuS, until 11/11, Coordinator Holger Sültmann, DKFZ).
Prostate cancer project *Proceed* - in this project we performed the analysis of the methylome and exome of prostate cancer patients, and the identification of a methylator phenotype in translocation negative tumours (NGFN, until 12/11, coordinator Holger Sültmann, DKFZ).
Colon Cancer Modifier: Identification of modifiers of the tumour formation and methylome in colorectal cancers. In this project we have contributed to the chromosomal localization of modifiers of colon cancer frequency/epigenetic trans factors by mouse genetics in chromosome substitution strains. The analysis of human colon cancer samples in this project has provided a number of biomarkers for early as well as advanced stages of colorectal cancer and the project has investigated the relation between epigenetic modification and transcriptional regulation in cancer (NGFN, until 12/11, coordinator Bernhard Herrmann, MPIMG).
The *Mutanom project* (www.mutanom.org) addressed a key problem in modelling tumour biology, the unknown functional consequences of novel mutations found at high frequency in tumours. As part of this project, we have identified specific mutation patterns in colorectal cancers as well as in prostate cancer and carried out a wide range of functional analyses on the mutated proteins (NGFN/BMBF-funding until 12/11, coordinator Bodo Lange).
PREDICT - Personalised modelling of therapy responses in lung cancer patients. In this project, funded within the BMBF MedSys program, we analyse pre-clinical models of non-small cell lung cancers using a systems biology approach. Several molecular technologies are used, such as gene expression analysis, targeted enrichment of exonic genomic regions followed by high-throughput analysis, mutation screening and proteome analysis, for the characterization of the lung tumours. Using a combination of data analysis and computational modelling, the effect of targeted therapies is then predicted and combinatorial therapies are suggested by the computer models that are beneficial for the individual patient (BMBF until 06/12, coordinator Ralf Herwig).

MoGLI - Within the MoGLI project, we study the transcriptional program and the molecular circuitries regulated by Hedgehog signalling and modulated by integration of other key oncogenic signal inputs, such as EGF signalling, in cancer (BMBF MedSys, until 03/12, coordinator Hans Lehrach).
Other cancer-related research has been carried out on the regulatory function of metabolic networks in cancer in the group of Markus Ralser and on the identification of cancer drug targets in collaboration with the group of Sylvia Krobitsch.

Other diseases

In addition to the very focussed research on cancer, a number of other projects (completely funded from external sources) contribute to our knowledge of processes leading to other diseases or phenotypes (e.g. ageing), or explore possible diagnostic/therapeutic concepts.

Systems medicine of liver diseases

(Marie-Laure Yaspo, James Adjaye, Christoph Wierling, Hans Lehrach)
Two separate projects in the department address the biology of liver disease:
Systems biology of steatohepatitis has the goal to identify modifiers of liver toxicity and epigenetic modification using A/J and PWD-derived chromosome substitution strains using detailed molecular analyses in combination with modelling (IMGuS, until 06/11, coordinator Kurt Zatloukal, Medical University of Graz).
The systems biology of network stress based on data generated from in vitro differentiated hepatocytes derived from steatosis-specific human iPS cells has the goal to generate liver cells by reprogramming of iPS cells derived from liver disease patients. These cells are then analysed in detail (genome, transcriptome, proteome and metabolome analysis) and are used for the development of a mathematical model of the affected pathways (LIVSYSiPS/ERASysBio+, until 02/13, coordinator James Adjaye).

Alzheimer´s disease

(Lars Bertram)
The research focus of the Neuropsychiatric Genetics Group lays predominantly on the characterization of neuropsychiatric diseases e.g. Alzheimer's disease, Parkinson's disease. To achieve this, the group applies genome-wide genotyping and "next generation" sequencing technologies combined with bioinformatics and *in vitro* assays. A ground breaking result was the identification of the gene CD33 (siglec-3) as a novel genetic risk factor for Alzheimer's disease.

Ageing

(Lars Bertram, Hans Lehrach, James Adjaye, Markus Ralser)
A number of different lines of research in the department address the biology of ageing. To contribute to the genetic analysis, the department has recently received funding to collaborate in the "Berlin Aging Study II" (BASE-II, www.base-berlin.mpg.de/de/BASE_II.html), a multicentre study that aims to identify and characterize factors associated with healthy and unhealthy aging (BMBF).
The research concept on ageing in the group of James Adjaye is based on the same signalling mechanisms that regulate the plasticity of stem cells, which are

altered during aging and in age-related diseases. The goal of this research topic is to reverse the aging phenotype of somatic cells by cellular reprogramming as a means to study genome stability, mitochondria biogenesis and stem cell fate. Accordingly, an understanding of molecular and signalling mechanisms underlying the aging process is likely to lead to novel approaches to preventing and treating age-related diseases. In addition, work of the group of Markus Ralser has contributed to our understanding of ageing in *S. cerevisiae*.

Autoimmune diseases

(Zoltán Konthur)

The research of the *in vitro* Ligand Screening Group is focusing on the application of large scale screening methods to develop a better understanding of autoimmune disorders. The group is looking at different disorders, e.g. Rheumatoid Arthritis, Multiple Sclerosis and Alzheimer from different angles by analysing the antibody repertoire of patients applying 454-pyrosequencing in parallel to monitoring the autoantigenicity profiles using protein array technology and phage display. First results from antibody sequencing of a healthy control cohort spanning a wide age-range provided further evidence that not V(D)J recombination but rather reduction of class-switch capability is the major cause of immune senescence and that this process is starting relatively early. Applying protein array technology, autoantigenicity profiles have been discovered and patented, which can be used as prognostic or predictive markers for Rheumatoid Arthritis. A prognostic clinical study to evaluate these profiles is currently ongoing in collaboration with the Department of Rheumatology and Clinical Immunology of the Charité. In the reporting period, the group received predominantly governmental funding from the BMBF, BMWi, DFG as well as grants from private foundations and from the pharma industry.

Heart disease

(Silke Sperling)

The group of Silke Sperling is focusing on cardiovascular diseases. The group studies molecular mechanisms underlying cardiac development and function by using molecular biological techniques and systems biology approaches. The main research focus lies on the transcriptional regulation process of cardiac gene expression, which plays a key role for normal and abnormal cardiogenesis leading in the latter case to congenital heart disease (CHD). The group analyses in detail the networks, which are regulating gene expression, the interplay between different transcription factors, co-regulatory elements as well as miRNAs and epigenetic factors. Furthermore the group identified the first chromatin-remodelling factor DPF3 (BAF45c), which is able of binding acetylated as well as methylated histone modifications through its double plant-homeodomain (PHD-finger) and bridges distinct regulatory signals of the histone code and enables a tissues-specific read-out based on its neural, cardiac and muscle specific expression.

iPS cells: construction, analysis, reprogramming

(James Adjaye)

Patient-derived induced pluripotent stem cells (iPS cells) can have both diagnostic and therapeutic consequences. The systems biology and personalised medicine research concept of the group of James Adjaye has been focussed on the formation and reprogramming of human stem cells/iPS cells, with relevance to liver and other disease, and mechanisms of ageing (e.g. EU FP6 project ENFIN, BMBF project MedSYS, ERASysBio Plus). During the report period, the group has successfully generated and characterized iPS cells from human skin-derived fibroblast, amniotic fluid and chorionic fluid cells and comparatively differentiated these into hepatocyte-like cells, which has further increased our knowledge of the mechanisms underlying cellular reprogramming. Also iPS cells generated from individuals suffering from diseases such as Alzheimer's, Type 1 Diabetes and Type 2 Diabetes have been established to provide novel *in vitro* models for studying the disease mechanisms.

Planned developments

Organisational

Over the last three years, a number of leaders of project groups have acquired new positions, a process, which will continue over the next two years. Examples are Silke Sperling (Heisenberg professorship, MDC/Charité), James Adjaye (chair, University of Düsseldorf), Markus Ralser (junior ERC group, Cambridge Systems Biology Centre and Dept. of Biochemistry, University of Cambridge, UK), Heinz Himmelbauer (Centre for Genomic Regulation (CRG), Barcelona, Spain), and Bodo Lange (CEO, Alacris Theranostics). Additional group leaders are in the process of transferring to other positions and centres. In view of the planned closure of the department at the end of 2014, we plan to transfer the remaining core of our research effort to a new institute (Dahlem Centre for Genomic Research and Medical Systems Biology, Berlin), which, among other functions, would serve both as coordinating entity and one site of research for the ITFoM project, if we are successful.

Scientific

Over the next years, we plan to continue to focus on the continuing development of model-based, truly personalised medicine and virtual pharmaceutical development. This effort is essential to bridge the gap between basic research and medical application, between genomics, high-end computing, and new hardware designs. The work we continue to do is novel, interdisciplinary, and could have enormous implications for both, patients and health care costs, a major contributor to economic problems in Europe and the world.

Cooperations within the institute

There are many interactions between the department and other departments and groups at the institute. With the department of Bernhard Herrmann, we have collaborated on systems genetics (Consomics, Modifier, M.-L. Yaspo, M. Schweiger, R. Herwig) and systems medicine projects, (Treat20, M.-L. Yaspo, Modifiers. M.-L. Yaspo, M. Schweiger, R. Herwig). Collaborations with the department of Martin Vingron have in particular been focussed on the bioinformatics

of transcription factor/promoter interactions (M.-L. Yaspo), other bioinformatics projects (S. Sperling, M. Schweiger) as well as the International Max Planck Research School on Computational Biology and Scientific Computing (R. Herwig). In the Eurexpress project (www.eurexpress.org), the department has collaborated with the group of Stefan Mundlos. We have collaborated closely with the groups of Sylvia Krobitsch (yeast genetics, ataxias, systems genetics of anti cancer drugs, H. Lehrach, gene regulation mechanisms, M.-L. Yaspo, H. Lehrach), Sascha Sauer (systems proteomics, H. Lehrach), and Ulrich Stelzl (interaction network analysis, R. Herwig). The department has also interacted closely with the sequencing facility of Bernd Timmermann on a number of projects (e.g. OncoTrack, 1000 Genomes Project, M.-L. Yaspo, H. Lehrach).
A number of projects have been carried out in collaboration with other Max Planck Institutes (e.g. H. Weckerle (MPI of Neurobiology) W. Nietfeld, H. Lehrach) as well as in the BASE-II project (U. Lindenberger (MPI for Human Development), L. Bertram, W. Nietfeld, H. Lehrach).

Neuropsychiatric Genetics Group

(Established: 10/2008)

Head
Dr. Lars Bertram, MD
Phone: +49 (0)30 8413-1542
Fax: +49 (0)30 8413-1139
Email: bertram@molgen.mpg.de

Scientists and Postdocs
Maria Liebsch* (since 09/11)
Sara Ansaloni* (since 01/11)
Esther Meissner* (since 10/10)
Christina M. Lill* (since 01/10)
Johannes T. Röhr* (since 07/09)

PhD student
Julia Biesenack* (since 11/10)

Student
Marcel Schilling* (since 03/11)

Technicians/Engineers
Brit-Maren Schjeide* (since 10/08)
Leif Schjeide (Software engineer)* (09/10-05/12)

Scientific overview

The scientific expertise of the Neuropsychiatric Genetics Group lies in the mapping and characterization of complex disease genes, predominantly in the field of neuropsychiatric diseases. This is achieved by combining genome-wide genotyping and sequencing approaches with bioinformatics and *in vitro* assays. One of the first papers published since the inception of the Neuropsychiatric Genetics Group was the first ever family-based genome-wide association study (GWAS) in the field of Alzheimer's disease. This study – which was selected as one of the "Top 10 Medical Breakthroughs in 2008" by Time Magazine – resulted in the identification of CD33 (siglec-3) as a novel genetic risk factor for Alzheimer's.

* externally funded

This finding was recently replicated by two large international genetic consortia, emphasizing the important role of this gene in Alzheimer's pathogenesis. In addition to the laboratory work, our group has pioneered the development of bioinformatic approaches that systematically and quantitatively assesses genetic data for a number of phenotypes including Alzheimer's disease, Parkinson's disease, schizophrenia, and multiple sclerosis. More recently, we have initiated several projects that apply "next generation" sequencing to genetically complex diseases. The main project applying these powerful methods searches for novel early-only familial, i.e. disease-causing, Alzheimer's mutations using exome sequencing. A related and separately funded project utilizes next-generation sequencing for fine-mapping the CD33 locus that our group identified by GWAS in 2008 (see above). The aim of this study is to identify the DNA sequence variant(s) that underlie the GWAS association signal functionally. Another focus of our group lies in the genetic and functional characterization of micro-RNAs in neuropsychiatric diseases. This project, made possible through a Special Research Award of the Hans-and-Ilse-Breuer-Stiftung, uses a combination of systematic *in silico* and *in vitro* assessments to predict and experimentally validate the impact of DNA sequence variants on micro-RNA function. In order to increase power and specificity of gene-finding efforts in Alzheimer's disease, the group has begun to use CSF (cerebrospinal fluid) biomarkers as endophenotypes in genetic association analyses. One successful application of this approach was recently completed in a study where we could show that CSF-Ab levels correlate with polymorphisms at the PICALM locus. Finally, the group has recently received funding to head the genetics core of the "Berlin Aging Study II" (BASE-II), a multicenter study

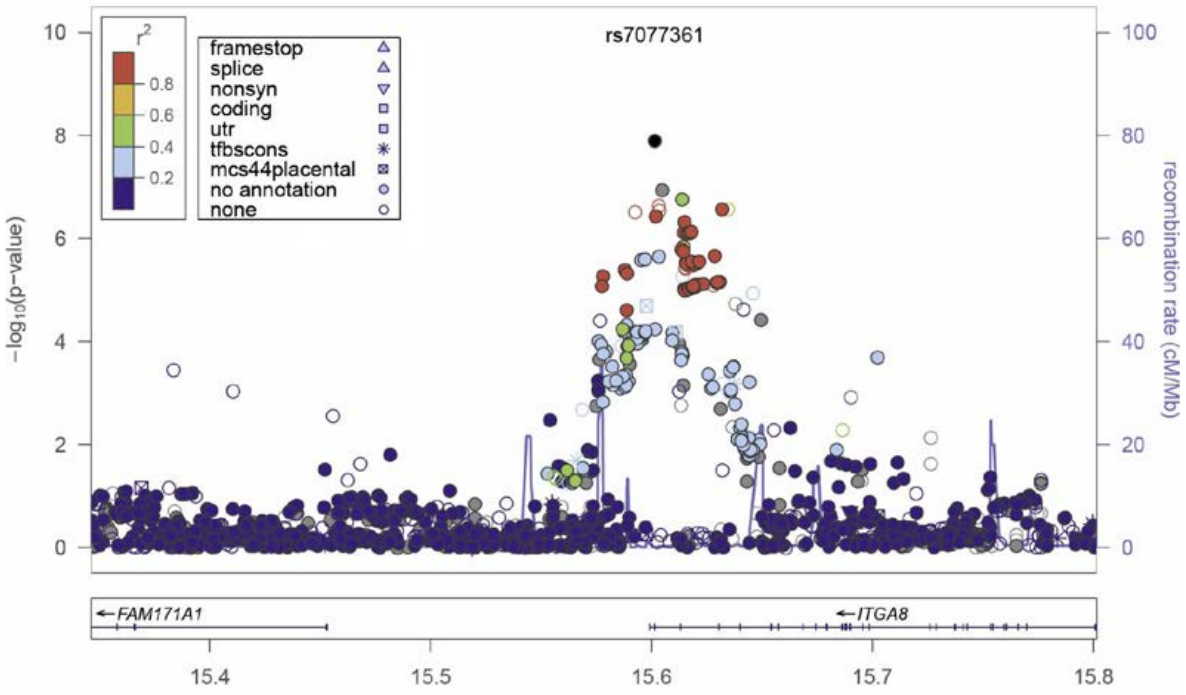


Figure 4: Locus plot of the ITGA8 region on chromosome 10p13 (15346353-15801533 bp, hg18) recently identified as a novel Parkinson's disease susceptibility gene by our group (Lill et al. [2012] PLoS Genet). Shown are association results for ~1,400 single nucleotide polymorphisms (SNPs) in the ITGA8 gene region on chromosome 10p13 based on meta-analyses including at least four independent datasets. SNPs are color-coded based on linkage disequilibrium (r2) estimates from the CEU dataset from the 1000 Genomes Projekt (release June 2010).

that aims to identify and characterize factors associated with healthy and unhealthy aging. BASE-II is a collaboration between the MPIMG, MPI for Human Development, Berlin, Charité – Universitätsmedizin Berlin, German Institute for Economic Research, Berlin, and Eberhardt-Karls University, Tübingen.

Selected publications

Lill CM, Roehr JT, McQueen MB, et al. (2012). *Comprehensive research synopsis and systematic meta-analyses in Parkinson's disease genetics: The PDGene database.* PLoS Genet; 8(3):e1002548

Schjeide BMM, Schnack C, Lambert JC, Lill CM, … Lehrach H, Amouyel P, von Arnim CA, Bertram L (2011). *The role of clusterin, complement receptor 1, and phosphatidylinositol binding clathrin assembly protein in Alzheimer disease risk and cerebrospinal fluid biomarker levels.* Arch Gen Psychiatry 68(2):207–213

Bertram L, Lill CM, Tanzi RE (2010). *The genetics of Alzheimer disease: back to the future.* Neuron 68(2):270–281

Bioinformatics Group

(Established: 02/2001)

Head
Ralf Herwig, PhD (since 02/01)
Phone: +49 (0)30 8413-1741
Fax: +49 (0)30 8413-1769
Email: herwig@molgen.mpg.de

Scientists and Postdocs
Atanas Kamburov* (since 03/12)
Kerstin Neubert* (since 07/11)
Axel Rasche* (since 09/09)
Mireia Vilardell* (since 09/09)
Christopher Hardt* (since 11/08)
Felix Dreher* (since 08/06)
Reha Yildirimman* (since 07/06)
Marcus Albrecht* (09/07-01/12)
Lukas Chavez Wurm* (01/11-12/11)
Anja Thormann* (06/10-12/11)
Thomas Meinel* (07/07–07/09)

PhD students
Matthias Lienhard* (since 11/11)
Atanas Kamburov (10/07-02/12)
Lukas Chavez Wurm (04/06-12/10)

Students
Jörn Dietrich (01/09–04/12)
Jevgeni Erehmann (07/10–12/11)
Hanna Galicka (12/09–11/11)

Scientific overview

The bioinformatics group develops methods and tools for the analysis and interpretation of biological data, predominantly in the domain of high-throughput sequencing, and subsequent interpretation of that data at the level of human interaction networks. The group has published 44 scientific publications during the reporting period. In several national and international consortia, we apply these tools and resources to the study of human disease processes (e.g. cancer, renal disorders and toxicology). The work of the group is structured in 1. methods development, 2. resources development and 3. applications to human diseases.

* externally funded

1. Methods developments

Ralf Herwig is member in several international consortia that are focused on bioinformatics methods for high-throughput sequencing, for example the 1000 Genomes Project, the SEQC consortium and the COST action SeqAhead funded by the EU Framework 7 programme. The group has developed computational methods for high-throughput sequencing applications in particular i) exome sequencing ii) RNA-seq and iii) MeDIP-seq.

i) We contributed with data analysis and technical support for the 1000 Genomes Project. Furthermore, we analyze exome enrichment data in human lung cancer patients and set up a full pipeline including SNP detection, SNP validation and correlation of results with gene expression and proteome data.

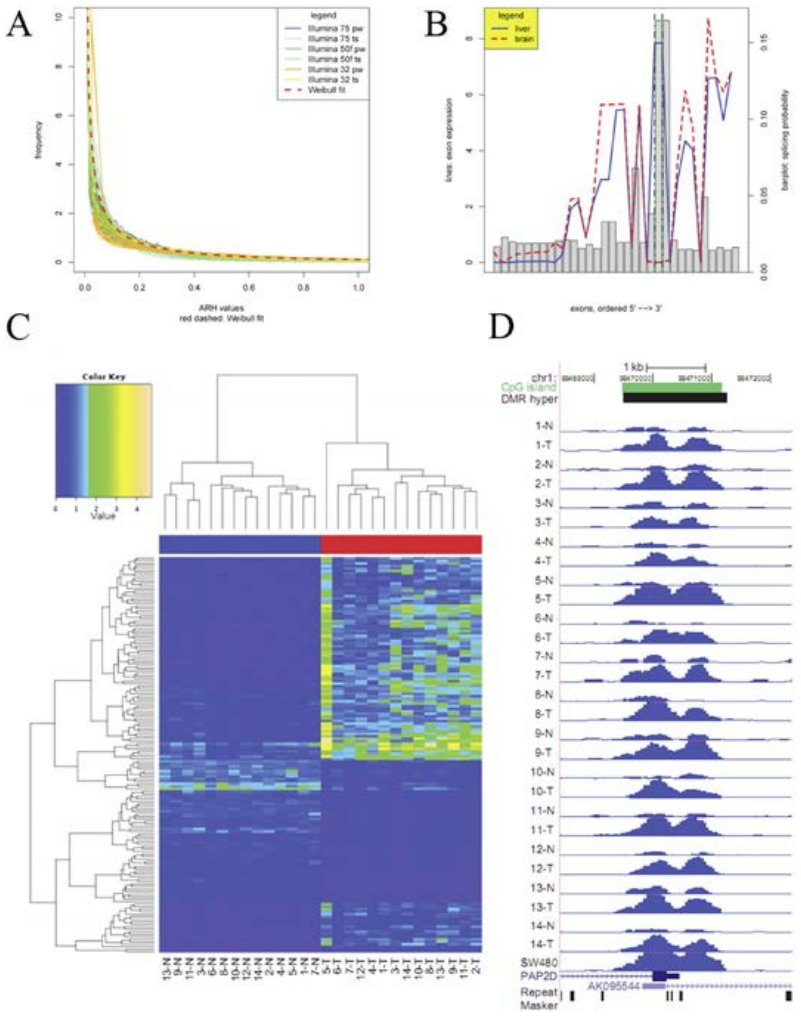


Figure 5: A) Distribution of ARH-seq values plotted for different sequencing data sets (in total 20 different human tissues). The resulting Weibull fit is superposed as dashed line and can be used to judge significance of differential splicing. B) ARH-seq predictions were evaluated against true positive splicing events derived from literature. Example shows a true positive splicing event in the gene MPZL1. RPKM values are visualised with the red dashed line for brain and blue solid line for liver. Splicing probabilities used for the entropy-based prediction are denoted as grey bars. Two recovered exons known for splicing are marked with green dot-dashed lines. C) Dendrogram of 158 genomic regions differentially methylated among tumour (red column labels) and normal tissue (blue column labels). D) Visualization of the region on chromosome 1 using the UCSC browser. RPM values are shown in wiggle format and show a consistent hypermethylation in the PAP2D promoter region. Panels show normal and tumour tissue for each patient as well as the SW480 cell line (bottom). Joint work with Michal Schweiger and Christina Grimm.

ii) A new method ARH – alternative splicing robust prediction by entropy - has been developed for the prediction of alternative splicing events from microarrays, based on the information theoretic concept of entropy. The method is non-parametric, robust and at low computational cost. An extension of the method, ARH-seq, has been adapted and intensively tested on high-throughput sequencing data (Figure 5A-B). Using published benchmark data on human tissues, we were able to show that the method outperforms existing approaches.

iii) We developed the MEDIPS software, the first comprehensive approach for normalization and differential analysis of genome-wide MeDIP-seq data.

MEDIPS is a full pipeline consisting of QC features and methods for data pre-processing and statistical analysis. MEDIPS has been intensively tested on data related to colon cancer patient data (figure 5C-D) and mouse models as well as the differentiation of human embryonic stem cells. MEDIPS has been made available for the community with an R/Bioconductor package.

2. Resources developments

Recent achievements include the integration of a critical mass of thirty human interaction databases into the unified system ConsensusPathDB. ConsensusPathDB integrates diverse heterogeneous interaction types such as protein-protein, signalling, metabolic, genetic and gene regulatory interactions and comprises, with 56,674 unique physical entities and 210,281 interactions, the largest collection of human interactions worldwide. This integrated network is regularly used for inferring genome-wide data at the network level, for example through gene enrichment analysis. Furthermore, we developed the web servers IMPaLA for the joint network analysis of metabolites and genes, and IntScore for the confidence assessment of interaction data.

Additionally, the group serves as data integration partner in several research collaborations and has developed the data integration system DIPSBC – data integration platform for systems biology collaborations - that is based on XML representation of different data types followed by indexing and querying data through a collaboration system. This system is currently used in five large collaborative projects, integrating a wide range of data types such as proteomics, transcriptomics, patient data, genotype data, networks and computational models.

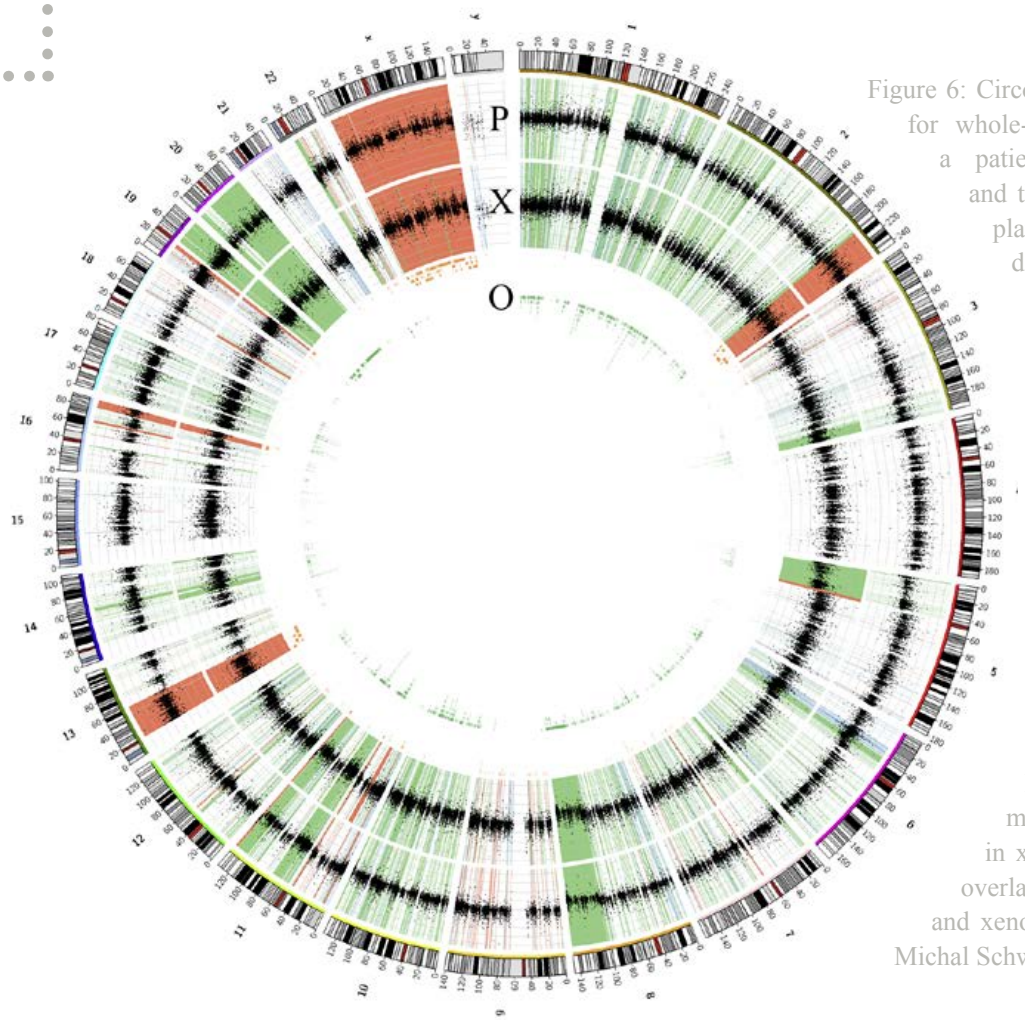


Figure 6: Circos plot showing results for whole-exome sequencing in a patient's primary tumour and the xenografted tumour planted on an immune-deficient mouse. Primary tumour tissue and xenograft tumour tissue show a high overlap of somatic mutations when compared with respective normal tissue (90%). Black: log2 coverage ratio; red: copy number gain region; green; copy number loss region; blue: LOH in tumour; P: CNVs in patient primary tumour; X: CNVs in xenografted tumour; O: overlapping CNVs in patient and xenograft. Joint work with Michal Schweiger.

3. Application to human diseases

The group is very well integrated in the genome analysis and systems biology research communities, for example with the following projects:

- In the APO-SYS project funded by the EU Framework 7 programme we have analyzed networks and computational models with respect to different aspects of apoptosis signalling in cancer.
- In the project "Modifiers of intestinal tumor formation and progression" within the National Genome Research Network Germany we analyze methylome, genome and transcriptome sequencing data of colon cancer patients and specific mouse models (consomics) in order to identify biomarker for colon cancer and to localize potential modifiers.
- In the project PREDICT of the Medical Systems Biology program we analyzing exome sequencing data along with multiple other resources in order to develop computational models and network modules that are able to predict the therapy success of targeted therapies in individual lung cancer patients (Figure 6).
- In the EC Framework 6 project carcinoGENOMICS and the Framework 7 project diXa we develop computational models of liver and kidney toxicity, in order to predict toxic effects of chemicals.

Selected publications

Kamburov A, Pentchev K, Galicka H, Wierling CK, Lehrach H, Herwig R (2011). *ConsensusPathDB - towards a more complete picture of cell biology*. Nucl Acids Res 39(1):712-717

The 1000 Genomes Project Consortium* (2010). *A map of human genome variation from population scale sequencing*. Nature 467:1061-1073 *Albrecht M, Herwig R

Chavez L, Jozefczuk J, Grimm C, Dietrich J, Timmermann B, Lehrach H, Herwig R*, Adjaye J* (2010). *Computational analysis of genome-wide DNA-methylation during the differentiation of human embryonic stem cells along the endodermal lineage.* Genome Res 20:1441-1450 *equal contribution

Genetic Variation, Haplotypes & Genetics of Complex Diseases Group

(Established: 2002)

Head
Dr. Margret Hoehe* (since 2002)
Phone: +49 (0)30 8413-1468
Fax: +49 (0)30 8413-1462
Email: hoehe@molgen.mpg.de

Secretary
Barbara Gibas* (10/05-03/11)

Scientists
Thomas Hübsch* (since 08/08)
Gayle McEwen* (02/10-02/12)
Eun-Kyung Suk* (07/05-08/11)
Sven Stengel* (04/10-04/11)
Britta Horstmann* (07/09-09/10)
Roger Horton * (09/08-03/10)
Dorothea Kirchner* (09/09-12/09)
Johannes Dapprich* (12/08-03/09)
Cornelia Platzer* (06/08-01/09)

Guest scientists
Gayle McEwen (since 03/2012)
Eun-Kyung Suk (since 09/11)
Thomas Sander (07/09-06/10)

PhD student
Jorge Duitama* (01/10-07/10)

Students
Anne-Susann Musahl* (06/08-03/12)
Matthias Linser* (07/09-09/11)

Technicians/Engineers
Sabrina Schulz* (07/08-05/11)
Stefanie Palczewski* (10/01-05/11)

Scientific overview

Human individuals are diploid by nature. Therefore, the independent determination of both haplotype sequences of an individual genome is essential to relate genetic variation to genome function, phenotype and disease. To address the importance of phase, we have 1) developed novel approaches to haplotype-resolve whole genomes; 2) generated the most comprehensively molecular haplotype-resolved human genome to date, “Max Planck One” (MP1); 3) conducted global analyses of existing molecular diplotypes encoding genes and larger functional

* externally funded

entities at the DNA sequence and protein level; 4) explored the role of phase for gene function, disease, and clinical interpretation of personal genomes through extraction and annotation of all genes carrying damaging mutations in either *cis* or *trans* configurations. We have expanded analyses to multiple genomes (70 and more) to systematically assess the nature and extent of molecular haplotype, and diplotype, diversity at the population level. This work will significantly advance our understanding of the (inherently individual) biology of genes and genomes, and prepare the ground for 'phase-sensitive' personal genomics and individualized medicine.

Development of a novel fosmid pool-based next generation sequencing (NGS) approach to haplotype-resolve whole genomes

The principle of this method (Figure 7) relies on the fact that sequencing of an individual fosmid produces a haploid 40 kb segment of the genome and that multiple fosmids can be pooled and sequenced together without significantly affecting the haploid nature of the output. Sequencing multiple pools leads to increasing saturation of molecular haplotype sequences across the entire genome.

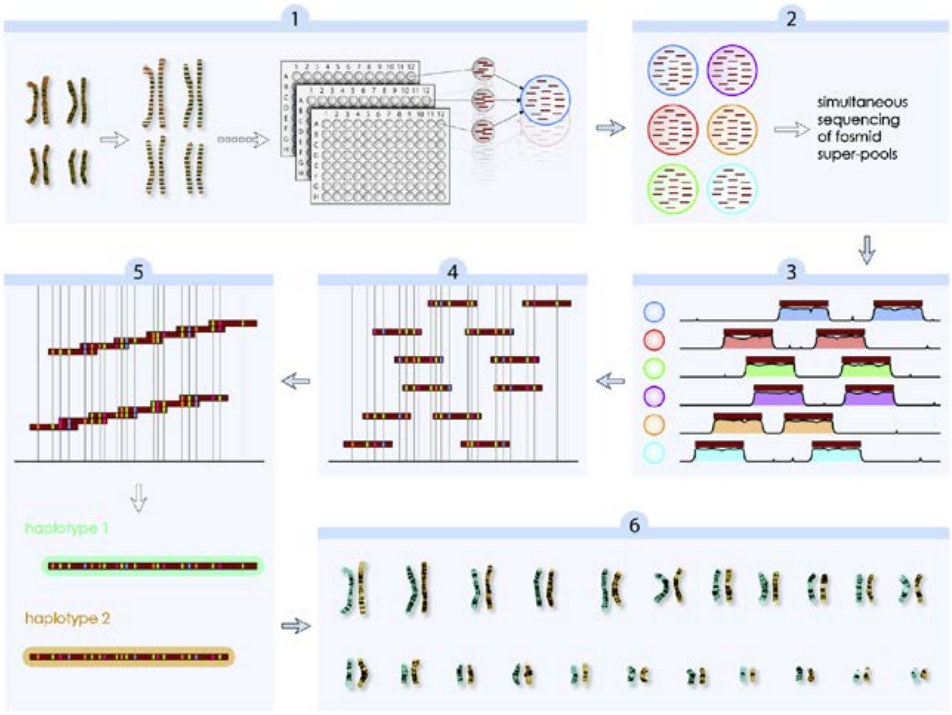


Figure 7: A fosmid pool-based NGS approach to haplotype-resolve an individual human genome. 1. Sheared genomic DNA of MP1 is used to prepare a complex fosmid library. About 1.44 million fosmid clones are generated, equivalent to ~7-fold coverage of each haploid genome, and distributed into three 96-well plates. Thus, each well comprises a pool of 5,000 fosmids, representing ~5% of the genome. To increase throughput, 3 fosmid pools are combined into 'super-pools' of 15,000 fosmids. The probability that complementary haplotypes may co-occur within a pool is $p<0.0112$. 2. Super-pools are bar-coded and multiplex-sequenced by use of NGS (SOLiD). 3. Fosmid sequences are detected from read coverage separately for each pool. 4. Variants are called on combined fosmid pools to identify heterozygous, haplotype-informative alleles for phasing. 5. Fosmids are then tiled into contiguous molecular haplotype sequences based on allelic identity at multiple heterozygous positions using ReFHap, a phasing algorithm developed by us. 6. Finally, the phased haploid contigs are anchored onto their homologous autosomes (from "Max Planck Research" 4.2011 "Neue Zweisamkeit im Erbgut").

Haplotype-informative, heterozygous positions within the genome are used to tile the haploid fosmids by allelic identity into long contiguous haploid sequences. To this end, we have established an independent NGS as well as bioinformatics pipeline, requiring the development of specific wet lab protocols and bioinformatic algorithms to be able to perform fosmid-based analyses. Our work has been highlighted in the Nature Methods Special 'Method of the Year 2011'.

Generating the most comprehensively molecular haplotype-resolved human genome to date

We have haplotype-resolved >90% of the genome of MP1, the first of 100 individuals of a German population cohort, from whom we established a unique 'Haploid Reference Resource' of 100 fosmid libraries. Virtually all SNPs (>99%) were phased into haploid sequences of up to 6.3 Mb (N50 ~1 Mb). These were anchored within each chromosome to provide long range chromosomal haplotypes. Importantly, we were able to phase >99% of the individual and rare SNPs of MP1, which existed in 75% of genes including upstream regions, making the majority of individual molecular haplotypes unique. Thus, we were able to capture an individual human genome in its molecular individuality.

We have haplotype-resolved a second genome, a HapMap trio child, reliable haplotypes of which had been deduced by 1000 Genomes Project. We were able to independently validate high accuracy of fosmid-based phasing and resolve 20% more of the variants, setting a new gold-standard. Notably, the necessary phase information for almost all potentially disease-relevant SNPs was generated.

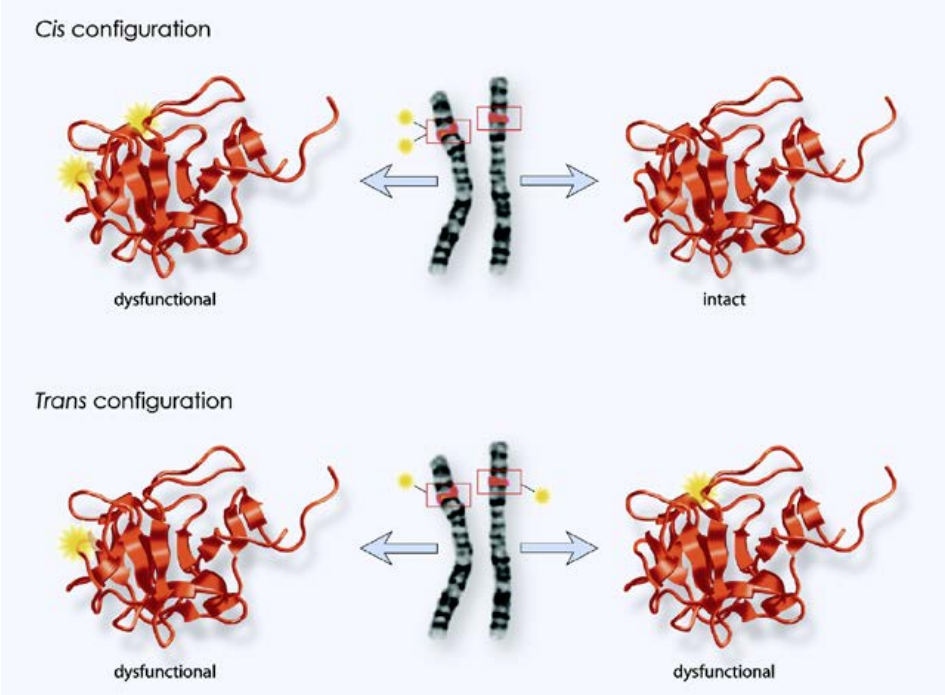


Figure 8: *Cis* vs. *trans* configurations of potentially protein damaging mutations (from "Max Planck Research" 4.2011 "Neue Zweisamkeit im Erbgut").

Systematic analyses of the importance of phase in biology and disease

Global survey of molecular diplotypes in 17,861 autosomal protein-coding genes: At the DNA sequence level, over 90% of all genes and upstream sequences contained at least one heterozygous SNP and so have two different molecular forms. Far most of those have two or more SNPs and therefore require phasing to uncover their *cis* or *trans* configurations. We determined the concrete molecular haplotype pairs for the vast majority of those.

Global analysis of cis vs. trans configurations in biology and disease: To gain first insights into importance of phase, we extracted the subset of phased genes that contained two or more potentially functionally significant mutations. Mutations residing on the same chromosome (in *cis)* leave the second protein intact; mutations on opposite chromosomes (in *trans)* may affect structure and function of both proteins (Figure 8). We conducted an in depth annotation of both *cis and trans* configurations in relation to a vast spectrum of disease phenotypes and pharmacogenomics, and exemplified in particular the potential clinical relevance of phase for prediction, treatment, and prevention of disease.

Moreover, we analyzed the role of phase in exerting biological functions at the example of *cis* versus *trans* configurations in zinc finger and olfactory receptor genes.

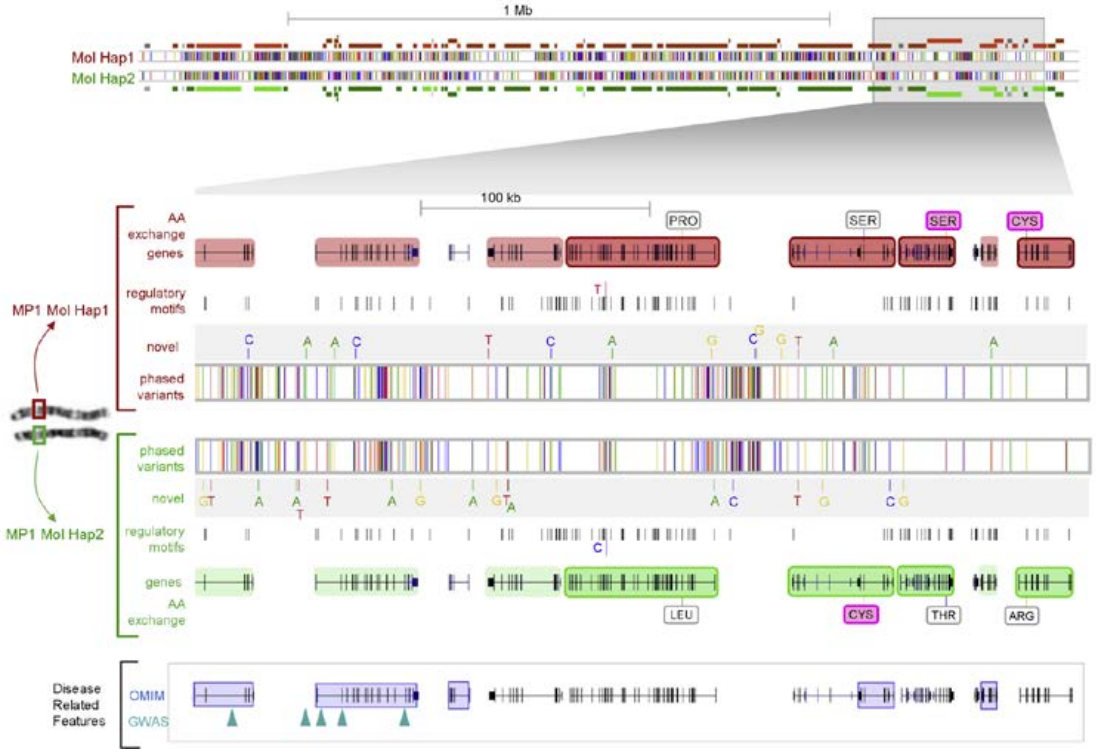


Figure 9: Example of a Megabase-size haploid landscape of functional variation on chromosome 19. Differences between the two molecular haplotypes are shown at the nucleotide level, regulatory level, and level of genome organization (from Suk et al., Genome Res 2011; 21(10):1672-85).

Determination of haploid landscapes of functional variation: Towards phase-sensitive personal genomics: The determination of contiguous molecular haplotype sequences in the Megabase range ('haploid landscapes') is crucial to translate individual genomic variation into the functionally active proteome. These may encode functional entities that affect gene expression in a coordinated way.

We have characterized and annotated over 700 haploid landscapes >1 Mb (Figure 9). In particular, we were able to resolve extended haplotypes in the MHC region, one of the most important regions in the human genome of key clinical relevance.

Systematic analysis of variation between multiple haplotype-resolved genomes

We are currently performing first systematic comparisons between a total of 71 (up to 386) genomes to assess the nature and extent of molecular haplotype, and diplotype, diversity at the population level. Our preliminary results indicate both, exorbitant diversity of individual gene forms at the DNA sequence level and a definable repertoire of diplotypic proteins. Our work has been selected as the Conference Abstract Winner of the 1st International Conference on Genomics in the Americas, Philadelphia, September 2012.

Conclusions and perspectives

The diplotypic nature of genes, regulatory sequences and larger functional entities is both substantial and global. Thus, the distinction of molecular haplotypes and diplotypes will be essential to resolve the inherently diploid biology of genes and genomes and prepare the ground for valid individualized medicine and personal genomics.

Selected publications

Duitama J, McEwen GK, Huebsch T, Palczewski S, Schulz S, Verstrepen K, Suk EK, Hoehe MR (2012). *Fosmid-based whole genome haplotyping of a HapMap trio child: Evaluation of Single Individual Haplotyping techniques*. Nucleic Acids Res 40(5):2041-53

Suk EK, McEwen GK, Duitama J, Nowick K, Schulz S, Palczewski S, Schreiber S, Holloway DT, McLaughlin S, Peckham H, Lee C, Huebsch T, Hoehe MR (2011). *A Comprehensively Molecular Haplotype-Resolved Genome of a European Individual*. Genome Res 21(10):1672-85

Lunshof JE, Bobe J, Aach J, Angrist M, Thakuria JV, Vorhaus DB, Hoehe MR*, Church GM* (2010). *Personal genomes in progress: from the Human Genome Project to the Personal Genome Project*. Dialogues Clin Neurosci 12(1):47-60 *co-last authors

Evolution & Development Group

(Established: 02/2003)

Head
Dr. Albert J. Poustka (since 02/03)
Dr. Georgia Panopoulou (01/04-12/10)
Phone: +49 (0)30 8413-1235
Fax: +49 (0)30 8413-1380
Email: poustka@molgen.mpg.de
panopoul@molgen.mpg.de

Scientists
Andrew Hufton (11/06-01/10)

PhD students
Udo Georgi (since 08/09)
Alexander Kuehn (04/05-07/09)

Students
Raik Otto (since 06/12)
Jonas Ibn-Salem (since 01/11)
Peggy Matz (04/12-06/12)
Sebastian Thieme (05/09-08/11)
Oliver Herrmann (07/09-02/10)
Christoph Standfuß (03/09-08/09)
Jana Tänczyk (10/08-05/09)
Lena Braun (09/08-05/09)

Scientific overview

Knowledge of evolutionary principles is essential for the understanding of both, the function of an organism and its relationship with its environment. In this respect, it is fundamental for the comprehension of human biology and disease. Evolutionary questions such as how species are formed and how environmental influences lead to the development of a range of phenotypes from a single genotype are very relevant for understanding the origin of phenotypic variation between individuals and how individuals might respond to environmental challenge. Whole genome duplications (WGD) are a recurrent feature of eukaryotic evolution, and the frequency of polyploid species suggests that WGD can confer selective advantages. WGD simultaneously generate a large amount of redundant genetic material, which when exploited can lead to functional novelty. Our group

has contributed to the understanding of the impact of WGDs in genomic order, the evolution of regulatory elements and duplicate gene function. We also study the above issues within an applied context, using medically relevant examples and additional animal model systems. Thus, we search for copy number variations (CNVs) and rearrangements linked to autism in whole genome and exome by next generation sequencing of autism families. We then functionally analyse SNPs and CNVs *via in vivo* assays in human cultures and in zebrafish embryos. As an example, we have analysed the expression patterns and the morpholino knockout phenotypes of the zebrafish orthologs of the human genes within the 16p11.2 600Kb region. One fifth of the patients carrying this deletion develop autism.

Taking advantage of the advance in sequencing technologies, we investigated the phylogenetic position of three marine worms, the Xenoturbellida, Acoelomorpha, and Nemertodermatida. Together with others, we have recently published a paper in Nature making use of the first draft assembly of the genome sequence of the enigmatic worm *Xenoturbella bocki*. We suggested that in contrast to previous studies acoelomorph flatworms (acoels and nemertodermatids) form a new phylum of deuterostomes together with Xenoturbella, which we have named the Xenacoelomorpha. Meanwhile, we have sequenced the complete genome and made the first draft assemblies of 5 members of Xenacoelomorpha. We also organised the first international Xenacoelomorpha genome sequencing project meeting in November 2011 in Berlin.

In another project recently completed, we have characterized *via* mass spectrometry the proteomes of several sea urchin skeletal elements.

Finally, in a recent DFG funded project and in collaboration with Dr. J. Ploetner (Natural History Museum, Berlin), we have sequenced the brain and testis transcriptomes of the European water frog *Pelophylax ridibunda*. The hybrid *Pelophylax esculenta* can only produce viable offspring by interbreeding with one of its parent species *P. ridibunda* or *P. lessonae* through a process called Hybridogenesis, where *via* an unknown mechanism during meiosis one parental genome is excluded.

Evolution of regulatory elements

Using a computational method developed in our lab, we identified phylogenetically conserved noncoding elements (PCNEs) in a manner that is not biased by rearrangement and duplication. We identified more than a thousand PCNEs that have been conserved between vertebrates and the basal chordate amphioxus. *Via* transgenic zebrafish assays we found that the majority of the computationally identified elements are functional enhancers. We could show that PCNEs are enriched around genes with ancient synteny conservation and that this association is strongest for extragenic PCNEs, suggesting that *cis*-regulatory interdigitation plays a key role in repressing genome rearrangement. Our results also demonstrate that subfunctionalization of conserved *cis*-regulation has not been the primary determinate of gene duplicate retention in vertebrates. Instead, the data supports the gene balance hypothesis, which proposes that duplicate retention has been driven by selection against dosage imbalances in genes with many protein connections.

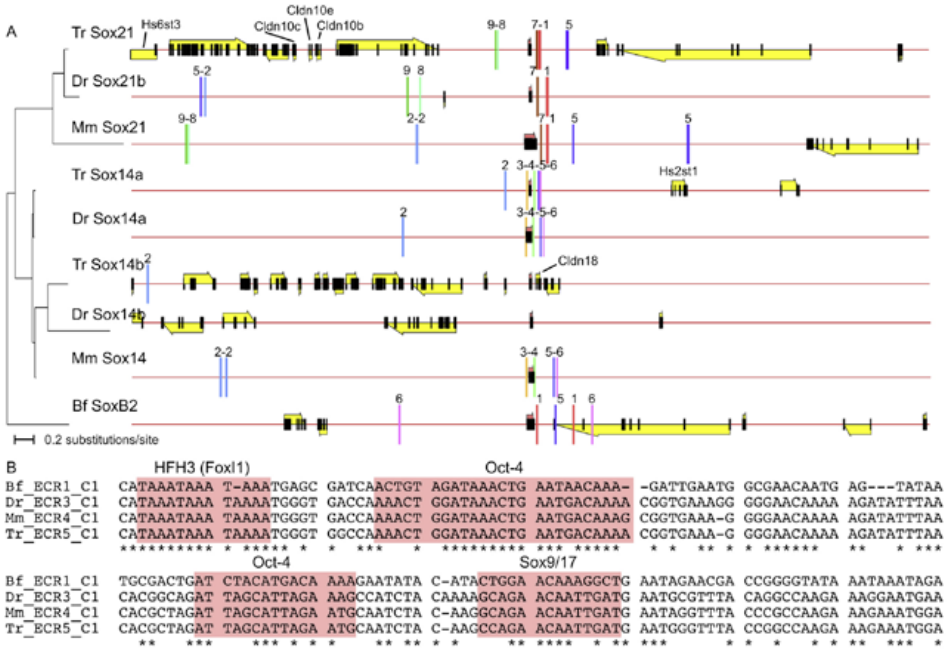


Figure 10: (A) PCNEs associated with the Sox14/21 gene family. Each genomic region is 204 kb long. Genes are shown as arrows indicating the direction of transcription. Black boxes indicate exons. The small single-exon Sox14/21 genes are labeled with red triangles. PCNEs are shown as colored lines; the colors and the numbers above the PCNEs indicate the group membership of each PCNE and reveal conservation of PCNE order. Neighboring genes that may represent conserved syntenic relationships are labeled; unlabeled genes are not syntenically conserved. A maximum likelihood phylogenetic tree of the Sox14/21 proteins is shown on the right.(B) A multiple alignment of a portion of the Sox14/21 group 1 PCNEs, with conserved binding motifs highlighted in pink. Bf, amphioxus; Tr, fugu; Dr, zebrafish; Mm, mouse.

New candidate autism susceptibility loci

In collaboration with the Clinic of Child Psychiatry of the University of Frankfurt (Prof. F. Poustka) and the German Cancer Research Center, Heidelberg (Dr. S. Klauck), we are carrying out whole genome and transcriptome sequencing of two multiplex autism families with at least two affected children and non affected parents and siblings. SNP analysis for the first family allowed us to identify a single *de novo* mutation in a gene of one of the patients that is likely to disrupt correct splicing of the corresponding transcript of an important regulatory protein, which is currently being analysed in cell-line assays. We anticipate enlarging this study and envision to setup an individualised diagnostics/treatment program for ASD in Germany.

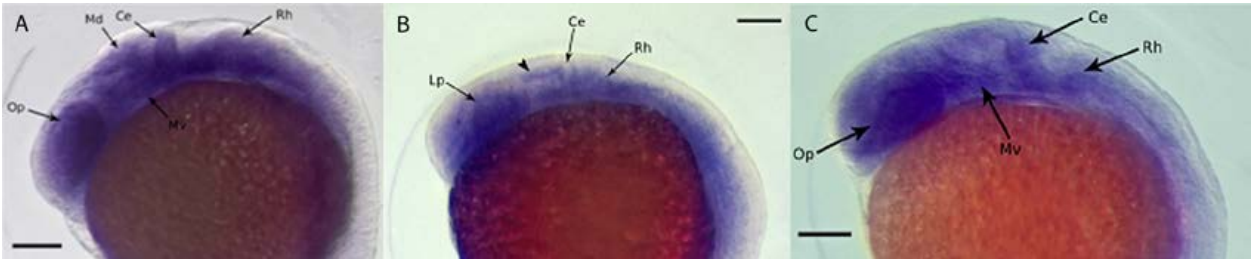


Figure 11: Lateral view of zebrafish 24hrs embryos showing the tissue expression of the zebrafish orthologs of the following three human 16p11.2 genes A: DOC2A, B: ypel3, C: ppp4ca genes at 24hrs. Abbreviations: Ce:Cerebellum, Op: oprtical plakode, Mv: ventral Mesencephalon (Tegmentum), Rh: Rhombencephalon, Lp: Linsenplakode

Zebrafish as model system to functionally analyse syntenic genes included in the human 16p11.2 ~600Kb region frequently deleted in autism patients

The tissue expression of nine zebrafish orthologs (*kctd13, ASPHD1, DOC2A, ypel3, gdpd3, mapk3, ppp4ca, ppp4cb, aldoab, shank3b, arsa*) of human genes located within the 16p11.2 region was analysed for six developmental stages between the gastrula (5.3-10hrs) and hatching (48-72 hrs) stages. Most of the genes are mainly expressed in neuronal tissues especially cerebellum, which is the brain region responsible for most of the functions that are affected in autistic patients.

Xenacoelomorpha: A new deuterostome Phylum

We have analysed three large datasets to investigate the phylogenetic position of Xenoturbella and Acoel worms: A) A complete set of mitochondrial genes from various acoels, nemertodermatids and Xenoturbella; B) A large phylogenomic data set of 38,330 aminoacid positions; and C) new micro RNA complements. Similar to previous studies, our phylogenomic analysis recovers a strong relationship between Xenoturbella and acoels, but shows also that the previous phyloge-

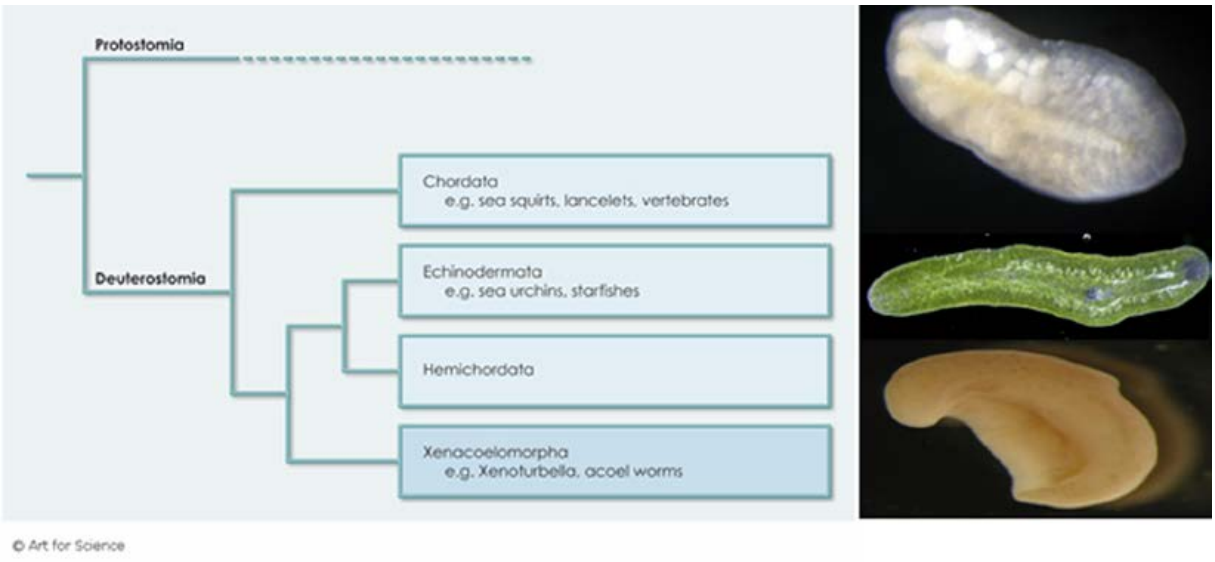


Figure 12: Modified phylogeny: The Xenacoelomorpha represent the fourth phylum of the deuterostomia (left). New deuterostomes, from top to bottom: The nemertodermatid Meara stichopi, the acoel Symsagitiffera roscofensis and Xenoturbella bocki (pictures on right side). Figures of worms are courtesy of Andreas Wallberg, Pedro Martinez and Hiroaki Nakano (top to bottom respectively).

netic analyses of acoels are strongly affected by a long-branch attraction (LBA) artefact. When we minimize LBA, we find consistent support for a position of both acoelomorphs and Xenoturbella within the deuterostomes. The most likely phylogeny links Xenoturbella and Acoelomorpha in a clade we call Xenacoelomorpha.

The Evolution of Simplicity: The Xenacoelomorpha Genome Project

We are currently sequencing the genomes of 5 different species of Xenacoelomorpha. Based on draft assemblies, we have started analyzing the phylogenetic order, intron conservation, gene loss and gain, synapomorphies, Hox and homeobox genes of these organisms. In collaboration with international experts we also annotate these genomes.

Proteomics of sea urchin skeletal elements

Using mass spectrometry-based methods, we have identified 231 proteins in the matrix of the *S. purpuratus* spicule matrix. Approximately two thirds of the identified proteins are either known or predicted to be extracellular proteins or transmembrane proteins with large ectodomains. The most abundant proteins of the spicule matrix are SM50-, SM30- and MSP130-related proteins, matrix metalloproteases and carbonic anhydrase.

Selected publications

Philippe H, Brinkmann H, Copley RR, Moroz LL, Nakano H, Poustka AJ, Wallberg A, Peterson KJ, Telford MJ (2011). *Acoelomorph flatworms are deuterostomes related to Xenoturbella*. Nature 470:255-258

Mann K, Wilt FH, Poustka AJ (2010). *Proteomic analysis of sea urchin (Strongylocentrotus purpuratus) spicule matrix*. Proteome Science 8(1):33

Hufton AL, Panopoulou G (2009). *Polyploidy and genome restructuring: a variety of outcomes*. Curr Opin Genet Dev 19(6):600-6

Cancer Biology Group

(Established: 01/2007

Head
Dr. med. Dr. rer. nat. Michal-Ruth Schweiger (since 01/07)
Phone: +49 (0)30 8413-1354
Fax: +49 (0)30 8413-1380
Email: mschweig@molgen.mpg.de

Scientists
Axel Fischer* (since 04/09)
Martin Kerick* (since 08/08)

PhD students
Melanie Isau* (since 04/09)
Andrea Wunderlich* (since 08/08)
Stefan Börno* (since 07/08)

Technicians
Nada Kumer* (since 04/08)

Students
Jana-Carina Elias (04/12-09/12)
Jerome Jatzlau (03/12-07/12)
Michelle Hussong (05/10-05/12)

Scientific overview

According to the world health organization (WHO), malignant neoplasms will be the most common cause of death worldwide in 2010. Despite intensive research on carcinogenesis this frightening scenario will persist mainly due to the overall increase of lifetime expectancy. Furthermore, most cancers are only diagnosed in an advanced stage, which prohibits curative treatment and a large proportion of patients do not respond to their chemotherapy. In a concerted action, based on the recent improvements of methodological techniques, we develop strategies for the identification of patients at risk and tumors and we intend to identify prognostic and predictive biomarkers as guides for patient's successful treatment at different stages of the disease. These goals are approached by means of newest high-throughput technologies combined with computational analyses. On the

* externally funded

other side, and at least of similar importance, we perform functional experiments to identify pathomechanisms underlying tumor development, progression and latency.

DNA Methylation analyses in prostate cancer

Prostate cancer (PC) accounts for more than 900,000 cases per year and is the second most common cancer among men worldwide. The clinical course of PC is heterogeneous, ranging from indolent tumours requiring no therapy during lifetime to highly aggressive PC developing into a metastatic disease. Despite its high prevalence, the clinical management of PC is limited by the low specificity of the existing diagnostic and prognostic tools and the lack of effective systemic therapeutic strategies.

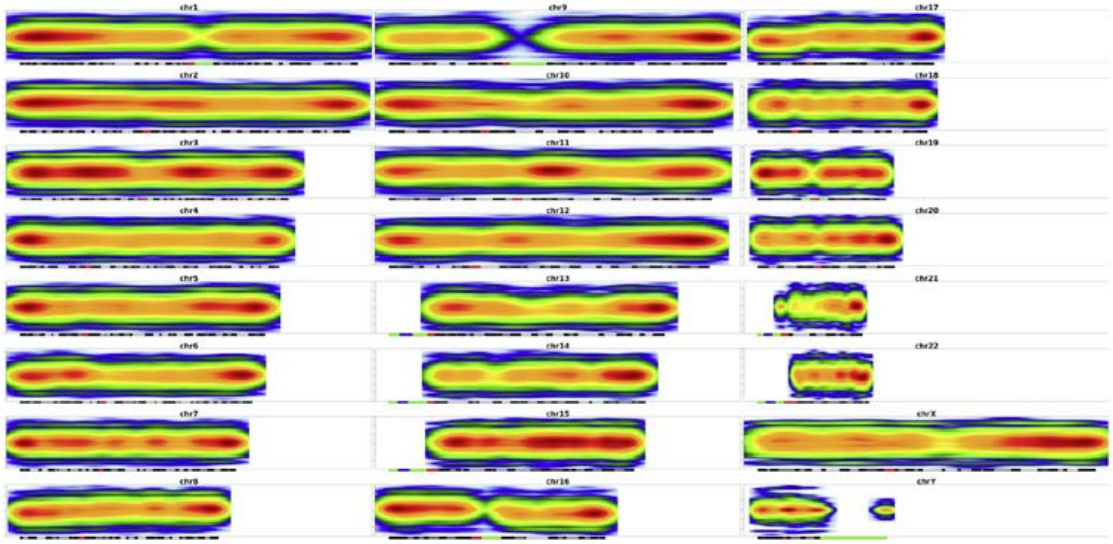


Figure 13: Genome-wide distribution of differential methylations in prostate cancer, shown are individual chromosomes with red indicating high accumulation of methylations and blue low levels.

A large proportion of PCs harbor gene fusions involving members of the ETS family and the androgen regulated *transmembrane protease serine 2* (*TMPRSS2)* gene, most commonly involving the *v-ets erythroblastosis virus E26 oncogene homolog ERG* that is observed in approximately 50% of all PC cases. The overexpression of *ERG* is thought to be sufficient for the initiation of PIN (prostate intraepithelial neoplasia) lesions, a precursor of PC. Other rearrangements are less frequent and tend to be present in PCs already harboring the *TMPRSS2:ERG* gene fusion (FUS+). This suggests that other molecular mechanisms than translocations like alterations in the methylation or gene expression pattern must play a driving role in the *TMPRSS2:ERG* negative (FUS-) subclass.
Using a MeDIP-Seq approach on 51 tumor (20 TMPRSS2:ERG fusion negative (FUS-); 17 TMPRSS2:ERG fusion positive (FUS+)) and 53 normal samples we identified 147.000 differentially methylated regions comparing tumour and normal samples of which marker sets for future prostate cancer detection could be derived and successfully tested in independent sample sets. Most importantly, comparing FUS+ and FUS- samples revealed a significantly altered methylation pattern in FUS- cancers, while FUS+ samples were more equal to normal samples. Interestingly, we found EZH2 (enhancer of zeste homolog 2) – a polycomb

group gene – significantly up-regulated in tumour samples. Increased expression can be explained by ERG in FUS+ samples while in FUS-samples we found miR26a – a suppressor of EZH2 - significantly down-regulated. We could show that hypermethylation of a 2kb region near miR26a is causative for miR26a supppression in FUS-samples. Thus, we developed a model for prostate tumour formation: In FUS+ cells, ERG overexpression results in overexpression of oncogenes like MYC, and EZH2 causing hypermethylation of homeobox genes leading to a reversion of differentiation and tumour formation. On the other hand FUS-samples exhibit a methylator phenotype accompagnied by hypermethylation of regulatory microRNA genes like miR26a (suppressing EZH2) or miR34 (suppressing MYC). Suppression of the regulatory microRNAs results in overexpression of MYC and EZH2, the latter augmenting aberrations in the DNA methylation pattern. Next steps are now to identify causes and consequences of the differential methylation patterns in FUS-samples. Our goal is to identify modifier enzymes responsible for the aberrant methylation pattern and to investigate mechanisms reverting the observed phenotype.

Identification of genetic and epigenetic alterations underlying colon cancer progression

Colon cancer (CRC) is the third most common cancer type worldwide and in 2004 over 1 million new cancer cases have been diagnosed. Thus, one focus of the group is to identify colon cancer progression markers e.g. methylation markers and to set them in relation to the pathomechanism of tumour formation and progression. We have already identified a set of biomarkers – miRNAs as well as differentially methylated regions (cDMRs) which we are validating. We have picked the candidate biomarkers out of genome-wide screens which we had performed on a large cohort of CRC cases. Now, having these data sets completed, we are in the privileged position to investigate them more intensively and to develop a picture of CRC alterations. Based on our extensive data sets on gene and miRNA expression, mutations and methylation alterations, we are now working on an integrative view of colorectal cancer. Using diverse bioinformatics visualization tools we try to get hold on genome-wide pathogenic alterations and to identify a temporal order of modifications.

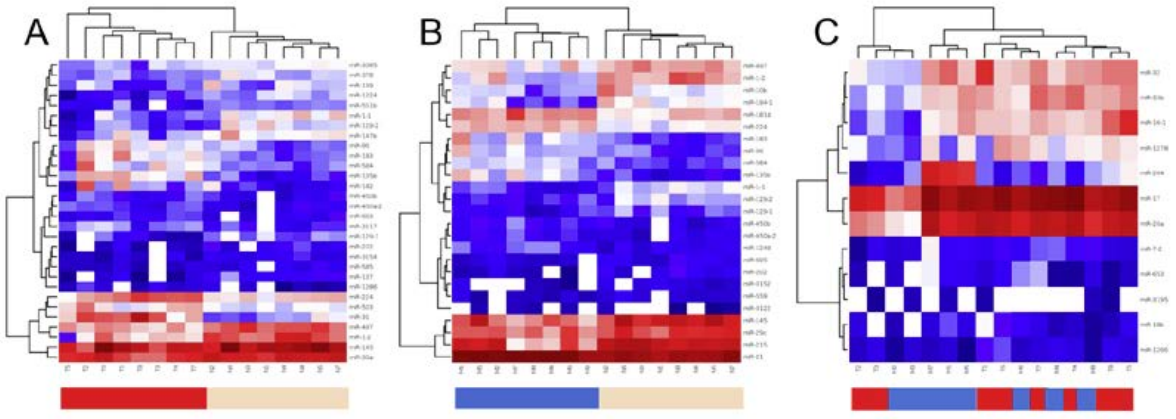


Figure 14: Epigenetic alterations can be used for a discrimination of colorectal normal, tumour and metastasis tissues. Heatmaps for (A) tumour vs. normal, (B) metastasis vs. normal, (C) metastasis vs tumour (bars below the heatmaps: red indicates tumour samples, pink normal, blue metastasis tissues.

Functional analyses of cervical cancer pathogenesis and development of antiviral substances

Cervical cancer is the second most common cancer among women with a worldwide prevalence of 2,27 millions and with an estimated 490,000 new cases and 270,000 deaths per year. “High risk” human papillomaviruses are detected in over 98% of cervical carcinoma cases. Together with Prof. Dr. Peter Howley (Harvard Medical School, Boston) we have found that the interaction between the viral E2 protein and the cellular bromodomain containing protein Brd4 is required for the genome maintenance and the viral transcriptional regulation functions and thereby regulate cervical cancer pathogenesis. We are now working on specific inhibitors of this interaction and we are using NGS technologies for an investigation of functional consequences for cervical cancer pathogenesis. Our experiments show that a disruption of the interaction between E2 and Brd4 leads to a curing of cells from papillomavirus infections and thereby might prevent cervical cancer.
Brd4 with its two bromodomains binds to acetylated histones, pTEFb (positive transcription elongation factor) and is dealt as master regulator of the epigenetic memory. We have found that Brd4 plays a significant role in transcriptional regulation, but that it is also a central partner in the oxidative stress response. Preliminary experiments indicate that an epigenetic regulation of the defense mechanisms against reactive oxygen species functions over histone methylations.

Taken together, the goals of the Cancer Genomics Group are to integrate different fields of medicine, biology and natural science in order to better understand how tumor cells work and how carcinogenic processes are regulated. Knowledge of disease- relevant alterations in gene sequences and molecules of the metabolic network will reveal targets for effective diagnostic and therapeutic applications. With the availability of these techniques we are on a turning point of cancer diagnosis and treatment.

Selected publications

Kerick M[Θ], Isau M[Θ], Timmermann B, Sueltmann H, Krobitsch S, Bartsch G, Schaefer G, Klocker H, Lehrach H, Schweiger MR (2011). *Technical challenges of targeted DNA Enrichment to Identify Genetic Variations in Cancer Patients. BMC Medical Genomics* 4:68 [Θ] shared co-first authors

Timmermann B[Θ], Kerick M[Θ], Röhr C, Fischer A, Isau M, Boerno S, Wunderlich A, Barmeyer C, Seemann P, Koenig J, Lappe M, Kuss AW, Garshasbi M, Bertram L, Trappe K, Werber M, Herrmann BG, Zatloukal K, Lehrach H, Schweiger MR (2010). *Somatic mutation profiles of MSI and MSS colorectal cancer identified by whole exome next generation sequencing and bioinformatics analysis.* PLoS One 5(12):e15661 . [Θ] shared co-first authors.

Krawitz PM[Θ], Schweiger MR[Θ], Rödelsperger C, Isau M, Fischer A, Dahl A, Jonske de Condor B, Kölsch U, Meisel C, Kinoshita T, Murakami Z, Hecht J, Brunner H, Meinecke P, Horn D, Mundlos S, Robinson PN (2010). *Identity-by-Descent Filtering of Exome Sequence Data identifies PIGV mutations in Hyperphosphatasia Mental Retardation (HPMR) syndrome*. Nat Genetics 42(10):827-9. [Θ] shared co-first authors.

Systems Biology Group

(Established: 01/2009)

Head
Christoph Wierling*, PhD
(since 01/09)
Phone: +49 (0)30 8413-1272
Fax: +49 (0)30 8413-1380
Email: wierling@molgen.mpg.de

Scientists
Thomas Kessler* (since 10/10)
Andriani Daskalaki* (since 01/09)
Hendrik Hache* (since 01/09)
Alexander Kühn* (08/09-02/12)
Svetlana Peycheva* (09/11-12/11)
Elisabeth Maschke-Dutz*
(01/09-12/11)

PhD students
George Davidescu* (since 10/11)
Vikash Pandey (since 08/10)
Jian Li (since 04/10)

Scientific overview

Research concept

Mathematical modeling and simulation techniques have turned out to be valuable tools for the understanding of complex systems in different areas of research and engineering. In recent years, this approach has come to application frequently also in biology, resulting in the establishment of the research area systems biology. Systems biology tries to understand the behaviour of complex biological systems by means of mathematical approaches. This requires the integration of qualitative and quantitative experimental data into coherent models. A challeng-

* externally funded

ing task is the development of comprehensive models that can be used in medical and pharmaceutical research for the establishment of a personalized medicine.
The Systems Biology Group has its research interests in the mathematical modeling of cellular processes with respect to complex diseases, such as cancer. Within the group different systems biology resources and tools for the modeling and simulation of biological systems have been designed and implemented. These tools are used in current projects for the modeling of cancer-related signal transduction and metabolic pathways, their subsequent gene regulatory network, and the effect of mutations and drugs. Moreover, the group is working on the modeling of stem cell biology. The research is driven by the integration of diverse 'omics data, as generated by current (high-throughput) technologies.

Scientific methods and findings

The Systems Biology Group is hosting the PyBioS modeling and simulation system (http://pybios.molgen.mpg.de). PyBioS has a web-based user interface (Figure 15) and it makes use of well established methods for the mathematical description of biochemical reaction systems based on ordinary differential equation systems and Petri nets, and novel interfaces to biochemical pathway databases (e.g., Reactome, KEGG, ConsensusPathDB). In addition PyBioS provides several functionalities for model analysis and visualization. Moreover, the systems biology group is involved in the development of the ConsensusPathDB database that is hosted by the Bioinformatics Group of Ralf Herwig.

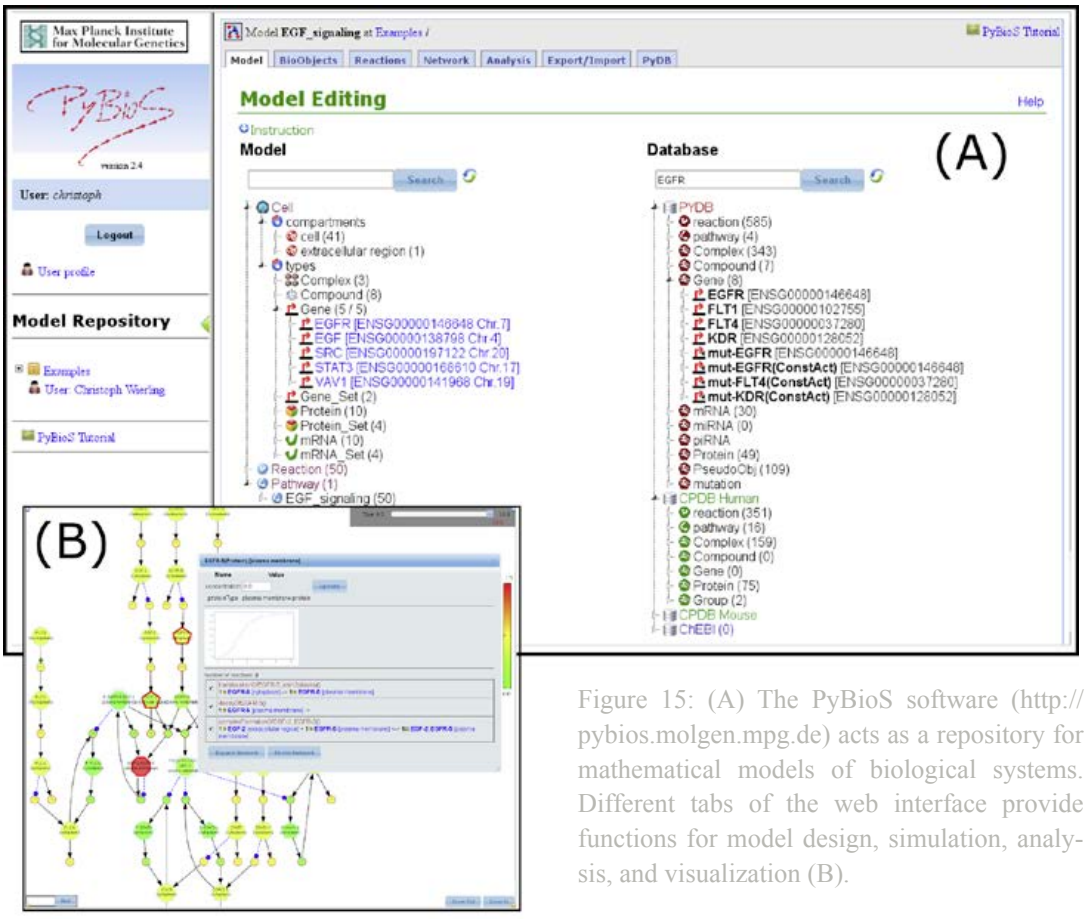


Figure 15: (A) The PyBioS software (http://pybios.molgen.mpg.de) acts as a repository for mathematical models of biological systems. Different tabs of the web interface provide functions for model design, simulation, analysis, and visualization (B).

Structure and behavior of any cell and any organism are determined by converting information in the genome and the environment into the phenotype through a series of molecular processes. Dysfunctions in the molecular interaction network can cause sever diseases such as cancer. Curing the disease often involves by itself complex disturbances in these networks. Progress in the treatment of tumours in individual patients will depend critically on being able to predict the effects of such treatments in the context of the genome involved. The development of predictive models is however complicated by the lack of information on many of the reaction kinetics needed. Information on the kinetics and kinetic parameters is either not available at all, or, at best, is based on experiments often carried out under conditions quite different from those in living cells. Thus, computational modeling approaches must primarily face the challenge of coping with this lack of information. One approach to overcome this limitation can be a rigorous analysis of the model's parameter space, e.g., by sampling unknown parameters from appropriate random distributions and a subsequent statistical evaluation. Such a kind of Monte Carlo-based approach makes it necessary to run thousands of simulations and thus can only be performed using distributed computing. This approach has been developed and implemented within the Systems Biology Group.

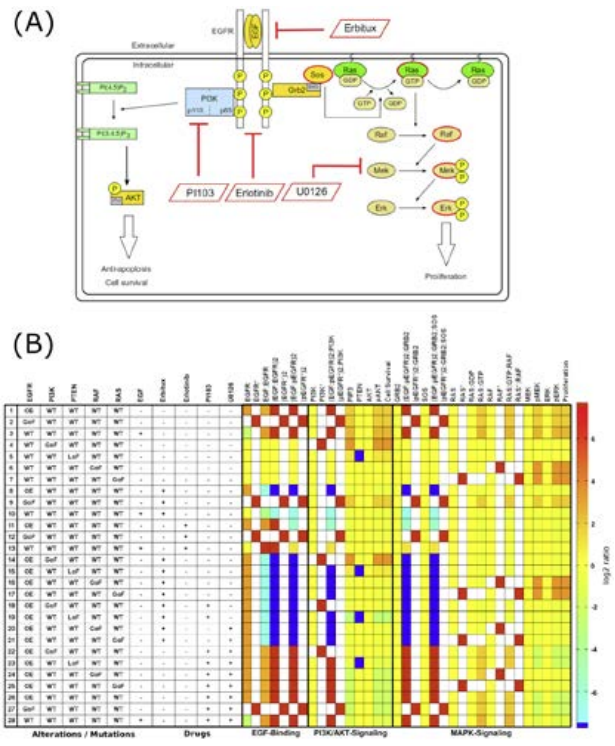


Figure 16: (A) Model of EGF-mediated signaling. (B) Heat map of simulation results of the Monte Carlo-based simulation approach of the EGF ligand, over-expression (EGFR), gain of function (EGFR, PI3K, RAS, RAF) and loss of function (PTEN) mutations, and different drugs on specific model components. Colors indicate log2-rations of the steady state concentrations versus a control state of low level activity in cell survival and proliferation (from Wierling et al., 2012).

In the course of current research projects (MUTANOM, MoGLI, TREAT20) the Systems Biology Group has established a large model of cancer-related pathways. Currently, the model comprises more than 2800 components and 4400 individual reactions and it covers different signaling pathways, such as EGF-, IGF-, NGF-, Wnt-, Notch-, Hedgehog-, Fas-, Trail-signaling, etc. Furthermore, the model has been extended by the integration of validated microRNA (miRNA) target information and subsequently it was used to analyze the impact of specific miRNA inhibitors. In cooperation with the groups of M.-R. Schweiger, M.-L. Yaspo, and B. Lange the model is used to study the effects of cancer-related somatic mutations and the identification of reasonable intervention points (Figure 16). Moreover, within the MoGLI project we study the transcriptional program and the molecular circuitries regulated by Hedgehog (HH) and GLI proteins (Figure 17). The results will improve our understanding of the complex molecular networks regulated by oncogenic HH/GLI signaling and will accelerate the search for novel molecular targets that represent an opportunity for therapeutic intervention.

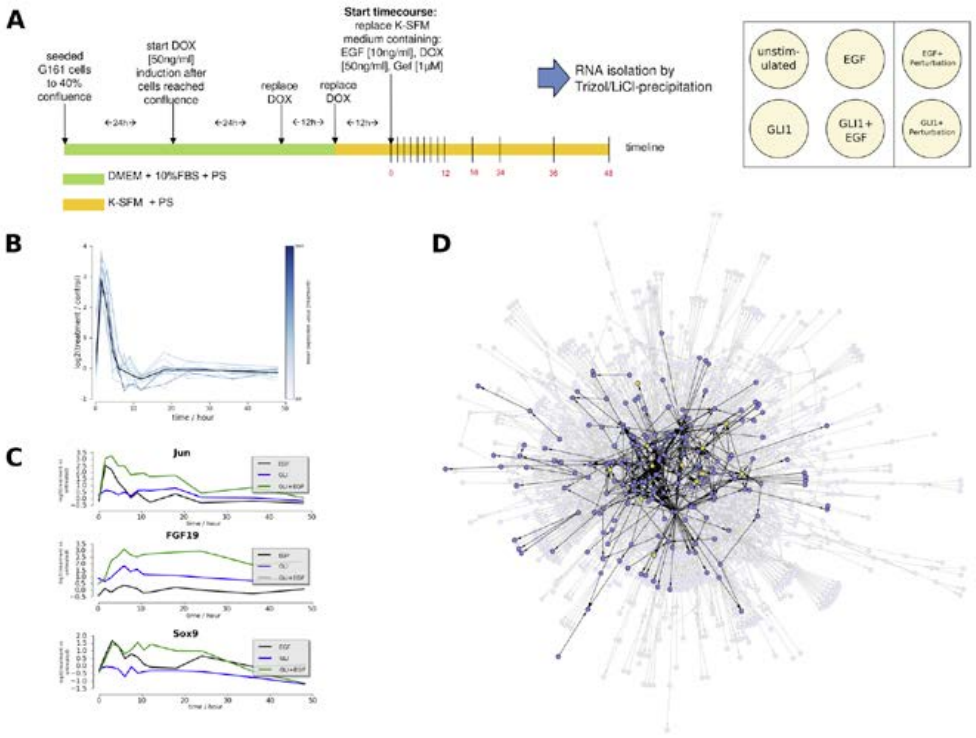


Figure 17: (A) Genome-wide time course expression data was used to identify synergistic effects of Hedgehog signaling with other pathways, such as EGF-mediated signaling. (B) Clusters of co-regulated genes (C) and individual synergistic genes were identified. (D) Further integration of gene regulatory data was used for the development of network models of differentially and synergistically regulated genes.

Within the IMGuS project on systems biology of steatohepatitis, a model of liver-related metabolic processes has been developed that is further analyzed and extended also in the LivSYSiPS project on stem cell reprogramming and differentiation related to steatohepatitis (in cooperation with J. Adjaye). The established resources, tools, algorithms, and models build a foundation for the application of systems biology strategies in medical and pharmaceutical research and, based on data from high-throughput genome, transcriptome, and proteome analysis, it enables the development of a personalized medicine.

Selected publications

Wierling C, Kühn A, Hache H, Daskalaki A, Maschke-Dutz E, Peycheva S, Li J, Herwig R, Lehrach H (2012). *Prediction in the face of uncertainty: a Monte Carlo-based approach for systems biology of cancer treatment.* Mutation Research 746(2):163-70

Li J, Pandey V, Kessler T, Lehrach H, Wierling C (2012). *Modeling of miRNA and Drug Action in the EGFR-Signaling Pathway.* PLoS ONE 7(1):e30140

Kamburov A, Pentchev K, Galicka H, Wierling C, Lehrach H, Herwig R (2011). *ConsensusPathDB: toward a more complete picture of cell biology*. Nucleic Acids Res 39:Database issue D712-717.

Research Group Gene Regulation and Systems Biology of Cancer

(Established: 1996)

Head
Dr. Marie-Laure Yaspo (since 1996)
Phone: +49 (0)30 8413-1356
Fax: +49 (0)30 8413-1380
Email: yaspo@molgen.mpg.de

Scientists
Marc Sultan* (since 09/07)
Hans-Jörg Warnatz* (since 05/09)

PhD students
Hans-Jörg Warnatz* (10/07-04/09)

Bioinformaticians
Vyacheslav Amstislavskiy (since 08/11)
Karin Schlangen* (since 08/11)
Meryem Avci* (since 01/09)
Madeleine Kittner* (01/12-05/12)
Henning Stehr* (07/11-12/11, guest)
Tina Stehr* (07/11-12/11, guest)

Engineers and technicians
Agnes Wolf* (since 06/11)
Katharina Kim* (since 01/11)
Alexander Kovacsovics* (since 01/11)
Matthias Linser* (since 01/11)
Simon Dökel* (since 04/10)
Daniela Balzereit* (since 02/02)
Sabine Schrinner* (01/02-09/12)
Asja Nuernberger* (10/05-06/12)
Kerstin Schulz* (08/10-12/11)
Emilie Dagand* (07/07-12/11)

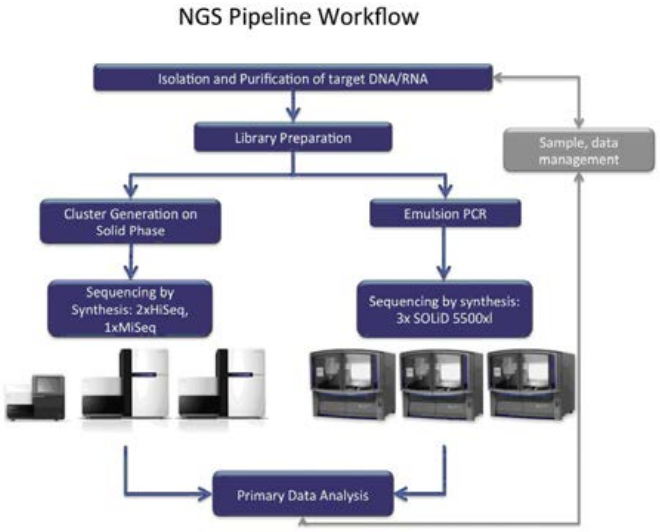
Scientific overview

The two EU projects initiated before 2009 have been completed: *EURExpress*, the Mouse expression atlas at E14.5 by non-radioactive ISH, where we generated > 3,000 expression patterns and ~50,000 images (www.eurexpress.org) and *ANEUploidy,* contributing a functional analysis of human chromosome 21 gene promoters in HEK293 cells, the identification of biological networks regulated by chromosome 21 transcription factors, e.g. BACH1 regulating oxidative stress and cell cycle, and the transcriptome response triggered by trisomy 21 in lymphoblastoid cells.
The group focuses its research activities on: 1) analysis of gene regulatory networks and 2) cancer genomics and systems biology of cancer. As these projects rely heavily on new generation sequencing (NGS), we have established a se-

* externally funded

quencing pipeline in the Lehrach department. One of the major strengths of our setup is the flexibility in protocol settings customized for each project, and the close interaction between sequence production and data analysis, optimizing logistics and complex data processing.

The *NGS pipeline* operates whole genome (paired-end and mate pairs) and exome sequencing, mRNA, miRNAseq, ChIPseq, and MeDip-seq. We enforced quality controls throughout experimental and data analysis processes, for instance the primary analysis includes sequence read mapping and data check reports (e.g. library complexity) that are automatically loaded in our sequencing database.



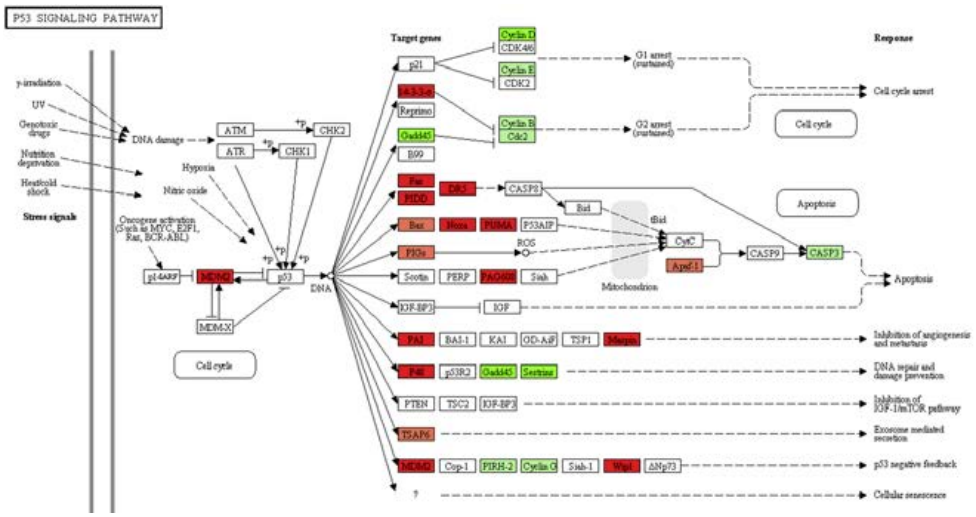
In addition to supporting the core lab projects, the sequencing pipeline carries out WGS for the 1,000 Genomes (http://www.1000genomes.org/), RNA- and mi-RNAseq from selected 1000 genomes samples in Geuvadis (www.geuvadis.org) and participates to the ESGI infrastructure (http://www.esgi-infrastructure.eu/), pooling efforts of European genomics and bioinformatics facilities (ESGI-PromElegan: nuclear capped RNAs, DNAseI hypersensitivity, ChIPseq from *C. elegans* samples; ESGI-CCGENSEQ: RNAseq for the collaborative mouse strain collection).

In parallel, we have built the *NGS analysis* for our core projects addressing: 1) Processing of raw data, management and quality monitoring. 2) Expression profiling, alternative splicing and gene fusions. 3) Identification of DNA variants and rearrangements. We use a variety of open source programs as well as customized steps to optimize the sensitivity/specificity threshold in SNV calls, in particular for samples with low tumor cellularity tumour. A strong emphasis is put on the integration of genome and transcriptome information, and on the prediction of damaging mutations.

Analysis of Gene Regulation Networks

Systems Biology of Steatohepatitis (www.imgus.at) coordinated by Kurt Zatlokal, Graz, analyses genetic mouse models of this metabolic liver disease. We are using consomic strains and exploiting the differential response of B6, PWD and A/J mouse strains to steatohepatitis inducing agents for identifying molecular determinants of the disease by means of RNAseq in consomic and parental strains of two panels, “B6 x PWD” and “B6 x A/J”. Data are currently being analysed in cooperation with the Zatlokal lab (Graz) and the group of C. Wierling in the department.

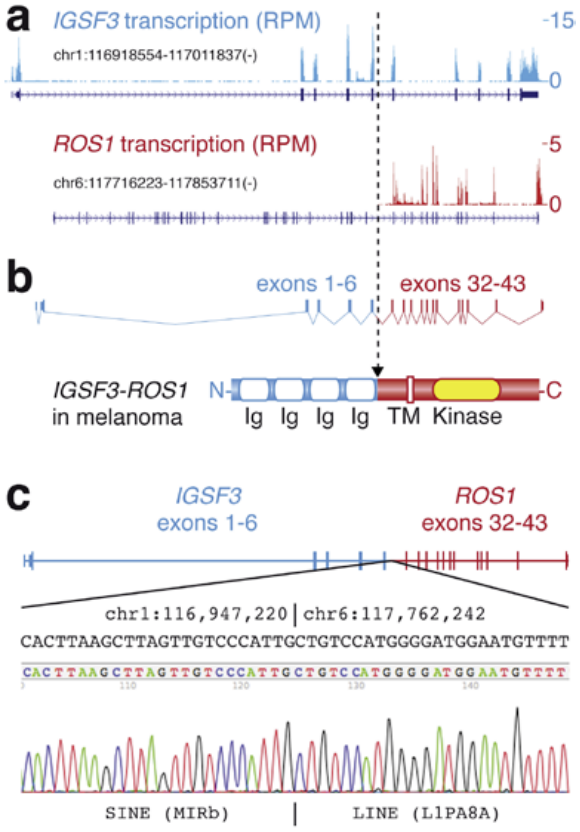
The European project *TRIREME* (Tackling the Response to Ionizing Radiation by Extensive Multilevel Exploration; http://www.tau.ac.il/medicine/trireme/) coordinated by Yossi Shilo, Tel Aviv University, aims at understanding signalling pathways associated to the DNA damage response (DDR). We investigated the gene expression dynamics in human CAL51 breast cancer cells subjected to ionizing radiations, a potent inducer of double-strand breaks. We performed ChIP-seq for six context-relevant transcription factors (TP53, RELA, E2F7, EGR1, ATF3 and CREB1) and for histone H3K4 mono- and trimethylation. We identified TP53 as a major regulator of DDR in time series experiments, and observed that chromatin



conformation played an important role in gene expression changes induced by DNA double-strand breaks. H3K4 trimethylation marks at transcription start sites correlated with increased gene activity, particularly those activated by DNA damage. We also identify a H3K4 monomethylation pattern (characterizing enhancer regions) among ~1,800 TP53 target regions already marked before irradiation. TP53-binding triggered by ionizing radiation lead to ATM-dependent changes in this monomethylation pattern. Further, we identified microRNAs induced by ionizing radiation, such as for two miRNAs located in the TP53-induced genes ANK1 and GDF15, showing ATM-dependent increase in H3K4me3. We are now extending this study for wild-type and TP53-deficient cells.

Systems biology of cancer

The analysis power of NGS has allowed a steep increase in the identification of somatic events associated to many types of cancer, where most studies focused on the cataloguing of DNA variants. Building on our experience with RNAseq, we are combining WGS, exome sequencing and expression analyses in order to identify expressed relevant mutations. We are also taking into account DNA ploidy and tumor cellularity. The central goal of the group is to gain insight in the biology of individual cancers, based on quantitative gene expression data. We are part of the **I**nternational **C**ancer **G**enome **C**onsortium (ICGC) aiming at deciphering the molecular landscape of cancer cells. In ICGC-*PedBrain* (http://www.pedbraintumor.org/), we are analysing pediatric medulloblastomas and astrocytomas by RNAseq, addressing mutant allelic expression and gene fusions.



Several recurrent mutations were identified, both in known medulloblastoma-related genes (CTNNB1, PTCH1, MLL2, SMARCA4) and in genes not previously linked to this tumour (DDX3X, CTDNEP1, KDM6A, TBR1). RNA-sequencing confirmed these alterations, and revealed the expression of the first medulloblastoma fusion genes. In Early Onset Prostate Cancer (http://www.icgc.org/icgc/cgp/70/345/53039), we contribute RNAseq and WGS, the integrative sequencing approaches of 11 cases revealed characteristic molecular pathomechanisms, associated with the accumulation of androgen-dependent genomic rearrangements, including gene fusion events affecting androgen receptor and WNT signaling pathways.

Innovative therapeutics options might be predicted based on the global molecular characteristics of individual tumors. *TREAT20* is a joint translational medicine project with Prof. P. Schlag (CCC, Charité, Berlin) aiming at predicting optimized therapies for 20 cases of metastatic melanomas of various origins. We observed a wide heterogeneity of DNA aberrations in different melanoma subtypes, as well as chromothripsis, and novel driver events such as translocations involving the ROS1 oncogene. In contrast, expression profiles showed strong similarities, but with striking patient-specific events. The data are used to fuel the cancer modelling system developed by C. Wierling and A. Kühn, predicting treatment outcome, which is then tested on cells and tumor xenographs (R. Schäfer, Charite and EPO, Berlin partners in TREAT20).
We are extending this concept to epigenetic marks and proteomics in the *EU-Oncotrack* (www.oncotrack.org) initiative, coordinated by H. Lehrach that aimed at isolating novel cancer diagnostics in colon cancer (see above). Further, we recently started the *BLUEPRINT of Haematopoietic Epigenomes project* **(**http://www.blueprint-epigenome.eu/**)** where we contribute RNAseq.

Planned developments

Analysing alternative splicing in medulloblastoma and melanoma, understanding cancer pathways, testing predicted activated pathways with proteomics, and drug effects on patient-derived cell lines and xenographs.

Cooperation within the institute

Within the MPIMG, the Research Group Gene Regulation and System Biology of Cancer cooperates with the following people and their departments/groups:

- Alexey Soldatov (NGS protocols)
- Ralf Herwig, Axel Rasche (alternative splicing RNAseq)
- Bernhard Hermann (Treat20)
- Stefan Mundlos (Eurexpress)
- Bernd Timmerman (Oncotrack)
- Martin Vingron (Blueprint, alternative splicing)
- Christoph Wierling (cancer modelling)

General information

Complete list of publications (2009-2012)

2012

Jones DTW, Jäger N, Kool M, Zichner T, Hutter B, Sultan M, Cho YJ, Pugh TJ, Hovestadt V, Stütz AM, Rausch T, Warnatz HJ, Ryzhova M, Bender S, Sturm D, Pleier S, Cin H, Pfaff E, Sieber L, Wittmann A, Remke M, Witt H, Hutter S, Tzaridis T, Weischenfeldt J, Raeder B, Avci M, Amstislavskiy V, Zapatka M, Weber UD, Wang Q, Lasitschka B, Bartholomae C, Schmidt M, von Kalle C, Ast V, Lawerenz C, Eils J, Kabbe R, Benes V, van Sluis P, Koster J, Volckmann R, Shih D, Betts MJ, Russell RB, Coco S, Tonini JP, Schüller U, Hans V, Graf N, Kim YJ, Monoranu C, Roggendorf W Unterberg A, Herold-Mende C, Milde T, Kulozik AE, von Deimling A, Witt O, Maass E, Rössle, J, Ebinger M, Schuhmann MU, Frühwald MC,

Hasselblatt M, Jabado N, Rutkowski S, von Bueren AO, Williamson D, Clifford SC, McCabe MG, Collins VP, Wolf S, Wiemann S, Lehrach H, Brors B, Scheurlen O, Felsberg J, Reifenberger G, Northcott PA, Taylor MD, Meyerson M, Pomeroy SL, Yaspo ML Korbel JO, Korshunov A, Eils R, Pfister SM, Lichter P (2012). *ICGC PedBrain: Dissecting the genomic complexity underlying medulloblastoma*, Nature, in press

Sultan M, Dökel S, Amstislavskiy V, Wuttig D, Sültmann H, Lehrach H, Yaspo ML (2012). *A simple strand-specific RNA-Seq library preparation protocol combining the Illumina TruSeq RNA and the dUTP methods*. Biochem Biophys Res Commun 422(4):643-6

2011

Diez-Roux G, Banfi S, Sultan M, Geffers L, Anand S, Rozado D, Magen A, Canidio E, Pagani M, Peluso I, Lin-Marq N, Koch M, Bilio M, Cantiello I, Verde R, De Masi C, Bianchi SA, Cicchini J, Perroud E, Mehmeti S, Dagand E, Schrinner S, Nürnberger A, Schmidt K, Metz K, Zwingmann C, Brieske N, Springer C, Hernandez AM, Herzog S, Grabbe F, Sieverding C, Fischer B, Schrader K, Brockmeyer M, Dettmer S, Helbig C, Alunni V, Battaini MA, Mura C, Henrichsen CN, Garcia-Lopez R, Echevarria D, Puelles E, Garcia-Calero E, Kruse S, Uhr M, Kauck C, Feng G, Milyaev N, Ong CK, Kumar L, Lam M, Semple CA, Gyenesei A, Mundlos S, Radelof U, Lehrach H, Sarmientos P, Reymond A, Davidson DR, Dollé P, Antonarakis SE, Yaspo ML, Martinez S, Baldock RA, Eichele G, Ballabio A (2011). *A high-resolution anatomical atlas of the transcriptome in the mouse embryo*. PLoS Biol 9(1):e1000582

Salton M, Elkon R, Borodina T, Davydov A, Yaspo ML, Halperin E, Shiloh Y (2011). *Matrin 3 binds and stabilizes mRNA*. PLoS One 6(8):e23882

Warnatz HJ, Schmidt D, Manke T, Piccini I, Sultan M, Borodina T, Balzereit D, Wruck W, Soldatov A, Vingron M, Lehrach H, Yaspo ML (2011). *The BTB and CNC homology 1 (BACH1) target genes are involved in the oxidative stress response and in control of the cell cycle.* J Biol Chem 286(26):23521-32

2010

Cheng X, Guerasimova A, Manke T, Rosenstiel P, Haas S, Warnatz HJ, Querfurth R, Nietfeld W, Vanhecke D, Lehrach H, Yaspo ML, Janitz M (2010). *Screening of human gene promoter activities using transfected-cell arrays*. Gene 450(1-2):48-54

Richard H, Schulz MH, Sultan M, Nürnberger A, Schrinner S, Balzereit D, Dagand E, Rasche A, Lehrach H, Vingron M, Haas SA, Yaspo ML (2010). *Prediction of alternative isoforms from exon expression levels in RNA-Seq experiments.* Nucleic Acids Res 38(10): e112

Warnatz HJ, Querfurth R, Guerasimova A, Cheng X, Haas SA, Hufton AL, Manke T, Vanhecke D, Nietfeld W, Vingron M, Janitz M, Lehrach H, Yaspo ML (2010). *Functional analysis and identification of cis-regulatory elements of human chromosome 21 gene promoters*. Nucleic Acids Res 38(18):6112-23

PhD Thesis

Hans-Jörg Warnatz: *Systematic cloning and functional analysis of the proteins encoded on human chromosome 21*. Freie Universität Berlin, 2009

Former Groups of the Department / Associated Research Groups

Molecular Embryology and Aging Group

(07/2001-04/2012)

Head
Dr. James Adjaye
now at Heinrich Heine University, Düsseldorf
Institute for Stem Cell Research and Regenerative Medicine
Phone: +49 (0)211 81-19545
Fax: +49 (0)211 81-19147
Email: adjaye@molgen.mpg.de

In May 2012, James Adjaye has accepted an appointment as Professor and Director of the Institute for Stem Cell Research and Regenerative Medicine, Heinrich Heine University, Düsseldorf, Germany

Scientists and postdocs
Katharina Drews
(née Wolfrum)* (since 06/12)
Justyna Jozefczuk* (since 02/06)
Ying Wang* (01/09-12/11)
Alessandro Prigione* (11/08-10/11)
Christian Regenbrecht* (03/09-02/11)
Raed Abu Dawud* (02/09-11/10)
Heiko Fuchs* (05/10-03/12)

PhD students
Peggy Matz* (since 07/12)
Barbara Mlody* (since 12/10)
Mathias Megges* (since 09/10)
Björn Lichtner* (since 02/10)

Katharina Drews
(née Wolfrum)* (10/08-05/12)

Technician
Elisabeth Socha* (since 05/10)

Scientific overview

Background

Terminal differentiation of embryonic progenitor cells to form distinct adult tissues is the hallmark of organogenesis in complex organisms. Stem cells, in particular embryonic stem cells, which can be expanded indefinitely and are pluripotent, have attracted considerable attention as a therapeutic approach for treating for example, diabetes, cardiovascular-, neurological-, and liver-based diseases. However, the use of inner cell mass-derived embryonic stem cells in cell replacement therapy remains problematic for a number of reasons, including ethical as well as host rejection of allogeneic cells. As a means of overcoming the problem

* externally funded

of host rejection, it has now been demonstrated by numerous laboratories that the combined expression of four transcription factors, OCT4, SOX2, NANOG and LIN28 or OCT4, SOX2, KLF4 and MYC is sufficient to reprogramme human or mouse somatic cells into induced pluripotent stem cells - iPS cells. These cells have normal karyotypes, express telomerase activity, express cell surface makers and genes that characterize ES cells, and maintain the developmental potential to differentiate into advanced derivates of all three primary germ layers. Patient-specific iPS cells can be used to derive tissue-matched differentiated donor cells for therapy, and a source of cells for research into the pathogenesis of complex diseases and also toxicology studies and drug screening.

The research in my group is divided into five inter-related areas:

(1) Transcriptional and signal transduction mechanisms regulating self renewal and pluripotency in human embryonic stem cells, carcinoma cells and iPS cells (induced pluripotent stem cells).

(2) Reprogramming of somatic cells (healthy and diseased individuals- Alzheimer's Diabetic, Nijmegen Breakage Syndrome and Steatosis patients) into an ES-like state (iPS cells) and studying the underlying disease mechanisms.

(3) Comparative characterization of functional hepatocytes and neuronal cells derived from human ES cells and patient-specific iPS cells with the aim of establishing a platform for toxicology studies and drug screens.

(4) Systems biology of stem cell fate and cellular reprogramming.

(5) Age-associated gene expression patterns and signal transduction mechanisms employing bone marrow-derived MSCs and fibroblast-derived iPS cells from young and aged individuals as model systems.

Current state of research and scientific findings

Knowledge gained from our earlier studies on OCT4-mediated gene regulatory networks and associated signalling pathways crucial for maintaining self-renewal and pluripotency in human embryonic stem cells has enabled us to efficiently derive and maintain induced pluripotent stem cells from somatic cells of distinct origins. In keeping with the new era of personalized medicine, we now derive (employing viral, episomal plasmids, mRNA and miRNA) and characterize iPS cells from various cell types (dermal fibroblasts amniotic fluid, chorionic villi and mesenchymal stem cells) of distinct ages and disease states (see Figure 21).

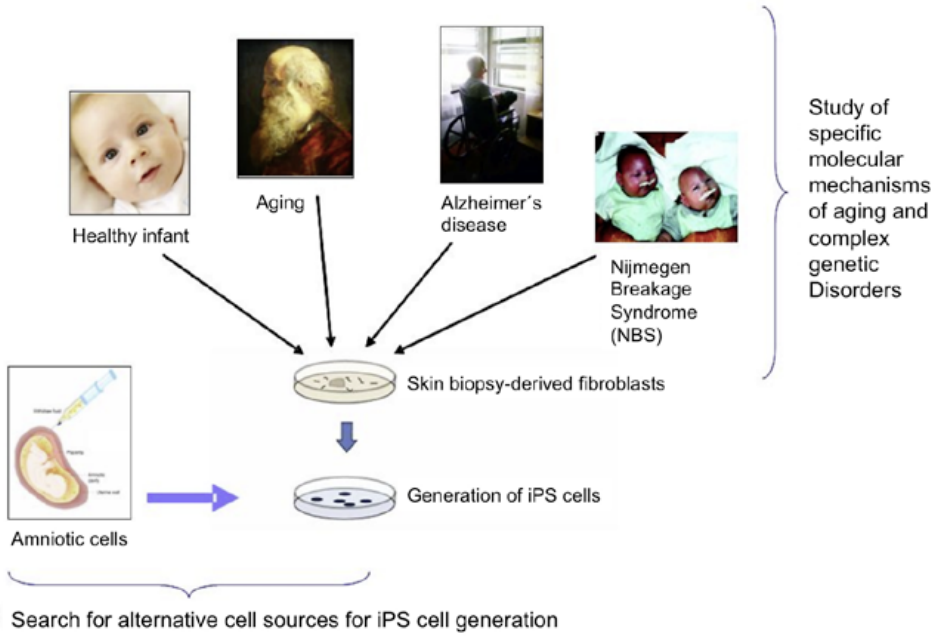


Figure 21: An illustration of the iPS based projects under investigation

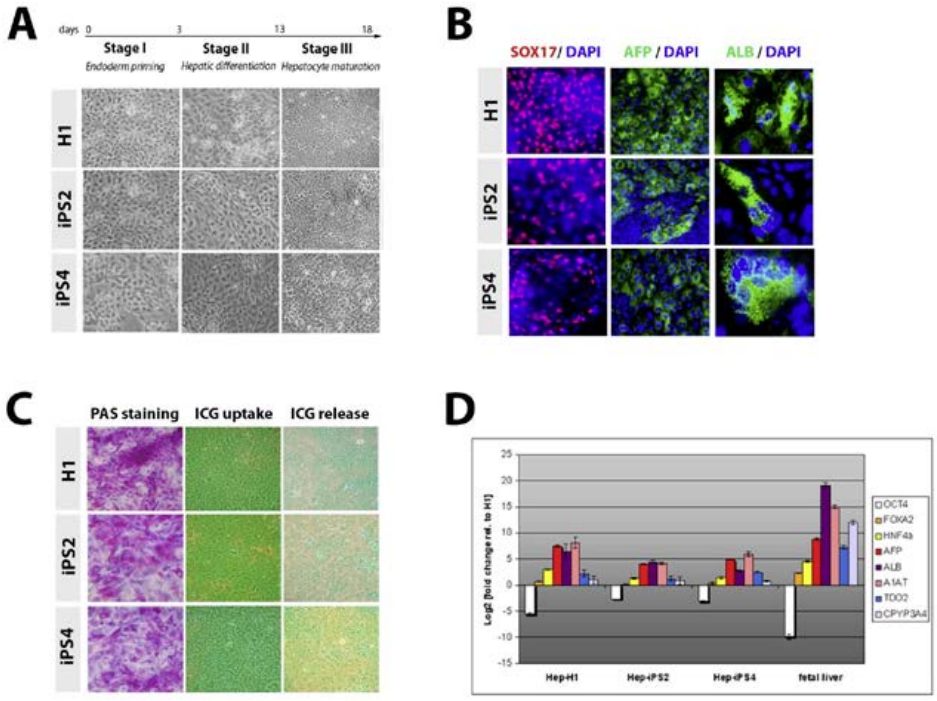


Figure 22: Characterisation of iPS-derived hepatocytes from human ES and iPS cells. (A) Phase contrast image showing similar cell morphologies. The differentiation protocol consists of endoderm priming induced by Activin-A, followed by hepatic differentiation and then maturation. This was applied to the embryonic stem cell line-H1 and two fetal foreskin fibroblast-derived iPS cell lines- iPS2 and iPS4 (Prigione et al. 2010) (B) Induction of the endoderm lineage and further differentiation to hepatocytes is confirmed by the expression of SOX17, AFP (Alpha fetoprotein) and ALB (Albumin). (C) Confirmation of glycogen storage by PAS staining, uptake of metabolites by ICG (indole cyano green) uptake and release. (D) Q-PCR demonstrating that ES and iPS-derived hepatocytes (Hep-H1 and Hep-iPS2, - iPS4) express similar complements of liver specific genes such as FOXA2, HNF4a, AFP, ALB, A1AT, TDO2 and CyP3A4.

We have successfully generated and characterised iPS cells from human skin-derived fibroblasts and comparatively differentiated these into hepatocyte-like cells (see Figure 22). iPS cells generated from individuals afflicted by diseases such as Alzheimer's, Type 1 Diabetes, Nijmegen Breakage Syndrome and Steatosis provide *in vitro* models for studying the mechanisms underlying these diseases. These studies are carried out in collaboration with relevant clinicians.
Our research on aging is based firmly on the concept that the same signalling mechanisms that regulate the plasticity of stem cells are altered during aging and in age-related diseases. We also attempt to reverse the aging phenotype of somatic cells by cellular reprogramming as a means to study genome stability, mitochondria biogenesis and stem cell fate. Accordingly, an understanding of molecular and signalling mechanisms underlying the aging process is likely to lead to novel approaches to preventing and treating age-related diseases.

Mechanisms underlying the induction of pluripotency in somatic cells

We derive (employing viral, episomal plasmids, mRNA and miRNA) and characterise iPS cells from distinct cell types (fibroblasts, amniotic fluid, chorionic villi and mesenchymal stem cells), ages and disease states. The main themes of investigation include

(1) Do iPS cells derived from distinct cell types retain cell type specific gene signatures?
(2) What is the long-term effect of these cell type specific gene signatures on self renewal and pluripotency?
(3) Modulating cellular homeostasis and mitochondrial bioenergetics as a means of enhancing the efficiency of deriving iPS cells.

These studies are intended to increase our meagre understanding of the mechanisms underlying cellular reprogramming.

Derivation and characterisation of hepatocyte-like cells from human ES and iPS cells

Drug development is a lengthy and costly procedure. Although animal models may represent valuable tools for toxicology studies, their application is hampered by the difficulties associated with the extrapolation of data to human, as several species-specific variations can be observed in response to drugs. The use of *ex vivo* adult primary human hepatocytes is the desirable option for safety evaluation of new compounds *in vitro*. However, this approach is undermined by the limited organ availability and the difficulty of developing efficient and reproducible protocols for the isolation of a pure population of primary hepatocytes. To circumvent these drawbacks, we employ ES and iPS cells to gain insights into hepatogenesis *in vitro*.

Induced pluripotent stem cells generated from human skin-derived fibroblast have been comparatively (with human ES cell lines H1 and H9) differentiated into hepatocyte-like cells. Detailed microarray-based gene expression analysis has been carried out to analyse dynamic regulatory events during hepatogenesis *in vitro*. We have also measured biochemical parameters specific to the liver, such as urea and albumin production and the uptake and release of metabolites.

This aspect of my research is currently funded by the BMBF under the framework of the transnational project ERASysBio Plus. I am the co-ordinator of LIVSYS-iPS, a multi-national project which uses a systems biology approach to investigate the etiology of non-alcoholic fatty liver disease (NAFLD) which comprises a broad spectrum of disease states ranging from manageable stress as in simple steatosis (S) to excessive stress as in steatohepatitis (SH). Here, we compare the properties of hepatocyte-like cells from healthy and patient skin-derived iPS cells. The knowledge gained from these studies will be invaluable for the early means of identifying drugs that cause side effects in patients and most importantly for understanding the molecular (genes and associated signalling pathways) mechanisms underlying the etiology of simple steatosis and steatohepatitis. A detailed description of the consortium can be seen on http://www.erasysbio.net/index.php?index=264.

Systems biology of stem cell fate and cellular reprogramming

The maintenance of pluripotency and self-renewal of human ES and iPS cells are intrinsically complex processes driven by the co-ordinated dynamic expression of a plethora of genes, their encoded proteins and associated signalling pathways in response to external signalling cues such as FGF2.

My aims are three-fold;

(1) *Network reconstruction*: Identify the gene networks involved in maintaining self-renewal of ES and iPS cells and also cell lineage specification.

(2) *Systems modelling: In silico* modeling of gene expression datasets (ES / iPS self-renewal perturbation experiments) generated in my lab will permit defining crucial components within the reconstructed networks supporting self-renewal and pluripotency.

(3) *Discrete function prediction*: Identification of protein domains of uncharacterised gene products that show similar profiles to OCT4, SOX2 and NANOG. The newly characterised proteins will be incorporated into the systems modeling and network reconstruction analyses.

Our systems biology approach combines high throughput approaches (ChIP-chip, ChIP-seq, RNAi, metabolomics and cytokine stimulations of hESCs) and advanced computational techniques to dissect the molecular mechanisms of stem cell fate and cellular reprogramming. Finally we have created a user-friendly database (http://biit.cs.ut.ee/escd/index.cgi?gene=oct4&p=0.05&entries) integrating all OCT4 and stem cell related datasets in both human and mouse ES and EC cells. This systems biology research activities are also funded by the BMBF Medical Systems Biology project: Drug-induced iPS cells and modelling reprogramming of somatic cells into induced Pluripotent Stem cells (BMBF - MedSYS project, for details see: http://dips.mdc-berlin.de/).

Selected publications

Jozefczuk J, Prigione A, Chavez L, Adjaye J (2011). *Comparative Analysis of Human Embryonic Stem Cell and Induced Pluripotent Stem Cell-Derived Hepatocyte-Like Cells Reveals Current Drawbacks and Possible Strategies for Improved Differentiation*. Stem Cells Dev. PMID: 21162674

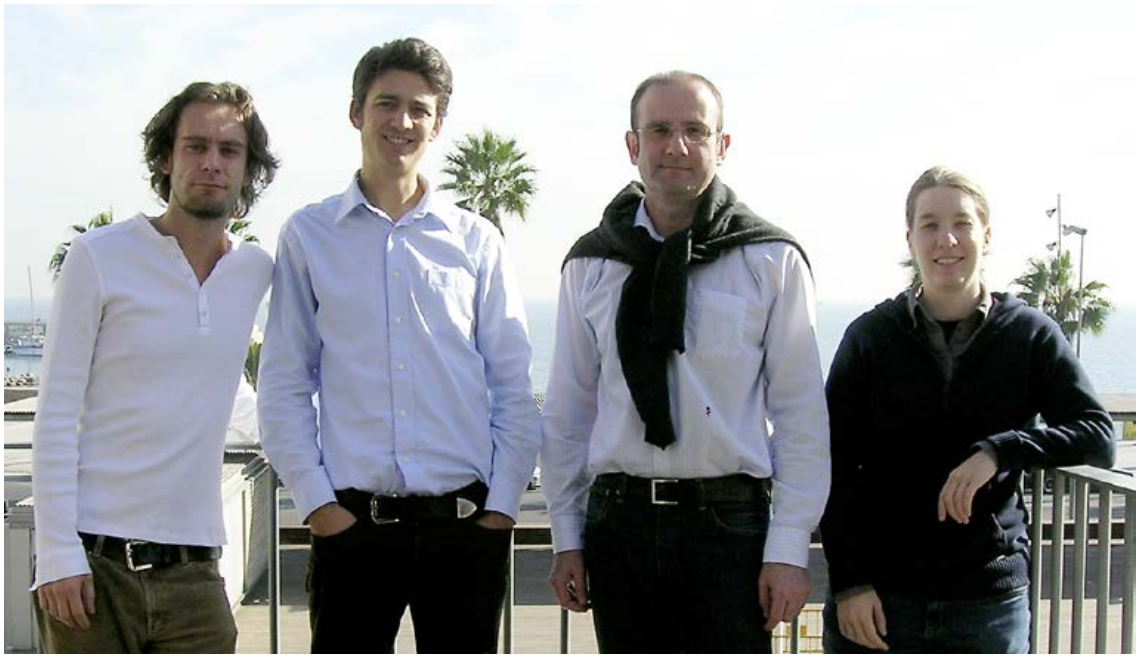
Wolfrum K, Wang Y, Prigione A, Sperling K, Lehrach H, Adjaye J (2010). *Defining the LARGE Principle of Cellular Reprogramming: Lost, Acquired and Retained Gene Expression in neonatal foreskin and amniotic fluid-derived iPS cells.* PLoS One 5(10):e13703.

Prigione A, Fauler B, Lurz R, Lehrach H, Adjaye J (2010). *The Senescence-Related Mitochondrial/Oxidative Stress Pathway is Repressed in Human Induced Pluripotent Stem Cells*. Stem Cells 4:721-733

Former Groups of the Department / Associated Research Groups

Comparative and Functional Genomics Group

(07/1995-05/2008)



Head

Dr. Heinz Himmelbauer
now at Centre for Genomic Regulation (CRG), Barcelona, Spain
Ultrasequencing Unit
Phone: +34 93 3160243
Email: Heinz.Himmelbauer@crg.es

The group moved to the Centre for Genomic Regulation (CRG) in June 2008, following Heinz Himmelbauer's appointment as head of the CRG Ultrasequencing Unit.

Scientists

Juliane Dohm* (since 10/08)
Robert Kofler* (08/08-07/09)
Darek Kedra* (12/09-12/10)

Engineers/Technicians

Hui Kang* (07/09-08/11)
Maik Zehnsdorf* (01/03-12/10)

PhD students

André Minoche* (since 01/09)
Cornelia Lange* (10/04-05/10)
Tobias Nolden* (05-04-08/10)

* externally funded

Scientific overview

Research concept

The focus of the group is in technology development (lab methods and data analysis tools) to support the efficient use of next generation sequencing (NGS) technologies. As central research project, we have sequenced and interpreted the genome of sugar beet, a crop plant only distantly related to other species with sequenced genomes. In addition, we pursue technology-driven cooperations in medical genomics, bioinformatics, and genetics in areas where innovative utilization of NGS is a strong asset.

Scientific methods and findings

Characterization of datasets from second generation sequencing

The generation and analysis of high-throughput sequencing data is becoming a major component of many studies in molecular biology and medical research. Illumina's Genome Analyzer (GA) and HiSeq instruments are currently the most widely used sequencing devices. We comprehensively evaluated properties of genomic HiSeq and GAIIx data derived from two plant genomes and one virus, with read lengths of 95 to 150 bases. We provide quantifications and evidence for GC bias, error rates, error sequence context, effects of quality filtering, and the reliability of quality values. By combining different filtering criteria we reduced error rates 7-fold at the expense of discarding 12.5% of alignable bases. While overall error rates were low in HiSeq data we observed regions of accumulated wrong base calls. Only 3% of all error positions accounted for 24.7% of all substitution errors. Analyzing the forward and reverse strands separately revealed error rates of up to 18.7%. Insertions and deletions occurred at very low rates on average but increased to up to 2% in homopolymers. A positive correlation between read coverage and GC content was found depending on the GC content range. The discovered errors and biases have implications for the use and the interpretation of Illumina sequencing data. Quality filtering is essential to minimize downstream analysis artifacts. Strand-specific biases provide a criterion to distinguish sequencing errors from low abundance polymorphisms.

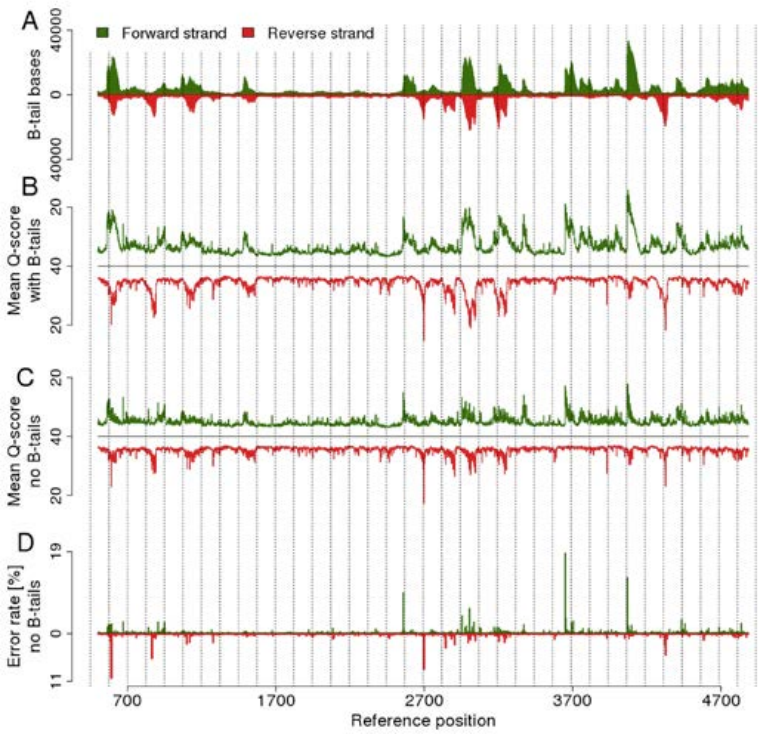


Figure 23: Biased distribution of low quality bases in PhiX read HiSeq data. Reads from forward and from reverse strands were analyzed separately (A) Number of bases within B-tails (consecutive bases of Q-score=2 at 3' end of a read) per position; (B) average Q-score of bases in untrimmed reads; (C) average Q-score of bases in B-tail trimmed reads; (D) observed per-base substitution error rate. Accumulation of low quality values is observed even after B-tail removal. Peaks of observed error rates occur at positions where increased low quality counts are detected, and in most cases peaks are seen only on one strand.

Development of a directional transcriptome sequencing protocol

Several studies support that antisense-mediated regulation may affect a large proportion of genes. However, methods available for transcriptome sequencing generally do not provide information on transcript orientation or are limited to the processing of short RNA fragments. Based on Illumina sequencing, we developed DSSS (Direct Strand Specific Sequencing), a strictly strand specific protocol for transcriptome sequencing. We tested DSSS with 150-200 nt single stranded RNA fragments from two samples, prokaryotic (*Mycoplasma pneumoniae*) as well as eukaryotic (*Mus musculus*), and obtained data containing strand specific information. We validated our results by comparison with a strand specific tiling array dataset for strain M129 of the simple prokaryote *M. pneumoniae*, and by quantitative PCR. Compared to arrays, DSSS provided higher dynamic range and single-base resolution, thus enabling efficient antisense detection and the precise mapping of transcription start sites and UTRs. Data for mouse confirmed the strand-specificity of the method and the general applicability of the approach to studying eukaryotic transcription.

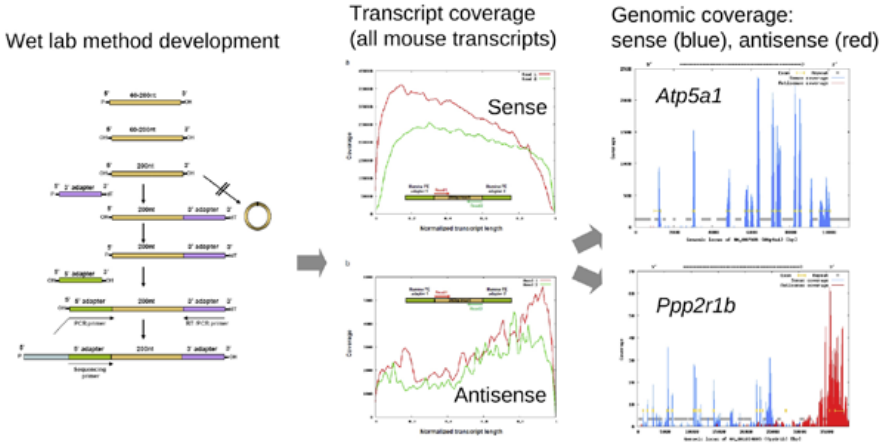


Figure 24: Outline of direct strand specific sequencing, a new method for generating directional transcript data.

Sequencing the genome of sugar beet (Beta vulgaris)

The genome of sugar beet (*Beta vulgaris*) has an estimated size of 758 Mbp in nine chromosomes and consists of about 65% repetitive sequence. We sequenced a homozygous line in a whole genome shotgun approach with data generated on 454, Illumina, and Sanger platforms. Single and paired read data were assembled

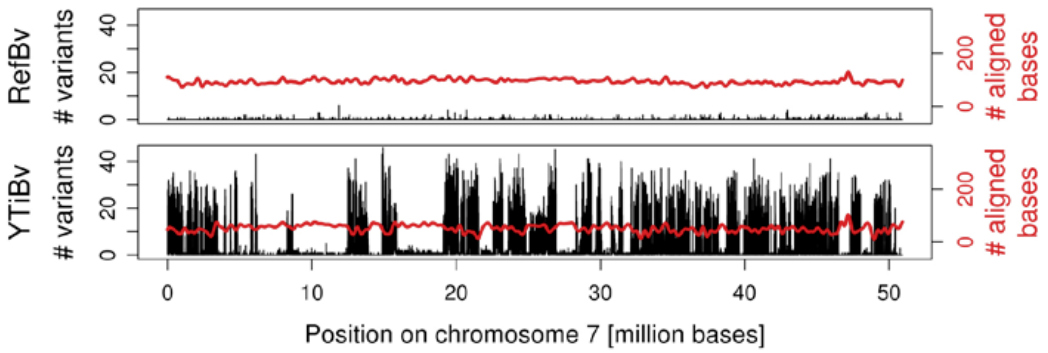


Figure 25: Comparison of sugar beet chromosome 7 (50 Mb) between two homozygous lines. Columns in black indicate SNV frequency per window of 2 kb. The red line indicates read coverage. The top panel shows a comparison between the reference genome to Illumina reads generated from the reference for background determination. The lower panel shows a comparison between the reference and another line. Conspicuous is the mosaic of regions exhibiting low or high sequence diversity.

into scaffolds using the Newbler software. Subsequent integration with genetic and physical maps of sugar beet substantially improved the scaffold size. Of the resulting 596 Mbp of sequence data two-thirds could be assigned to chromosomes, and each chromosome is covered on average by 25 large scaffolds. The total N50 size is 1.53 Mbp with 87 sequences covering 50% of the assembly. After consensus correction using Illumina read data we performed evidence-based gene prediction using 583 million sugar beet mRNAseq reads. We predicted 32,064 genes supported by evidence of which 51% were completely covered. Predictions without mRNAseq support were generally less accurate and less reliable. The majority of genes supported by evidence is located within genetically anchored scaffolds and will be a valuable resource for sugar beet genomics and comparative studies.

Planned developments

The sugar beet genome sequencing project is carried out in cooperation between CRG and MPIMG in the context of the BeetSeq and AnnoBeet projects, both funded by the Ministry of Education and Research (BMBF).

Selected publications

Dohm JC, Lange C, Holtgräwe D, Rosleff Sörensen T, Borchardt D, Schulz B, Lehrach H, Weisshaar B, Himmelbauer H (2012). *Palaeohexaploid ancestry for Caryophyllales inferred from extensive gene-based physical and genetic mapping of the sugar beet genome.* Plant J 70:528-540

Minoche AE, Dohm JC, Himmelbauer H (2011). *Evaluation of genomic high-throughput sequencing data generated on Illumina HiSeq and Genome Analyzer systems*. Genome Biology 12:R112

Vivancos AP, Güell M, Dohm JC, Serrano L, Himmelbauer H (2010). *Strand-specific sequencing of the transcriptome*. Genome Research 20:989-999

Former Groups of the Department / Associated Research Groups

in vitro Ligand Screening Group

(05/2002-09/2012)

Head

Dr. Zoltán Konthur* (05/02-09/12)
Phone: +49 (0)30 8413-1586
Fax: +49 (0)30 8413-1365
Email: konthur@molgen.mpg.de

In October 2012, Zoltán Konthur moved to the Dahlem Centre for Genome Research and Medical Systems Biology (DCGMS) in Berlin.

Scientists

Volker Sievert* (10/05-12/11)

Technician

Carola Stoschek (since 05/02)

PhD students

Yuliya Georgieva* (since 06/09)
Stephan Klatt* (since 10/08)
Florian Rubelt* (since 10/08)
Theam Soon Lim* (09/06-08/09)
Chenna Reddy Galiveti* (11/04-08/09)

Students

Grischa Fuge* (02/11-05/12)
Sunniva Förster (01/11-12/11)
Friederike Braig (03/11-09/11)
Miriam Baradari (09/10-04/11)
Dejan Gagoski* (07/09-05/10)
Sandra Schmöckel (09/09-02/10)
Tonio Schütze (05/09-04/10)
Svetlana Mollova* (03/09-11/09)
Susanne Weber (02/09-05/09)

* externally funded

Scientific overview

Research concept

Normally, the immune system protects us from foreign substances or pathogens by generating specific antibodies eliciting an immune response. However, in a variety of diseases – especially autoimmune disorders – dysfunction of the immune system occurs leading to self-reactive (auto)antibodies. In some cases, these antibodies can cause severe damage to the body, while in other cases their presence is not understood. Our knowledge about their role in disease progression, whether being of significance or simply a bystander effect, is vague. The scientific focus of the group has shifted since the last evaluation and now centres on (1) the analysis of V(D)J recombination patterns in immunoglobulin repertoires in healthy individuals and autoimmune patients and (2) the elucidation of autoantigenicity pattern in health and disease. The methodological portfolio includes the use of Next Generation Sequencing (NGS), Phage Display as well as Protein Array Technologies.

Scientific methods and findings

Analysis of V(D)J recombination patterns in immunoglobulin repertoires of healthy individuals and autoimmune patients.

Our goal is to explore whether there is a difference in the nature of how healthy individuals and autoimmune patients shape their antibody repertoires. The variety of immunoglobulin (Ig) paratopes for antigen recognition is a result of a V(D)J recombination mechanism in the heavy and the light chain of the antibody molecule, while a fast and efficient immune response is mediated by specific Ig-isotypes obtained through class switch recombination (CSR). Hence, we believe that it is not enough to analyse V(D)J recombination. The effector function of an antibody encoded in the isotype is of equal importance and need to be addressed as well.

As no adequate analytical tools were available to tackle this question, we have established a new method of yet unpaired sensitivity to amplify and sequence the expressed antibody repertoire of an individual. The method is based on V-gene independent amplification of rearranged immunoglobulin repertoires in combination with emulsion PCR to minimize primer- and PCR-induced bias. We first analysed the obtained sequences using the IMGT/High V-Qest online tool. Further, in cooperation with the group of Prof. Edda Klipp (Humboldt University of Berlin), we developed a novel avenue of bioinformatic analysis based not only on information on V(D)J recombination, but also on class-switch recombination of individual donors by incorporating isotype-specific analysis of the antibody sequences.

First, we established a large data set from 14 healthy Caucasian donors of different age and gender. Hierarchical clustering of the donors only according to the V(D)J recombination information revealed neither correlation by age nor gender. However, when CSR information was introduced into the analysis, for the first time, donors clustered hierarchically according to age. We could observe changes in Ig-isotype repertoires to be age-dependent indicating reduction of class-switch capability. This is in good agreement with recent findings suggesting that the dramatically reduced vaccination efficacy in elderly populations is not because of a lack of specific antibodies due to reduction of V(D)J recombination, but rather a problem in antibody titre and lacking specificity in the right immunoglobulin

class to elicit response. Unexpected however is the fact, that the decline of class-switch ability starts already relatively early. The age of fifty and beyond already defines the onset of immune senescence.
In a second step, we now analyse a large data set obtained from autoimmune patients with rheumatoid arthritis (RA). For this, we have obtained patient samples from our collaboration partners in the Department of Rheumatology and Clinical Immunology of the Charité. The two data sets will be compared to see whether there is a difference in the nature of V(D)J recombination patterns between the antibody repertoires of healthy individuals and autoimmune patients. The bioinformatic analysis is conducted in collaboration with the group of Peter Arndt (Department Vingron).

Elucidation of autoantigenicity patters in health and disease

Everybody has circulating self-reactive antibodies in their blood. Although individual repertoires of autoantibodies can significantly overlap, they differ between healthy and diseased individuals. Differential analysis can lead to the identification of biomarker sets that can clearly separate different autoimmune diseases or even allow subdiagnosis of patients within a certain disease.
We apply two complementary screening methods for the discovery of autoantigenicity patters, namely Protein Array and Phage Display Technology. The methods comprise different subsets of the human proteome and offer different means of selection. While most antigens on the array are denatured, the proteins on the bacteriophage surface are presented as folded structures. Our protein arrays consist of ~25.000 expression constructs of a human foetal brain cDNA library representing 2055 human genes in multiple copies. For phage display, we have now generated various versions of full-ORF libraries of 4452 genes applying the Gateway-technology. While the identity of each spot on the protein array is known, the phage display libraries require downstream processing. Selection is carried out in an iterative process based on affinity enrichment using patient-derived immunoglobulin fractions as selection targets. The identity of the enriched clones is revealed by NGS of the DNA inserts.
In a current project (BMWi-ZIM) we have applied these screening technologies for the elucidation of autoantigenicity profiles of healthy donors, Multiple Sclerosis and Alzheimer's patients. The healthy donors are age and sex matched with the Alzheimer's cohort and have all been diagnosed CERAD negative. The bioinformatic analyses are conducted in collaboration with the group of Ralf Herwig, applying ConsensusPathDB and Ingenuity Pathway Analysis.
Additional ongoing projects centre on RA and are conducted in close collaboration with Karl Skriner (Charité). We were able to find biomarkers, which allow discrimination between early stages of RA and systemic lupus erythematosus even in rheuma factor negative patients (patent filed). The downstream characterisation and evaluation of the biomarkers is subject of a recently started project (BMBF – KMU-Innovativ 3).
Another ongoing project is based on a set of protein biomarkers we have found to predict therapy response in RA patients to be treated with tumour necrosis factor alpha (TNFα) blocking biologicals (patent filed), such as monoclonal antibodies (infliximab, adalimumab) or soluble TNFα-receptor (etanercept). Currently, these biologicals are only second line treatment medications in RA patients refractory to standard disease-modifying antirheumatic drugs (DMARDS, e.g. Methotrexate), even though early treatment with biologicals is anticipated to achieve full remission if only treated within the first 3-6 month after disease

onset. The major drawback of these biologicals is the high rate of therapy non-responders (30-40%) as well as their high cost. Our biomarker panel can be used to quickly identify the patients, which could benefit from such medication prior treatment. We are shortly finalising an evaluation of our predictive markers in an investigator initiated clinical study with 150 RA patients, conducted at the Charité and sponsored by Pfizer.

Selected publications

Georgieva Y, Konthur Z (2011). *Design and Screening of M13 phage display cDNA libraries.* Molecules 16:1167-1183

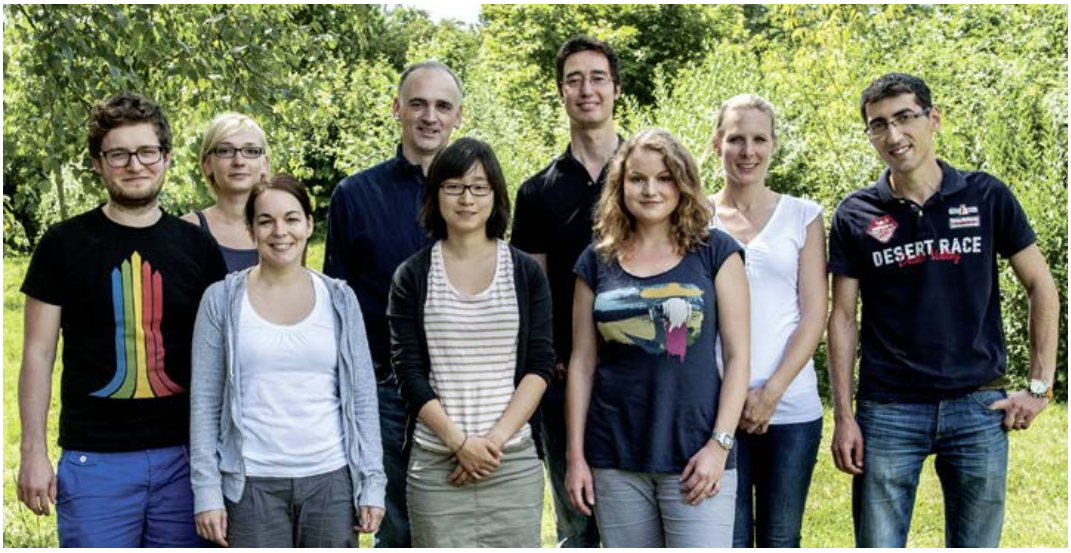
Bourbeillon J, Orchard S, Benhar I, Borrebaeck C, de Daruvar A, Dübel S, Frank R, Gibson F, Gloriam D, Haslam N, Hiltke T, Humphrey-Smith I, Hust M, Juncker D, Koegl M, Konthur Z, Korn B, Krobitsch S, Myldermans S, Nygren PA, Palcy S, Polic B, Rodriguez H, Sawyer A, Schlapschy M, Snyder M, Stoevesandt O, Taussig MJ, Templin M, Uhlen M, van der Maarel S, Wingren C, Hermjakob H, Sherman D. (2010) *Minimum information about a protein affinity reagent (MIAPAR).* Nat Biotechnol 28:650-653

Lim TS, Mollova S, Rubelt F, Sievert V, Dübel S, Lehrach H, Konthur Z (2010). *V–gene amplification revisited – An optimised procedure for human antibody isotype and idiotype amplification*. N Biotechnol 27:108-117

Former Groups of the Department / Associated Research Groups

Protein Complexes and Cell Organelle Assembly Group

(05/2003-11/2012)



Head

Priv. Doz. Dr. Bodo M.H. Lange
(05/03-12/11)
Phone: +49 (0)30 8413-1645
Fax: +49 (0)30 8413-380
Email: lange_b@molgen.mpg.de
b.lange@alacris.de

Since January 2012, Bodo Lange has been appointed as managing director of Alacris Theranostics GmbH in Berlin.

Personal assistant

Antje Lazarovic* (08/08-04/11)

Scientists

Artur Muradyan* (since 12/09)
Hannah Müller* (since 06/08)
Irina Czogiel* (01/10-12/11)
Karin Habermann* (11/10-03/11)
Thomas Kessler* (07/07-09/10)
Marie-Laure Fogeron* (04/05-09/09)

Technicians and engineers

Verena Lehmann* (since 07/03)
Melanie Hessler* (08/09-12/11)
Tanja Kurtz* (07/06-12/11)
Karl-Heinz Rak (05/03-08/11)
Anne Kühnel* (11/07-05/11)

PhD students

Sophia Schade* (since 10/10)
Seon-Hi Julia Jang* (since 09/08)
Nicole Hallung* (since 06/08)
Armin Haupt* (since 10/07)
Anne-Kathrin Scholz* (05/06-12/10)
Karin Habermann* (12/05-10/10)

* externally funded

Scientific overview

Background

New opportunities are presently available, for the direct investigation of the molecular basis of diseases caused by mutation of genes. The technology in the area of functional genomics has matured and can now be applied to basic questions in the fields of developmental biology, cell biology and molecular medicine. This, in turn, facilitates the study of many parameters in parallel (mRNA and protein profiles, morphological cellular or subcellular analysis) and their dynamic changes upon modulating cues (*e.g.* cellular or environmental stress). Such complex data sets require now system models to understand and predict diseases development. In this context, the study of the protein complexes and protein-protein interactions of the components of cellular proliferation and signalling pathways is now an opportunity and highly relevant for our understanding of the molecular basis of development and diseases.

Current state of research and scientific findings

The main research focus of my group is the regulation of cell division and cell proliferation pathways in human cells and *Drosophila*. Critical for cellular division and tissue homeostasis is the cytoskeleton and the centrosome (the microtubule organizing centre in the eukaryotic cell).

We have set up an experimental system that combines cell biological, functional genomics and proteomics approaches. Using these approaches, we identified the proteome of the *Drosophila* mitotic centrosome and we functionally characterized all the components *via* RNAi. From the 260 identified proteins, we could assign new functions to 27 proteins. We could describe a core set of 11 proteins that is required for maintaining centrosome integrity. We focused subsequently our work on these 11 proteins as centrosome stability is frequently compromised in human cancer cells. Characterization of the human orthologues of these 11 proteins identified 4 functionally highly conserved components. Initially some results of this study focussing on the spindle assembly checkpoint were published, followed with a detailed functional characterisation of almost all identified components and its human orthologues.

From the identified centrosomal proteins, those with novel centrosome-associated functions are analysed *via* two basic approaches: First they are characterized on the developmental, cellular and biochemical level in the fly. To this end, we employ tissue-specific overexpression of target proteins and conditional depletion of proteins through RNAi, transgenic or mutant fly stocks. Second, we identify the protein complexes, *via* which these molecules exert their function, using protein complex isolation from *Drosophila* embryo extracts and tandem affinity purification (TAP) from human cells. Both approaches have defined overlapping and complementary lists of proteins that we exploited for our subsequent work. Our TAP experiments have identified an interaction network of 2700 proteins. From this set, we selected 20 proteins that were annotated in databases to be deregulated in cancer. These proteins have been characterized by RNAi and overexpression approaches in human tissue culture revealing striking effects for centrosome function, reflecting the deregulation of centrosome function *in situ* in human cancer tissues as confirmed on morphological, mRNA and protein level.

In addition, my group is participating in three medically related and complementary network projects. As part of *MUTANOM,* (www.mutanom.org) which I coordinate, we study the consequence of cancer-related mutations on the cellular

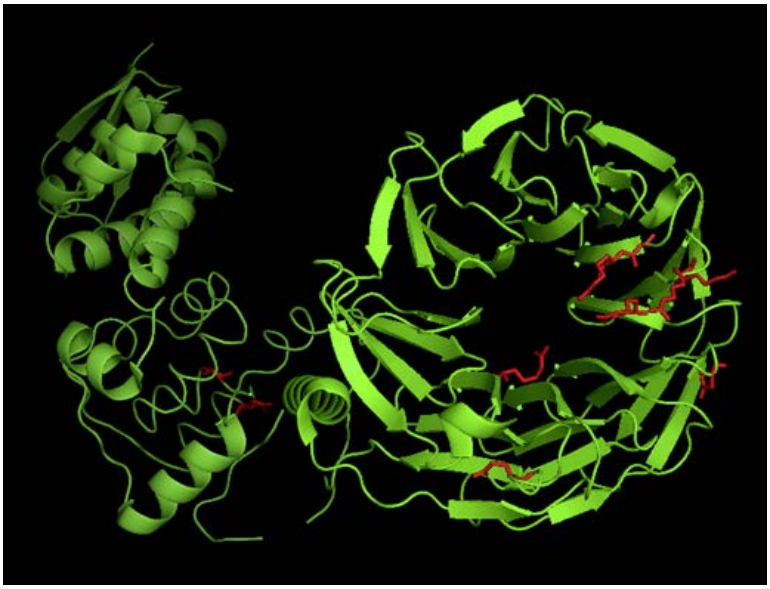


Figure 26: Frequently the effect of the mutations in cancer genes leads to a change of protein-protein interactions. The image here shows an example of a molecular model of the tumour suppressor FBXW7. Positions of frequent mutations are marked in red. The consequence of mutations in oncogens and tumours suppressors is investigated on a biochemical and functional level as part of the MUTANOM project.

and organism level (Figure 26), aiming to build predictive models for disease development and for the identification of new molecular targets for site-specific therapies. Second, as part of *NeuroNet*, we are investigating the protein-interaction networks implicated in the development of neurodegenerative diseases employing our TAP approach. Third, as part of the network *Systems Biology of the Cell,* we are studying estrogen receptor (ER) signalling upon modulating the ER pathway activity. In all three projects we contribute with the methods established in the group (human isogenic cell line system, expertise in protein complex isolation, phenotypical assays) to analyse the consequence of disease-related genes on a biochemical and cellular functional level. After the conclusion of some of these projects, we are now participating in the EU project OncoTrack (www.oncotrack.eu) characterizing the consequence of cancer mutations in colon cancer.

Selected publications

Müller H, Schmidt D, Steinbrink S, Mirgorodskaya E, Lehmann V, Habermann K, Dreher F, Gustavsson N, Kessler T, Lehrach H, Herwig R, Gobom J, Ploubidou A, Boutros M, Lange BMH (2010). *Proteomic and functional analysis of the mitotic Drosophila centrosome.* EMBO J 29:3344-3357

Stehr H, Jang SHJ, Duarte JM, Wierling C, Lehrach H, Lappe, M, Lange BMH (2011). *The structural impact of cancer-associated missense mutations in oncogenes and tumor suppressors*. Mol Cancer 10:54

Haupt A, Joberty G, Bantscheff M, Fröhlich H, Stehr H, Schweiger MR, Fischer A, Kerick M, Boerno ST, Dahl A, Lappe M, Lehrach H, Gonzalez C, Drewes G, Lange BMH (2012). *Hsp90 inhibition differentially destabilises MAP kinase and TGF-beta signalling components in cancer cells revealed by kinase-targeted chemoproteomics*. BMC Cancer 12(1):38

Former Groups of the Department / Associated Research Groups

Molecular Biology of Metabolism Group

(12/2007-12/2011)

Head
Markus Ralser (12/07-12/11)
now at University of Cambridge, UK
Cambridge Systems Biology Centre
and Dept. of Biochemistry
Phone: +44 1223-761346
Email: mr559@cam.ac.uk

The group moved to the University of Cambridge in 2011/2012.

Scientists
Nana-Maria Grüning (since 02/12)
Katharina Blümlein (01/10–12/11)
Jakob Vowinckel (01/10-10/11)

PhD students
Antje Krüger (since 04/09)
Michael Mülleder (since 06/11)
Steve Michel (since 03/10)
Nana-Maria Grüning (02/09-01/12)

Guest
Floriana Capuano, 24.10.-01.11.2011
Kate Campbell, 24.10.-01.11.2011

Scientific overview

Research concept

We are interested in the regulatory function of the metabolic network and investigate, how metabolic intermediates are implicated in the control of biological systems. An important situation, where the metabolome has regulatory function, is the oxidative stress response. Exposure to a toxic oxidant dose leads to a (temporal) reconfiguration of cellular metabolism. As a result, the concentration of several metabolites is altered; and these - in turn - are implicated in adapting the cell to oxidative conditions. The main biological questions we focus on are:

- How are metabolic transitions regulated, e.g. the change from oxidative to non-oxidative metabolism?
- How flexible are metabolic pathways? How do they interact with the regulation of cellular macromolecules?
- Why do cells age? Which metabolic processes are involved? Can we modulate aging by changing these processes?

Scientific methods and findings

Answering of these research topics is achieved by the combination of yeast and mammalian functional genomics with advanced bioanalytics. We use quantitative mass spectrometry coupled to high-pressure and/or nano-flow liquid chromatography for targeted quantitative and qualitative analysis of small molecules and peptides, and explore functional yeast genetics to detect the genetic interactions. In doing so, we could identify crucial metabolic hubs that help the cell to adjust its metabolism according to the current needs. For instance, the glycolytic enzyme GAPDH gets inactivated by high oxidant doses and this blockage causes a redirection of the carbohydrate flux into the adjacent Pentose Phosphate Pathway (PPP). Immediate activation of the PPP is essential to cellular survival in an oxidative stress situation. This deviation stabilizes the pool of redox co-factors required for the antioxidative machinery. Importantly, the PPP also plays a crucial role in accurate timing of the transcriptional response to oxidative stress, which is necessary for the cell to return to normal metabolism after the prompt metabolic reaction has ceased.

To study the effects of the glycolysis / PPP transition in mammals, we generated a transgenic mouse with a reduced activity of the glycolytic enzyme triose phosphate isomerase (TPI) due to a point mutation in the catalytic core (Ile170Val). The mouse line also gives insights into the pathogenic causes of TPI deficiency, a fatal genetic disease characterized by haemolytic anaemia and neurological dysfunction, immune deficiency or motorneuronal symptoms, usually leading to death in early childhood. However, recent molecular analysis indicated that rather than possessing catalytically deficient alleles, patients as well as fly and mouse models express TPI forms deficient in stability and dimerization. Therefore, our work focuses on the pathogenic causes of the TPI deficiency and on the question how the glycolysis / PPP transition influences mammalian cellular functions.

Furthermore, another glycolytic enzyme – pyruvate kinase (PYK) – was reported to be involved in metabolic rearrangements in proliferating cells. The so-called Warburg effect describes a transition from oxidative phosphorylation to anaerobic glycolytic metabolism, and its reversal inhibits cancer cell proliferation. Oxidative metabolism also produces reactive oxygen species (ROS) which, in excess, can lead to oxidative stress. Studying the metabolic conse-

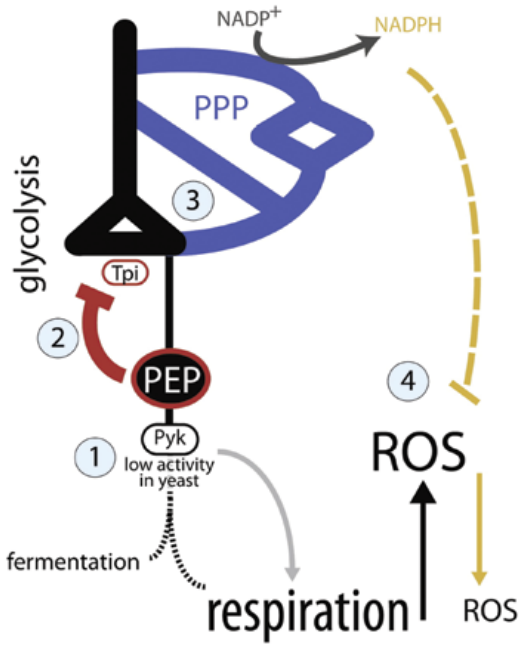


Figure 27: The PYK-PEP-TPI feedback loop synchronizes respiration and redox metabolism. Low pyruvate kinase (PYK) activity activates respiration in yeast (1), Accumulation of the PYK substrate PEP inhibits the glycolytic enzyme TPI (2), TPI inhibition stimulates the pentose phosphate pathway (3), preventing ROS accumulation (4) (Grüning et al., 2011).

quences of low PYK activity, we discovered that yeast central metabolism is self-adapting to synchronize redox metabolism when respiration is activated. Low PYK activity activated respiration. However, levels of reactive oxygen species (ROS) did not increase, as cells gained resistance to different oxidizing substances. Revealed by HILIC-MS/MS and enzymatic analyses, this was attributable to accumulation of the PYK substrate phosphoenolpyruvate (PEP), which acted as feedback inhibitor of the glycolytic enzyme triosephosphate isomerase (TPI). LC-MS/MS measurements revealed that TPI inhibition stimulated an activation of the pentose phosphate pathway (PPP), which is crucial part of the cellular anti-oxidative machinery. PPP activation proved to be essential for cellular protection against macromolecular damage caused by increased ROS levels arising from elevated respiration. Thus, a metabolic feedback loop, initiated by PYK, mediated by its substrate PEP and acting on TPI, stimulates redox metabolism in respiring cells. Originating from a single catalytic step, this autonomous metabolic reconfiguration prevents oxidative stress upon increased cellular respiration (Figure 27), and demonstrates how the cell uses its own metabolism to balance fundamental processes.

In human cells, it has been reported that the Warburg effect originates from a switch in the expression of alternative splice forms (PKM1 and PKM2) of PYK, which is also important for malignant transformation. However, by using mass spectrometry, we performed an absolute quantification of PKM1 and PKM2 splice isoforms in different human malignant cancers, benign oncocytomas, tissue matched controls, and several cell lines. PKM2 was the prominent isoform in all analyzed cancer samples and cell lines. However, this PKM2 dominance was not a result of a change in isoform expression, since PKM2 was also the predominant PKM isoform in matched control tissues. Thus, we could clarify that an exchange in PKM1 to PKM2 isoform expression during cancer formation is not occurring, nor do these results support conclusions that PKM2 is specific for proliferating and PKM1 for non-proliferating tissue. Ongoing work focuses on metabolic signals that prevent cells from excess proliferation under oxidative stress and also on metabolic factors involved in the cellular aging process.

Selected publications

Bluemlein K, Ralser M (2011). *Monitoring Protein Expression in whole-cell extracts by targeted label- and standard-free LC-MS/MS*. Nature Protocols 6(6):859-69

Grüning NM, Rinnerthaler M, Bluemlein K, Mülleder M, Wamelink MMC, Lehrach H, Jakobs C, Breitenbach M, Ralser M (2011). *Pyruvate Kinase Triggers a Metabolic Feedback Loop that Controls Redox Metabolism in Respiring Cells*. Cell Metabolism 14(3):415-427

Krüger A, Grüning NM, Wamelink MM, Kerick M, Kirpy A, Parkhomchuk D, Bluemlein K, Schweiger MR, Soldatov A, Lehrach H, Jakobs C, Ralser M (2011). *The pentose phosphate pathway is a metabolic redox sensor and regulates transcription during the antioxidant response.* Antioxid Redox Signal. 2011 Jul 15;15(2):311-24.

Former Groups of the Department / Associated Research Groups

Technology Development Group

(06/2007-07/2012)

In August 2012, Alexey Soldatov has moved to the Dahlem Centre for Genome Research and Medical Systems Biology (DCGMS) in Berlin.

Scientists

Tatiana Borodina (06/07-12/11)
Dmitri Parkhomchuk (06/07-09/11)

PhD students

Vera Rykalina (since 11/10)
Alexey Shadrin (since 10/10)
Maria Metsger (since 09/10)
Alexey Davydov (since 07/08)
Vyacheslav Amstislavsliy (11/07-09/11)

Engineer

Berit Lenz (11/09-05/11)
Jeannine Wilde (06/08-03/10)

Head

Alexey V. Soldatov (06/07-07/12)
Phone: +49 (0)30 8413-1128
Fax: +49 (0)30 8413-1380
Email: soldatov@molgen.mpg.de

Scientific overview

Till June 2007, A. Soldatov's team was within Lehrach's group and has specialised in the area of SNP genotyping and sequencing. A new SNP-genotyping method was developed, patented (PCT/EP2004/009546) and successfully applied for genotyping of *Arabidopsis thaliana* and human DNA samples. We have also worked on the development of new sequencing technologies in the frame of the European MolTools project: (i) microbead based sequencing (PCT/EP2006/008535) and (ii) ligation based sequencing in collaboration with U. Landegren laboratory (Uppsala University, Uppsala, Sweden).
Since June 2007, our group specializes in technology development and bioinformatics related with second generation sequencing (NGS). Till spring 2010 we were responsible for the NGS service in the department. We have set up the optimized laboratory workflows for both Illumina and ABI SOLiD sequencing platforms. We have developed a NGS bioinformatics pipeline, which brings together the information about the whole sequencing process from sample handling to results analysis. The system enables each authorized user (also external col-

laborators) to track processing of samples, check the sequencing quality and view data in the genomic browser.
The group uses NGS platforms for analysis of genome (resequencing) and transcriptome (gene expression profiling, splice junctions search, allele-specific expression, reverse transcription, etc.). We have performed RNA-Seq, ChIP-Seq, MedIP-Seq and genomic DNA sequencing (including sequencing library preparation, sequencing and sometimes data analysis) for several groups within the MPIMG (AG Yaspo, AG Nietfeld, AG Himmelbauer, AG Schweiger, AG Hoehe, AG Lange, AG Krobitsch, AG Ralser, dept. Herrmann) and for external collaborators (Y. Shiloh (U. Tel Aviv), C. Koncz (MPI for Plant Breeding Research, Cologne), J. Klose (Charite, Berlin)).
We have developed, patented and published a convenient method for strand-specific sequencing of cDNA. The method is based on incorporation of uridine bases during first (or second) cDNA strand synthesis and subsequent destruction of this strand before sequencing. Knowledge of transcript orientation is important for transcriptome studies. It allows (i) to annotate novel genes correctly, (ii) to investigate antisense transcription, which plays an important regulatory role in all eukaryotes, (iii) to resolve transcript overlaps, which are abundant in compact genomes of prokaryotes and lower eukaryotes, and (iv) to correctly determine gene expression levels in the presence of overlapping antisense transcription. The method is licensed by New England Biolabs company.
We have developed and patented a method for preservation of information about spatial distribution of nucleic acid molecules transferred from a solid surface to another solid surface or into solution. This method permits to use a power of NGS for studying of two-dimensional distributions of nucleic acid molecules on tissue sections or other biological objects. Biological processes are spatially organized and rely upon the interplay of many different components forming an intricate structure of cells, tissues and organisms. Molecules participating in these processes have a certain spatial distribution. Understanding the biological processes is critically dependent on a detailed knowledge of this distribution.
At present, the group's activities are focused on NGS-related technology development:

- early coding for parallel preparation of a number of sequencing libraries;
- RNA structural analysis;
- methylation analysis;
- preparation of NGS libraries from ultra small amounts of starting material;
- long-run NGS sequencing for haplotyping.

We are also involved in several biological projects requiring NGS for analysis of genome and transcriptome (mainly in collaboration with AG Yaspo, MPIMG):

- resequencing of the genome of the PWD mouse strain;
- 1000 Genomes project;
- GEUVADIS RNA-Seq;
- Treat 1000 project;
- Treat 20 project;
- ICGC PedBrain;
- ICGC prostate cancer;
- IMI Oncotrack;
- IMGuS staetohepatitis project;
- mouse transcriptional network;
- EU FP7 ADAMS project.

Selected publications

Borodina T, Adjaye J, Sultan M (2011). *A strand-specific library preparation protocol for RNA sequencing*. Methods Enzymol 500:79-98

Parkhomchuk D, Amstislavskiy V, Soldatov A, Ogryzko V (2009). *Use of high throughput sequencing to observe genome dynamics at a single cell level.* Proc Natl Acad Sci U S A 106(49):20830-5.

Parkhomchuk D, Borodina T, Amstislavskiy V, Banaru M, Hallen L, Krobitsch S, Lehrach H, Soldatov A (2009). *Transcriptome analysis by strand-specific sequencing of complementary DNA*. Nucleic Acids Res 37(18):e123

Former Groups of the Department / Associated Research Groups

Cardiovascular Genetics Group

(2001-02/2011)

Head

Univ.-Prof. Dr. med. Silke Rickert-Sperling,
now at Experimental Clinical Research Center (ECRC), Berlin
Phone: +49 (0)30 8413-1232
Fax: +49 (0)30 8413-1699
Email: sperling@molgen.mpg.de

In 03/2011, Silke Sperling accepted an appointment as W3 Heisenberg professor at the Charité - Universitätsmedizin Berlin. She established a new Department of Cardiovascular Genetics at the Experimental Clinical Research Center (ECRC), a joint institute between Charité - Universitätsmedizin Berlin and Max Delbrück Center for Molecular Medicine. In 09/2011, she was also co-appointed at the Faculty of Biology, Chemistry and Pharmacy at Freie Universität Berlin.

Secretary

Martina Luig (since 03/12#)
Barbara Gibas (10/06–08/10 and 04/11-03/12)

Technicians

Kerstin Schulz (since 01/12#)
Susanne Thomsen (06/11-12/11)
Ilona Dunkel (01/02-08/11)

Scientists and postdocs

Marcel Grunert (since 03/12#)
Jenny Schlesinger (since 01/12#)
Markus Schüler (04/11-12/11)
Martje Tönjes (04/09-09/09)
Andreja Brodarac (03/08-03/09)

PhD students

Vikas Bansal (since 11/11#)
Katherina Bellmann (since 08/11#)
Huanhuan Cui (since 05/11#)
Cornelia Dorn (since 02/10)
Marcel Grunert (07/08-02/12)
Jenny Schlesinger (04/08-12/11)
Markus Schüler (01/07-03/11)
Zhang Qin (10/08-09/10)
Lucas Rudigier (10/09-03/10)
Julia Kofent (04/09-09/09)
Martje Tönjes (02/06-03/09)
Katharina Rost (07/07-01/08)

Students

Cornelia Dorn (01/09-06/09)
Michalina Mańkowska (05/09-10/09)

appointed after transfer to ECRC

Scientific overview

Most cardiovascular diseases have complex genetic and environmental origins. Our lab studies molecular mechanisms underlying cardiac development and function using molecular biological techniques and bioinformatics expertise. We focus on the transcriptional regulation process, which plays a key role for normal and abnormal cardiogenesis leading in the latter case to congenital heart disease (CHD). A rapidly growing number of factors have been shown to be involved in regulating the pattern and timing of expression of genes responsible for the cardiac lineage determination, heart chamber formation, valvulogenesis and conduction-system development. Spatiotemporal and quantitative regulation of cardiac transcription factors must occur in a precise manner to ensure fine regulation of downstream targets. However, the ability of transcription factor binding to DNA is highly influenced by the chromatin status, and epigenetic mechanisms have an important role in establishing and maintaining transcriptional programs. To understand networks directing gene expression, the interplay between different transcription factors, co-regulatory elements and epigenetic factors has to be considered.

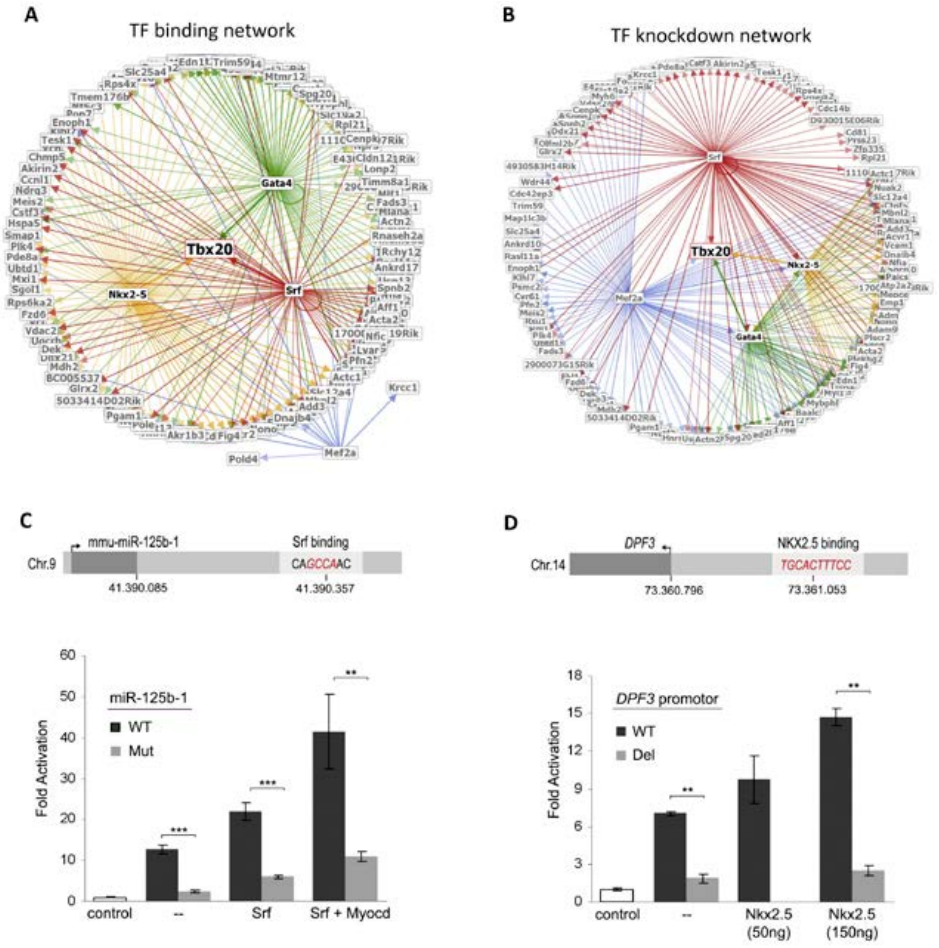


Figure 28: Regulation of the cardiac transcriptome by combinatorial binding of transcription factors (TF) Gata4 (green), Mef2a (blue), Nkx2.5 (yellow) and Srf (red). Shown are in total 1671 genes, which are targets of at least one factor. (A) Direct downstream targets identified by chromatin-immunoprecipitation. (B) Direct and indirect downstream targets identified by siRNA knockdown of respective factors. Promoter analyses of miR-125b-1 (C) and DPF3 (D) using luciferase reporter gene assays show regulation by Srf and Nkx2.5 respectively.

Regulation of cardiac gene expression

We performed a systems biology study in murine cardiomyocytes integrating mRNA profiles with DNA–binding events of key cardiac transcription factors (Gata4, Mef2a, Nkx2.5, and Srf), activating histone modifications (H3ac, H4ac, H3K4me2, and H3K4me3), and microRNA profiles obtained in wild-type and RNAi–mediated knockdown. Even these non-paralogous transcription factors partially compensate each other's function. Srf and Gata4 driven gene expression is highly dependent on the co-presence of H3ac, whereas Nkx2.5 and Mef2a activate transcription in a H3ac independent manner. A significant proportion of indirect Srf downstream targets are potentially regulated through Srf-regulated microRNAs. These findings show the impact of interdependencies between different factors and layers of regulation on the cardiac transcriptome. Further functional tests are required to evaluate regulatory circuits on a single gene basis. It is of interest to study how the interplay of the different factors stabilizes the overall function of a given network, and how this contributes to the resistance to external disturbances as well as pathogenic mutations missing phenotypic impact.

Chromatin-remodelling factor DPF3 (BAF45c)

We have identified the first chromatin-remodelling factor, namely DPF3 (BAF45c), which is capable of binding acetylated as well as methylated histone modifications through its double plant-homeodomain (PHD-finger). Thus it bridges distinct regulatory signals of the histone code and enables a tissues-specific read-out based on its neural, cardiac and muscle specific expression. DPF3 is associated with the BAF chromatin-remodelling complex and came to our attention due to its differential expression in hypertrophic hearts of Tetralogy of Fallot. DPF3 is evolutionarily conserved and knockdown of dpf3 in zebrafish leads to incomplete cardiac looping and severely reduced ventricular contractility, with disassembled muscular fibers. Changes in chromatin structure and gene transcription are frequently induced by external stimuli. In particular phosphorylation of chromatin-associated proteins is mediated by different kinases, such as p38, CaMK and CKII, and represents a powerful interface for the transmission of extracellular signals to chromatin. So far, the upstream signalling pathway of DPF3 is undiscovered and a particular focus of our research.

Congenital heart disease

Cardiac malformations represent a broad panel of in part overlapping phenotypes, reflecting the modular background of cardiogenesis. The biological network modifying the impact of key regulators is still widely a black box. One important lesson learned from the study of animal models, as well as patients with complex CHD is that our primary hope that one gene would simply refer to one phenotype is not fulfilled. Even one particular mutation can be associated with a panel of different cardiac malformations and the majority of CHD is not following Mendelian inheritance. Nevertheless, it is undoubtful that there is a clear genetic impact. We currently enter a novel era of research, which is characterized by fast evolving technological advances enabling the generation, study and interpretation of more and more complex data. This will allow the study of congenital heart disease in a manner, which provides insights into the underlying biological network influenced by genetic, epigenetic and environmental factors, and stochastic events.

Selected publications

Schueler M, Zhang Q, Schlesinger J, Tönjes M, Sperling SR (2012). *Dynamics of Srf, p300 and histone modifications during cardiac maturation in mouse*. Mol Biosyst 8: 495–503

Sperling SR (2011). *Systems biology approaches to heart development and congenital heart disease*. Cardiovasc Res 91: 269–278

Schlesinger J, Schueler M, Grunert M, Fischer JJ, Zhang Q, Krueger T, Lange M, Tönjes M, Dunkel I, Sperling SR (2011). *The cardiac transcription network modulated by Gata4, Mef2a, Nkx2.5, Srf, histone modifications, and microRNAs*. PLoS Genet 7:e1001313

General information about the whole Department

Complete list of publications (2009-2012)

2012

Adams D, Altucci L, Antonarakis SE, Ballesteros J, Beck S, Bird A, Bock C, Boehm B, Campo E, Caricasole A, Dahl F, Dermitzakis ET, Enver T, Esteller M, Estivill X, Ferguson-Smith A, Fitzgibbon J, Flicek P, Giehl C, Graf T, Grosveld F, Guigo R, Gut I, Helin K, Jarvius J, Küppers R, Lehrach H, Lengauer T, Lernmark Å, Leslie D, Loeffler M, Macintyre E, Mai A, Martens JH, Minucci S, Ouwehand WH, Pelicci PG, Pendeville H, Porse B, Rakyan V, Reik W, Schrappe M, Schübeler D, Seifert M, Siebert R, Simmons D, Soranzo N, Spicuglia S, Stratton M, Stunnenberg HG, Tanay A, Torrents D, Valencia A, Vellenga E, Vingron M, Walter J, Willcocks S (2012). *BLUEPRINT to decode the epigenetic signature written in blood.* Nat Biotechnol 30(3):224-6

Barberis M, Linke C, Adrover MA, Lehrach H, Krobitsch S, Posas F, Klipp E (2012). *Sic1 potentially regulates timing and oscillatory behaviour of Clb cyclins*. Biotechnol Adv 30:108-130

Berndt H, Harnisch C, Rammelt C, Stöhr N, Zirkel A, Dohm JC, Himmelbauer H, Tavanez JP, Hüttelmaier S, Wahle E (2012). *Maturation of mammalian H/ACA box snoRNAs: PAPD5-dependent adenylation and PARN-dependent trimming*. RNA 18:958-972

Breitenbach M, Laun P, Dickinson JR, Klocker A, Rinnerthaler M, Dawes IW, Aung-Htut MT, Breitenbach-Koller L, Caballero A, Nystrom T, Buttner S, Eisenberg T, Madeo F, Ralser M (2012). *The role of mitochondria in the aging processes of yeast.* Subcellular biochemistry 57:55-78

Cesuroglu T, van Ommen B, Malats N, Sudbrak R, Lehrach H and Brand A. (2012). *Public health perspective: from personalized medicine to personal health.* Personalized Medicine 9(2):115-119

Chin ST, Choong YS, Konthur Z, Lim TS (2012). *Generation of recombinant antibodies by display technologies*. In: Sustainable Diagnostics for the Bottom Millions, Oxford Academic Press (A. Acosta, N.M. Nor & M.E. Sarmiento, Eds.), in press

Clark MS, Denekamp NY, Thorne MAS, Reinhardt R, Drungowski M, Albrecht MW, Klages S, Beck A, Kube M, Lubzens E (2012). *Long-term survival of hydrated resting eggs from Brachionus plicatilis*. PLoS One 7(1):e29365

Clarke L, Zheng-Bradley X, Smith R, Kulesha E, Xiao C, Toneva I, Vaughan B, Preuss D, Leinonen R, Shumway M, Sherry S, Flicek P; The 1000 Genomes Project Consortium* (2012). *The 1000 Genomes Project: data management and community access*. Nat Methods 27;9(5):459-462. *Lehrach H, Sudbrak R (Project Leader), Borodina T, Davydov A, Marquardt P, Mertes F, Nietfeld W, Soldatov A, Timmermann B, Tolzmann M, Albrecht M, Amstislavskiy V, Herwig R, Parkhomchuk D.

Dohm JC, Lange C, Holtgräwe D, Rosleff Sörensen T, Borchardt D, Schulz B, Lehrach H, Weisshaar B, Himmelbauer H (2012).

Palaeohexaploid ancestry for Caryophyllales inferred from extensive gene-based physical and genetic mapping of the sugar beet genome. Plant J 70:528-540

Dreher F, Kamburov A, Herwig R (2012). *Construction of a pig physical interactome using sequence homology and a comprehensive reference human interactome*. Evolutionary Bioinformatics 8:119-126

Dreher F, Kreitler T, Hardt C, Kamburov A, Yildirimman R, Schellander K, Lehrach H, Lange BMH, Herwig R (2012). *DIPSBC – a data integration platform for systems biology cooperations*. BMC Bioinformatics 3(1):85

Drews K, Tavernier G, Demeester J, Lehrach H, De Smedt SC, Rejman J, Adjaye J (2012). *The cytotoxic and immunogenic hurdles associated with non-viral mRNA-mediated reprogramming of human fibroblasts.* Biomaterials 33(16):4059-68

Duitama J, McEwen GK, Huebsch T, Palczewski S, Schulz S, Verstrepen K, Suk E-K, Hoehe MR (2012). *Fosmid-based whole genome haplotyping of a HapMap trio child: Evaluation of Single Individual Haplotyping techniques*. Nucleic Acids Res 40(5):2041-53

Eberl M, Klingler S, Mangelberger D, Loipetzberger A, Damhofer H, Zoidl K, Schnidar H, Hache H, Bauer HC, Solca F, Hauser-Kronberger C, Ermilov AN, Verhaegen ME, Bichakjian CK, Dlugosz AA, Nietfeld W, Sibilia M, Lehrach H, Wierling C, Aberger F (2012). *Hedgehog-EGFR cooperation response genes determine the oncogenic phenotype of basal cell carcinoma and tumour-initiating pancreatic cancer cells*. EMBO Mol Med 4(3):218-33.

Haupt A, Joberty G, Bantscheff M, Fröhlich H, Stehr H, Schweiger MR, Fischer A, Kerick M, Boerno ST, Dahl A, Lappe M, Lehrach H, Gonzalez C, Drewes G, Lange BMH (2012). *Hsp90 inhibition differentially destabilises MAP kinase and TGF-beta signalling components in cancer cells revealed by kinase-targeted chemoproteomics.* BMC Cancer 12(1):38

Hegele A, Kamburov A, Grossmann A, Sourlis C, Wowro S, Weimann M, Pena V, Will C, Lührmann R, Stelzl U (2012). *Dynamic protein-protein interaction wiring of the human spliceosome.* Molecular Cell 45(4):567-580

Hooli B, Mohapatra G, Mattheisen M et al (2012). *The role of common and rare APP DNA sequence variants in Alzheimer's disease.* Neurology 78(16):1250-7

Jones DTW,*, Jäger N, Kool1 M, Zichner T, Hutter B, Sultan M, Cho YJ, Pugh TJ, Hovestadt V, Stütz AM, Rausch T, Warnatz HJ, Ryzhova M, Bender S, Sturm S, Pleier S, Cin H, Pfaff E, Sieber L, Wittmann A, Remke M, Witt H, Hutter S, Tzaridis T, Weischenfeldt J, Raeder B, Avci M, Amstislavskiy V, Zapatka M, Weber UD, Wang Q, Lasitschka B, Bartholomae CC, Schmidt M, von Kalle C, Ast V, Lawerenz C, Eils J, Kabbe R, Benes V, van Sluis P, Koster J, Volckmann R, Shih D, Betts MJ, Russell RB, Coco S, Tonini JP, Schüller U, Hans V, Graf N, Kim YJ, Monoranu C, Roggendorf W, Unterberg A, Herold-Mende C, Milde T, Kulozik AE, von Deimling A, Witt O, Maass E, Rössler J, Ebinger M, Schuhmann MU, Frühwald MC, Hasselblatt M, Jabado N, Rutkowski S, von Bueren AO, Williamson D, Clifford SC, McCabe MG, Collins VP, Wolf S, Wiemann S, Lehrach H, Brors B, Scheurlen W, Felsberg J, Reifenberger G, Northcott PA, Taylor

MD, Meyerson M, Pomeroy SL, Yaspo ML, Korbel JO, Korshunov A, Eils R, Pfister SM, Lichter P (2012). *Dissecting the genomic complexity underlying medulloblastoma*. Nature 2012, in press

Kamburov A, Stelzl U, Herwig R (2012). *IntScore: a web tool for confidence scoring of biological interactions*. Nucleic Acids Res; in press

Klatt S, Konthur Z (2012). *Secretory signal peptide modification for optimized antibody-fragment expression-secretion in Leishmania tarentolae.* Microb Cell Fact 11(1):97

Kühn A, Lehrach H (2012). *The "Virtual Patient" system: modeling cancer using deep sequencing technologies for personalized cancer treatment.* J Cons Prot Food Safety 7(1):55-62

Lebedeva E, Stingl JC, Thal DR, Ghebremedhin E, Strauss J, Ozer E, Bertram L, von Einem B, Tumani H, Otto M, Riepe MW, Hogel J, Ludolph AC, von Arnim CA (2012). *Genetic variants in PSEN2 and correlation to CSF beta-amyloid42 levels in AD.* Neurobiol Aging 33(1):201.e9–18

Lehrach H, The IT Future of Medicine Consortium (2012). *A revolution in healthcare: challenges and opportunities for personalized medicine*. Personalized Medicine 9(2):105-108

Li J, Pandey V, Kessler T, Lehrach H, Wierling C (2012). *Modeling of miRNA and Drug Action in the EGFR-Signaling Pathway. PLoS ONE* 7(1):e30140

Lill CM, Roehr JT, McQueen MB, et al., Bertram L (2012). *Comprehensive research synopsis and systematic meta-analyses in Parkinson's disease genetics: The PDGene database*. PLoS Genet 8(3):e1002548

Lill CM, Schjeide BMM, Akkad DA, et al. (2012). *Independent replication of STAT3 association with multiple sclerosis risk in a large German case-control sample.* Neurogenetics 13(1):83–86

Menzel G, Krebs C, Diez M, Holtgräwe D, Weisshaar B, Minoche AE, Dohm JC, Himmelbauer H, Schmidt T (2012). *Survey of sugar beet (Beta vulgaris L.) hAT transposons and MITE-like hATpin derivatives.* Plant Mol Biol 78:393-405

Montagut C, Dalmases A, Bellosillo B, Crespo M, Pairet S, Iglesias M, Salido M, Gallen M, Marsters S, Tsai SP, Minoche A, Somasekar S, Serrano S, Bellmunt J, Himmelbauer H, Rovira A, Settleman J, Bosch F, Albanell J (2012). *Identification of a mutation in the extracellular domain of the epidermal growth factor receptor conferring cetuximab resistance in colorectal cancer.* Nature Medicine 18:221-223

Pin PA, Zhang W, Vogt SH, Dally N, Büttner B, Schulze-Buxloh G, Jelly NS, Chia TYP, Mutasa-Göttgens ES, Dohm JC, Himmelbauer H, Weisshaar B, Kraus J, Gielen JJL, Lommel M, Weyens G, Wahl B, Schechert A, Nilsson O, Jung C, Kraft T, Müller AE (2012). *A pseudo-response regulator gene controls lifecycle adaptation in beet*. Current Biology, in press

Querfurth R, Fischer A, Schweiger MR, Lehrach H, Mertes F (2012). *Creation and application of immortalized bait libraries for targeted enrichment and next-generation sequencing*. BioTechniques 2012, doi: 10.2144/0000113877

Ralser M, Michel S, Breitenbach M (2012). *Sirtuins as regulators of the yeast metabolic network*. Frontiers in pharmacology 3:32

Schlachter C, Lisdat F, Frohme M, Erdmann VA, Konthur Z, Lehrach H, Glökler J (2012). *Pushing the detection limits: the evanescent field in surface plasmon resonance and analyte-induced folding observations of long human telomeric repeats.* Biosens Bioelectron 31:571-574

Schroeder IS, Sulzbacher S, Nolden T, Fuchs J, Czarnota J, Meisterfeld R, Himmelbauer H, Wobus AM (2012). *Induction and selection of Sox17 expressing endoderm cells generated from murine embryonic stem cells*. Cells Tissues Organs 195(6):507-23

Schueler M*, Zhang Q*, Schlesinger J, Tönjes M and Sperling SR (*shared authorship) (2012). *Dynamics of Srf, p300 and histone modifications during cardiac maturation in mouse.* Mol. BioSyst 8(2):495 - 503

Solomun T, Sturm H, Wellhausen R, Seitz H (2012). *Interaction of a single-stranded DNA-binding protein g5p with DNA Oligonucleotides immobilized on a gold surface*. Chem Phys Lett, in press

Sultan M, Dökel S, Amstislavskiy V, Wuttig D, Sültmann H, Lehrach H, Yaspo ML (2012). *A simple strand-specific RNA-Seq library preparation protocol combining the Illumina TruSeq RNA and the dUTP methods*. Biochem Biophys Res Commun 2012 May 15 [Epub ahead of print]

Tavernier G, Wolfrum K, Demeester J, De Smedt SC, Adjaye J, Rejman J (2012). *Activation of pluripotency-associated genes in mouse embryonic fibroblasts by non-viral transfection with in vitro-derived mRNAs encoding Oct4, Sox2, Klf4 and cMyc.* Biomaterials 33(2):412-7

Tsend-Ayush E, Kortschak RD, Bernard P, Lim SL, Ryan J, Rosenkranz R, Borodina T, Dohm JC, Himmelbauer H, Harley VR, Grützner F (2012). *Identification of mediator complex 26 (Crsp7) on platypus X1 and Y5 sex chromosomes: a candidate testis determining genes in monotremes?* Chromosome Res 20(1):127-138

Wellhausen R, Seitz H (2012). *Facing Current Quantification Challenges in Protein Microarrays*. J Biomed Biotech 2012:831347

Welzel F, Kaehler C, Isau M, Lehrach H, Krobitsch S. (2012). *Fox-2 dependent splicing of ataxin-2 transcript is affected by ataxin-1 overexpression*. PLoS ONE 7(5):e37985

Wierling C, Kühn A, Hache H, Daskalaki A, Maschke-Dutz E, Peycheva S, Li J, Herwig R, Lehrach H. (2012). *Prediction in the face of uncertainty: A Monte Carlo-based approach for systems biology of cancer treatment*. Mutation Research 746(2): 163-70

Won S, Lu Q, Bertram L, Tanzi RE, Lange C (2012). *On the Meta-Analysis of Genome-Wide Association Studies: A Robust and Efficient Approach to Combine Population and Family-Based Studies*. Hum Hered 73(1):35–46

Zorman S, Seitz H, Sclavi B, Strick T (2012). *Topological characterization of the DnaA-oriC complex using single-molecule nanomanipulation*, Nucleic Acids Res 2012

2011

Becla L, Lunshof JE, Gurwitz D, Schulte in den Bäumen T, Westerhoff HV, Lange BMH, Brand A (2011). *Health Technology Assessment in the era of personalized healthcare*. Int J Technol Assessment in Health Care 30:1-9

Bertram L (2011). *Alzheimer's genetics in the GWAS era: a continuing story of "replications and refutations*." Curr Neurol Neurosci Rep 11(3):246–253

Bertram L, Hampel H (2011). *The role of genetics for biomarker development in neurodegeneration*. Prog Neurobiol 95(4):501–504

Bluemlein K, Gruning NM, Feichtinger RG, Lehrach H, Kofler B, Ralser M (2011). *No evidence for a shift in pyruvate kinase PKM1 to PKM2 expression during tumorigenesis*. Oncotarget 2:393-400

Bluemlein K, Ralser M (2011). *Monitoring Protein Expression in whole-cell extracts by targeted label- and standard-free LC-MS/MS.* Nature Protocols 6(6):859-869

Bormann F, Sers C, Seliger B, Handke D, Bergmann T, Seibt S, Lehrach H, Dahl A (2011). *Methylation-specific ligation detection reaction (msLDR): a new approach for multiplex evaluation of methylation patterns*. Mol Genet Genomics 286(3-4):279-91

Borodina T, Adjaye J, Sultan M (2011). *A strand-specific library preparation protocol for RNA sequencing.*Methods Enzymol 500:79-98

Cavill R, Kamburov A, Blagrove M, Athersuch T, Ellis J, Herwig R, Ebbels T, Keun HC. (2011). *Consensus-phenotype integration of transcriptomic and metabolomic data implies a role for metabolism in the chemosensitivity of tumor cells*. PLoS Computat Biol 7(3):e1001113

Chatzinasiou F, Lill CM, Kypreou K, et al. (2011). *Comprehensive field synopsis and systematic meta-analyses of genetic association studies in cutaneous melanoma.* J Natl Cancer Inst 103(16):1227–1235

Chavez L (2011). *Epigenetic Analysis of Human Embryonic Stem Cells – Basics.* Computational Methods Applications SVH ISBN-13: 978-3838127620

Conrad DF, Keebler JE, DePristo MA, Lindsay SJ, Zhang Y, Casals F, Idaghdour Y, Hartl CL, Torroja C, Garimella KV, Zilversmit M, Cartwright R, Rouleau GA, Daly M, Stone EA, Hurles ME, Awadalla P; 1000 Genomes Project* (2011) *Variation in genome-wide mutation rates within and between human families.* Nat Genet 43(7):712-4. *Lehrach H, Sudbrak R (Project Leader), Borodina T, Davydov A, Marquardt P, Mertes F, Nietfeld W, Soldatov A, Timmermann B, Tolzmann M, Albrecht M, Amstislavskiy V, Herwig R & Parkhomchuk D

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R; 1000 Genomes Project Analysis Group* (2011). *The variant call format and VCFtools.* Bioinformatics 27(15):2156-8. *Lehrach H, Sudbrak R (Project Leader), Borodina T, Davydov A,* Marquardt P, Mertes F, Nietfeld W, Soldatov A, Timmermann B, Tolzmann M, Albrecht M, Amstislavskiy V, Herwig R, Parkhomchuk D

Daskalaki A (Editor) (2011). *Digital Forensics for the Health Sciences: Applications in Practice and Research.* IGI-Global Publishing, ISBN: 9781609604837

Diez-Roux G, Banfi S, Sultan M, … Lehrach H, Sarmientos P, Reymond A, Davidson DR, Dollé P, Antonarakis SE, Yaspo ML, Martinez S, Baldock RA, Eichele G, Ballabio A. (2011). *A high-resolution anatomical atlas of the transcriptome in the mouse embryo*. PLoS Biol 9(1):e1000582

Dobbernack G, Meinl W, Schade N, Florian S, Wend K, Voigt I, Himmelbauer H, Gross M, Liehr T, Glatt HR (2011). *Altered tissue distribution of 2-amino-1methyl-6-phenylimidazol[4,5-b]pyridine-DNA adducts in mice transgenic for human sulfotransferases 1A1 and 1A2.* Carcinogenesis 32:1734-1740

Esteve A, Kofler R, Himmelbauer H, Ferretti L, Vivancos AP, Groenen MAM, Folch JM, Rodríguez MC, Pérez-Enciso M (2011). *Partial short-read sequencing of a highly inbred Iberian pig and genomics inference thereof.* Heredity 107:256-264

Georgieva Y, Konthur Z (2011). *Design and Screening of M13 phage display cDNA libraries.* Molecules 16:1167-1183

Göke J, Jung M, Behrens S, Chavez L, O'Keeffe S, Timmermann B, Lehrach H, Adjaye J, Vingron M (2011). *Combinatorial Binding in Human and Mouse Embryonic Stem Cells Identifies Conserved Enhancers Active in Early Embryonic Development*. PLoS Comput Biol 7(12):e1002304

Gravel S, Henn BM, Gutenkunst RN, Indap AR, Marth GT, Clark AG, Yu F, Gibbs RA; 1000 Genomes Project*, Bustamante CD. (2011). *Demographic history and rare allele sharing among human populations*. Proc Natl Acad Sci USA 108(29):11983-8 *MPIMG contributors: Lehrach H, Sudbrak R, Borodina A, Davydov A, Mertes F, Nietfeld W, Soldatov A, Albrecht M, Amstislavskiy V, Herwig R, Parkhomchuk D

Grüning NM, Ralser M (2011). *Cancer: Sacrifice for survival.* Nature 480:190-191

Grüning NM, Rinnerthaler M, Bluemlein K, Mülleder M, Wamelink MMC, Lehrach H, Jakobs C, Breitenbach M, Ralser M (2011). *Pyruvate Kinase Triggers a Metabolic Feedback Loop that Controls Redox Metabolism in Respiring Cells*. Cell Metabolism 14(3) 415-427

Haapasalo A, Viswanathan J, Kurkinen KM, et al., Bertram L (2011) *Involvement of ubiquilin-1 transcript variants in protein degradation and accumulation.* Commun Integr Biol 4(4):428–432

Hallen L, Klein H, Stoschek C, Wehrmeyer S, Nonhoff U, Ralser M, Wilde J, Rohr C, Schweiger MR, Zatloukal K, Vingron M, Lehrach H, Konthur Z, Krobitsch S (2011). *The KRAB-containing zinc-finger transcriptional regulator ZBRK1 activates SCA2 gene transcription through direct interaction with its gene product, ataxin-2.* Hum Mol Genet 20:104-114

Hampel H, Prvulovic D, Teipel S, et al. Bertram L (2011). *The future of Alzheimer's disease: the next 10 years*. Prog Neurobiol 95(4):718–728

Häsler R, Kerick M, Mah N, Hultschig C, … Lehrach H, Nietfeld W, Schreiber S, Rosenstiel P (2011). *Alterations of pre-mRNA splicing in human inflammatory bowel disease*. Eur J Cell Biol 90(6-7):603-11

Haworth A, Bertram L, Carrera P, et al. (2011). *Call for participation in the neurogenetics consortium within the Human Variome Project*. Neurogenetics 12(3):169–173.

Herrmann F, Garriga-Canut M, Baumstark R, Fajardo-Sanchez E, Cotterell J, Minoche A, Himmelbauer H, Isalan M (2011). *p53 Gene Repair with Zinc Finger Nucleases Optimised by Yeast 1-Hybrid and Validated by Solexa Sequencing*. *PLoS One* 6:e20913

Hiltunen M, Bertram L, Saunders AJ (2011). *Genetic risk factors: their function and comorbidities in Alzheimer's disease.* Int J Alzheimers Dis 2011:925362

Iwaniuk KM, Schira J, Weinhold S, Jung M, Adjaye J, Müller HW, Wernet P, Trompeter HI (2011). *Network-Like Impact of MicroRNAs on Neuronal Lineage Differentiation of Unrestricted Somatic Stem Cells from Human Cord Blood.* Stem Cells Dev 20(8):1383-94

Jozefczuk J, Adjaye J (2011). *Quantitative Real-Time PCR-Based Analysis of Gene Expression.* Methods Enzymol 500:99-109

Jozefczuk J, Prigione A, Chavez L, Adjaye J (2011). *Comparative Analysis of Human Embryonic Stem Cell and Induced Pluripotent Stem Cell-Derived Hepatocyte-Like Cells Reveals Current Drawbacks and Possible Strategies for Improved Differentiation.* Stem Cells Dev 20(7):1259-1275

Kamburov A, Cavill R, Ebbels TMD, Herwig R, Keun HC (2011). *Integrated pathway-level analysis of transcriptomics and metabolomics data with IMPaLA*. Bioinformatics 27(20):2917-2918

Kamburov A, Pentchev K, Galicka H, Wierling C, Lehrach H, Herwig R (2011). *ConsensusPathDB: toward a more complete picture of cell biology*. Nucleic Acids Research 39(1):712-717

Kenyon EJ, McEwen GK, Callaway H, Elgar G (2011). *Functional analysis of conserved non-coding regions around the short stature hox gene (shox) in whole zebrafish embryos.* PLoS One 6(6):e21498

Kerick M, Isau M, Timmermann B, Sültmann H, Herwig R, Krobitsch S, Schäfer G, Verdorfer I, Bartsch G, Klocker H, Lehrach H, Schweiger MR (2011). *Targeted High Throughput Sequencing in Clinical Cancer Settings: Formaldehyde fixed-paraffin embedded (FFPE) tumor tissues input amount and tumor heterogeneity*. BMC Medical Genomics 4:68

Kerick M$^{\Theta}$, Isau M$^{\Theta}$, Timmermann B, Sültmann H, Krobitsch S, Bartsch G, Schaefer G, Klocker H, Lehrach H, Schweiger MR (2011). *Technical challenges of targeted DNA Enrichment to Identify Genetic Variations in Cancer Patients.* BMC Medical Genomics 4:68 $^{\Theta}$ shared co-first authors

Kohlmann A, Klein HU, Weissmann S, Bresolin S, Chaplin T, Cuppens H, Garicochea B, Grossmann V, Hanczaruk B, Hebestreit K, Hofer K, Iacobucci I, Jansen JH, Kronnie T, Locht L, Martinelli G, McGowan K, Stabentheiner S, Schweiger MR, Timmermann B, Vandenberghe P, Young BD, Dugas M, Torsten Haferlach (2011). *The Interlaboratory Robustness of Next-Generation Sequencing (IRON) Study: Deep-Sequencing Investigating TET2, CBL, and KRAS Mutations in 5,580 Amplicons by an International Consortium Involving 10 Laboratories.* Leukemia 25(12):1840-8

Köster DM, Haselbach D, Lehrach H, Seitz H (2011). *A DNAzyme based label-free detection system for miniaturized assays*. Mol Biosyst 7(10):2882-9

Kowald A, Wierling C (2011). *Standards, Tools and Databases for the Analysis of Yeast 'omics Data.* In: Castrillo J. I. and Oliver S. G. (Eds.) 'Yeast Systems Biology. Methods and Protocols'. Methods in Molecular Biology (MiMB) series. Vol.

759(20):345-365. (Series Ed: Walker, John M. ISSN: 1064-3745), Humana Press, Springer, New York.

Krüger A, Grüning NM, Wamelink MM, Kerick M, Kirpy A, Parkhomchuk D, Bluemlein K, Schweiger MR, Soldatov A, Lehrach H, Jakobs C, Ralser M (2011). *The pentose phosphate pathway is a metabolic redox sensor and regulates transcription during the antioxidant response*. Antioxid Redox Signal 15(2):311-24

Kühn A, Lehrach H (2011). *The "Virtual Patient" system: modeling cancer using deep sequencing technologies for personalized cancer treatment*. Journal of Consumer Protection and Food Safety. DOI 10.1007/s00003-011-0755-7

Lange BMH, Schweiger RM, Lehrach H (2011). *Global molecular and cellular measurement technologies.* In Cesario A, Marcus FB (editors): *Cancer Systems Biology*: Bioinformatics and Medicine: Research and Clinical Applications, ISBN 978-94-007-1566-0, e-ISBN 978-94-007-1567-7, DOI10.1007/978-94-007-1567-7, Springer Dordrecht Heidelberg London New York

Lehrach H, Sudbrak R, Boyle P, Pasterk M, Zatloukal K, Müller H, Hubbard T, Brand A, Girolami M, Jameson D, Bruggeman FJ, Westerhoff HV (2011). *ITFoM - The IT Future of Medicine*. Procedia Computer Science 7:26-29

Lill CM, Abel O, Bertram L, Al-Chalabi A (2011). *Keeping up with genetic discoveries in amyotrophic lateral sclerosis: the ALSoD and ALSGene databases.* Amyotroph Lateral Scler 12(4):238–249

Lill CM, Bertram L (2011) *Towards unveiling the genetics of neurodegenerative diseases.* Semin Neurol 31(5):531–541

Lim TS, Schütze T, Lehrach H, Glökler J, Konthur Z (2011). *Diversity Visualisation by Endonuclease – A rapid assay to monitor diverse nucleotide libraries.* Anal Biochem 411:16-21

Llorens F, Hummel M, Pastor X, Ferrer A, Pluvinet R, Vivancos A, Castillo E, Iraola S, Mosquera AM, Gonzalez E, Lozano J, Ingham M, Dohm JC, Noguera M, Kofler R, Del Rio JA, Bayés M, Himmelbauer H, Sumoy L (2011). *Multiple platform assessment of the EGF dependent transcriptome by microarray and deep tag sequencing analysis*. BMC Genomics 12:326

Mah N, Wang Y, Liao MC, Prigione A, Jozefczuk J, Lichtner B, Wolfrum K, Haltmeier M, Flöttmann M, Schaefer M, Hahn A, Mrowka R, Klipp E, Andrade-Navarro MA, Adjaye J (2011). *Molecular Insights into Reprogramming-Initiation Events Mediated by the OSKM Gene Regulatory Network.* PLoS One 6(8):e24351

Manolopoulos VG, Dechairo B, Huriez A, Kühn A, LLerena A, van Schaik RH, Yeo K-TJ, Ragia G, Siest G (2011). *Pharmacogenomics and personalized medicine in clinical practice*. Pharmacogenomics 12(5):597-610

Meinel T, Schweiger MR, Ludewig AL, Chenna R, Krobitsch S, Herwig R (2011). *Ortho2ExpressMatrix--a web server that interprets cross-species gene expression data by Gene family information*. BMC Genomics 12:483

Mertes F, Elsharawy A, Sauer S, van Helvoort JM, van der Zaag PJ, Franke A, Nilsson M, Lehrach H, Brookes AJ (2012). *Targeted enrichment*

of genomic DNA regions for next-generation sequencing. Brief Funct Genomics 10(6):374-86

Meunier D, Patra K, Smits R, Hägebarth A, Lüttges A, Jaussi R, Wieduwilt MJ, Quintanilla-Fend L, Himmelbauer H, Fodde R, Fundele R (2011). *Expression analysis of Proline Rich 15 (Prr15) in mouse and human gastrointestinal tumours*. *Mol Carcinogenesis* 50(1):8-15

Mills RE, Walter K, Stewart C, Handsaker RE, Chen K, Alkan C, Abyzov A, Yoon SC, Ye K, Cheetham RK, Chinwalla A, Conrad DF, Fu Y, Grubert F, Hajirasouliha I, Hormozdiari F, Iakoucheva LM, Iqbal Z, Kang S, Kidd JM, Konkel MK, Korn J, Khurana E, Kural D, Lam HY, Leng J, Li R, Li Y, Lin CY, Luo R, Mu XJ, Nemesh J, Peckham HE, Rausch T, Scally A, Shi X, Stromberg MP, Stütz AM, Urban AE, Walker JA, Wu J, Zhang Y, Zhang ZD, Batzer MA, Ding L, Marth GT, McVean G, Sebat J, Snyder M, Wang J, Ye K, Eichler EE, Gerstein MB, Hurles ME, Lee C, McCarroll SA, Korbel JO; 1000 Genomes Project* (2011). *Mapping copy number variation by population-scale genome sequencing.* Nature 470(7332):59-65. *Lehrach H, Sudbrak R (Project Leader), Borodina T, Davydov A,* Marquardt P, Mertes F, Nietfeld W, Soldatov A, Timmermann B, Tolzmann M, Albrecht M, Amstislavskiy V, Herwig R & Parkhomchuk D

Minoche AE, Dohm JC, Himmelbauer H (2011). *Evaluation of genomic high-throughput sequencing data generated on Illumina HiSeq and Genome Analyzer systems*. Genome Biology 12:R112

Müller H, Schmidt D, Dreher F, Herwig R, Ploubidou A, Lange B (2011). *Gene ontology analysis of the centrosome proteomes of Drosophila and human.* Communicative & Integrative Biology 4(3):308-311

Nikolova EV, Herwig R, Nikolov SG, Petrov VG (2011). *Predictive Dynamical Modelling MicroRNAs Role in Complex Networks.* (Ed. Daskalaki A) IGI-Global Publishing ISBN-13: 978-1609604837

Pfeiffer T, Bertram L, Ioannidis JPA (2011). *Quantifying selective reporting and the Proteus phenomenon for multiple datasets with similar bias*. PLoS ONE 6(3):e18362

Postma AV, van Engelen K, van de Meerakker J, Rahman T, Probst S, Baars MJ, Bauer U, Pickardt T, Sperling SR, Berger F, Moorman AF, Mulder BJ, Thierfelder L, Keavney B, Goodship J, Klaassen S (2011). *Mutations in the Sarcomere Protein Gene MYH7 in Ebstein Anomaly.* Circ Cardiovasc Genet 4(1):43-50

Prigione A, Hossini AM, Lichtner B, Serin A, Fauler B, Megges M, Lurz R, Lehrach H, Makrantonaki E, Zouboulis CC, Adjaye J (2011). *Mitochondrial-Associated Cell Death Mechanisms Are Reset to an Embryonic-Like State in Aged Donor-Derived iPS Cells Harboring Chromosomal Aberrations.* PLoS One 6(11):e27352

Prigione A, Lichtner B, Kuhl H, Struys EA, Wamelink M, Lehrach H, Ralser M, Timmermann B, Adjaye J (2011). *Human iPSCs Harbor Homoplasmic and Heteroplasmic Mitochondrial DNA Mutations While Maintaining hESC-Like Metabolic Reprogramming*. Stem Cells 2011. doi: 10.1002/stem.683.

Puente XS, [..] Dohm JC Himmelbauer H, [..] Campo E (2011). *Identification of recurrent mutations in chronic lymphocytic leukaemia*

through whole-genome sequencing. *Nature* 475:101-105

Quesada V, [27 authors], Himmelbauer H, [9 authors] López-Otín C (2011). *Large scale exome-sequencing identifies SF3B1 as a novel gene recurrently mutated in chronic lymphocytic leukemia*. Nature Genetics 44:47-52

Rodelsperger C, Krawitz P, Bauer S, Hecht J, Bigham AW, Bamshad M, de Condor BJ, Schweiger MR, Robinson PN (2011). *Identity-by-descent filtering of exome sequence data for disease-gene identification in autosomal recessive disorders*. Bioinformatics 27(6):829-36

Salton M, Elkon R, Borodina T, Davydov A, Yaspo ML, Halperin E, Shiloh Y (2011). *Matrin 3 binds and stabilizes mRNA*. PLoS One 6(8):e23882

Sarajärvi T, Tuusa JT, Haapasalo A, et al. (2011). *Cysteine 27 variant of the delta-opioid receptor affects amyloid precursor protein processing through altered endocytic trafficking.* Mol Cell Biol 31(11):2326–2340

Schjeide B-MM, Schnack C, Lambert J-C, Lill CM, … Lehrach H, Amouyel P, von Arnim CA, Bertram L (2011). *The role of clusterin, complement receptor 1, and phosphatidylinositol binding clathrin assembly protein in Alzheimer disease risk and cerebrospinal fluid biomarker levels.* Arch Gen Psychiatry 68(2):207–213

Schlesinger J*, Schueler M*, Grunert M*, Fischer JJ*, Zhang Q, Krueger T, Lange M, Tönjes M, Dunkel I, Sperling SR (2011). *The Cardiac Transcription Network Modulated by Gata4, Mef2a, Nkx2.5, Srf, Histone Modifications, and MircoRNAs*. PloS Genet 7(2):e1001313 *shared authorship

Schulz N, Himmelbauer H, Rath M, van Weeghel M, Houten S, Suhre K, Scherneck S, Vogel H, Kluge R, Wiedmer P, Joost HG, Schürmann A (2011). *Role of medium and short-chain L-3-hydroxyacyl-CoA dehydrogenase in the regulation of body weight and thermogenesis.* Endocrinology 152:4641-4651

Schütze T, Rubelt F, Repkow J, Greiner N, Erdmann VA, Lehrach H, Konthur Z, Glökler J (2011). *A streamlined protocol for emulsion PCR and subsequent purification.* Anal Biochem 410:155-157

Schütze T, Wilhelm B, Greiner N, Braun H, Peter F, Mörl M, Erdmann VA, Lehrach H, Konthur Z, Menger M, Arndt PF, Glökler J (2011). *Probing the SELEX process with next-generation sequencing.* PLoS One 6:e29604

Schweiger MR, Kerick M, Timmermann B, Isau M (2011). *The power of NGS technologies to delineate the genome organization in cancer: From mutations to structural variations and epigenetic alterations*. Cancer Metastasis Rev 30(2):199-210

Sperling S (2011). *Systems biology approaches to heart development and congenital heart disease.* Cardiovascular Research 2011; doi: 10.1093/cvr/cvr126

Stehr H, Jang SJ, Duarte JM, Wierling C, Lehrach H, Lappe M, and Lange BMH (2011). *The structural impact of cancer-associated missense mutations in oncogenes and tumor suppressors*. Molecular Cancer 10:54

Suk E-K, McEwen GK, Duitama J, Nowick K, Schulz S, Palczewski S, Schreiber S, Holloway DT, McLaughlin S, Peckham H, Lee C, Huebsch T, and Hoehe MR (2011). *A Comprehensively Molecular*

Haplotype-Resolved Genome of a European Individual. Genome Res 21(10):1672-85

Thormann A, Rasche A (2011). *Functional Context Network of T2DM Medical Complications of Type 2 Diabetes*. (Ed. Croniger C.) InTech Publishing ISBN-13: 978 9533073637

van Engelen K, Postma AV, van de Meerakker JB, Roos-Hesselink JW, Helderman-van den Enden AT, Vliegen HW, Rahman T, Baars MJ, Sels JW, Bauer U, Pickardt T, Sperling SR, Moorman AF, Keavney B, Goodship J, Klaassen S, Mulder BJ (2011). *Ebstein´s anomaly may be caused by mutations in the sarcomere protein gene MYH7*. Netherlands Heart Jounal. doi: 10.1007/s12471-011-0141-1

Vilardell-Nogales M, Rasche A, Thormann A, Maschke-Dutz E, Perez-Jurado LA, Lehrach H, Herwig R (2011). *Meta-analysis of heterogeneous Down Syndrome data reveals_consistent genome-wide dosage effects related to neurological processes*. BMC Genomics 12:229

Viswanathan J, Haapasalo A, Böttcher C, et al. (2011). *Alzheimer's disease-associated ubiquilin-1 regulates presenilin-1 accumulation and aggresome formation*. Traffic 12(3):330–348

von Keyserling H, Bergmann T, Schuetz M, Schiller U, Stanke J, Hoffmann C, Schneider A, Lehrach H, Dahl A, Kaufmann AM (2011). *Analysis of 4 Single-Nucleotide Polymorphisms in Relation to Cervical Dysplasia and Cancer Development Using a High-Throughput Ligation-Detection Reaction Procedure.* Int J Gynecol Cancer 21(9):1664-71

Warnatz HJ, Schmidt D, Manke T, Piccini I, Sultan M, Borodina T, Balzereit D, Wruck W, Soldatov A, Vingron M, Lehrach H, Yaspo ML (2011). *The BTB and CNC homology 1 (BACH1) target genes are involved in the oxidative stress response and in control of the cell cycle.* J Biol Chem 286(26):23521-32

Yildirimman R, Brolén G, Vilardell M, Eriksson G, Synnergren J, Gmuender H, Kamburov A, Ingelman-Sundberg M, Castell J, Lahoz A, Kleinjans J, van Delft J, Björquist P, Herwig R. (2011). *Human embryonic stem cell derived hepatocyte-like cells as a tool for in vitro hazard_assessment of chemical carcinogenicity*. Toxicological Sciences 124(2):278-290

Zakrzewski F, Weisshaar B, Fuchs J, Bannack E, Minoche AE, Dohm JC, Himmelbauer H, Schmidt T (2011). *Epigenetic profiling of heterochromatic satellite DNA*. Chromosoma 120:409-422

Ziegelberger G, Baum C, Borkhardt A, Cobaleda C, Dasenbrock C, Dehos A, Grosche B, Hauer J, Hornhardt S, Jung T, Kammertoens T, Lagroye I, Lehrach H, Lightfoot T, Little MP, Rossig C, Sanchez-Garcia I, Schrappe M, Schuez J, Shalapour S, Slany R, Stanulla M, Weiss, W (2011). *Research recommendations toward a better understanding of the causes of childhood leukemia*. Blood Cancer Journal 1:e1; doi:10.1038/bcj.2010.1

Ziegler G, Freyer D, Harhausen D, Khojasteh U, Nietfeld W, Trendelenburg G (2011). Blocking *TLR2 in vivo protects against accumulation of inflammatory cells and neuronal injury in experimental stroke*. J Cereb Blood Flow Metab 31(2):757-66

Zuccotti M, Merico V, Bellone M, Mulas F, Sacchi L, Rebuzzini P, Prigione A, Redi CA, Bellazzi R, Adjaye J, Garagna S. (2011).

Gatekeeper of pluripotency: A common Oct4 transcriptional network operates in mouse eggs and embryonic stem cells. BMC Genomics 12(1):345

2010

Abrahams JP, Apweiler R, Balling R, Bertero MG, Bujnicki JM, Chayen NE, Chène P, Corthals GL, Dylag T, Förster F, Heck AJ, Henderson PJ, Herwig R, Jehenson P, Kokalj SJ, Laue E, Legrain P, Martens L, Migliorini C, Musacchio A, Podobnik M, Schertler GF, Schreiber G, Sixma TK, Smit AB, Stuart D, Svergun D, Taussig MJ (2010). *4D Biology for Health and Disease" Workshop Report*. New Biotechnology 28(4):291-293

Balbach ST, Esteves TC, Brink T, Gentile L, McLaughlin KJ, Adjaye JA, Boiani M (2010). *Governing cell lineage formation in cloned mouse embryos*. Dev Biol 343:71-83

Bertram L (2010). *Field Synopsis of Genetic Association Studies in Schizophrenia*. In Human Genome Epidemiology (Second Edition);. Khoury MJ, G.M., Bradley L, Little J, Higgins J, Ioannidis JPA, editor. Oxford Press.

Bertram L, Heekeren H (2010). *Obesity and the brain: a possible genetic link*. Alzheimers Res Ther 2(5):27

Bertram L, Lill CM, Tanzi RE (2010). *The genetics of Alzheimer disease: back to the future.* Neuron 68(2):270–281

Bertram L, Tanzi RE (2010). *Alzheimer disease: New light on an old CLU*. Nat Rev Neurol 6(1):11–13

Boerno ST, Grimm C, Lehrach H, Schweiger MR (2010). *Next Generation Sequencing Technologies for DNA Methylation Analyses in Cancer Genomics*. Epigenomics 2(2):199-209

Bourbeillon J, Orchard S, Benhar I, Borrebaeck C, de Daruvar A, Dübel S, Frank R, Gibson F, Gloriam D, Haslam N, Hiltke T, Humphrey-Smith I, Hust M, Juncker D, Koegl M, Konthur Z, Korn B, Krobitsch S, Myldermans S, Nygren PA, Palcy S, Polic B, Rodriguez H, Sawyer A, Schlapschy M, Snyder M, Stoevesandt O, Taussig MJ, Templin M, Uhlen M, van der Maarel S, Wingren C, Hermjakob H, Sherman D (2010). *Minimum information about a protein affinity reagent (MIAPAR).* Nat Biotechnol 28:650-653

Castaldi PJ, Cho MH, Cohn M, et al. Bertram L (2010). *The COPD genetic association compendium: a comprehensive online database of COPD genetic associations.* Hum Mol Genet 19(3):526–534

Chavez L, Jozefczuk J, Grimm C, Dietrich J, Timmermann B, Lehrach H, Herwig R*, Adjaye J* (2010). *Computational analysis of genome-wide DNA-methylation during the differentiation of human embryonic stem cells along the endodermal lineage*. Genome_Research 20:1441-1450 *equal contribution

Cheng X, Guerasimova A, Manke T, Rosenstiel P, Haas S, Warnatz HJ, Querfurth R, Nietfeld W, Vanhecke D, Lehrach H, Yaspo ML, Janitz M (2010). *Screening of human gene promoter activities using transfected-cell arrays*. Gene 450(1-2):48-54

Chung HR, Dunkel I, Heise F, Linke C, Krobitsch S, Ehrenhofer-Murray AE, Sperling SR, Vingron M (2010). *The effect of micrococcal nuclease digestion on nucleosome positioning data*. PLoS ONE 5:e15754.

Dahl A, Mertes F, Timmermann B, Lehrach H (2010). *The application of massively parallel sequencing technologies in diagnostics*. *F1000 Biol Rep* 2:59

Daskalaki A, Rasche A (2010). *Meta-Analysis approach for the identification of molecular networks related to infections of the oral cavity*. In Informatics in Oral Medicine: Advanced Techniques in Clinical and Diagnostic Technologies. Daskalaki, Andriani (Ed.), Hershey, PA: Medical Information Science Reference. ISBN: 978-160566733-1.

Denekamp NY, Reinhardt R, Albrecht MW, Drungowski M, Kube M, Lubzens E (2010). *The expression pattern of dormancy-associated genes in multiple life-history stages in the rotifer Brachionus plicatilis*. Hydrobiologia 662(1):51-63

Di Stefano B, Buecker C, Ungaro F, Prigione A, Chen HH, Welling M, Eijpe M, Mostoslavsky G, Tesar P, Adjaye J, Geijsen N, Broccoli V (2010). *An ES-like Pluripotent state in FGF-dependent murine iPS cells*. PLoS ONE 5:e16092

Dolan SM, Hollegaard MV, Merialdi M, Betran AP, Allen T, Abelow C, Nace J, Lin BK, Khoury MJ, Ioannidis JP, Bagade S, Zheng X, Dubin RA, Bertram L, Velez Edwards DR, Menon R (2010). *Synopsis of preterm birth genetic association studies:the preterm birth genetics knowledge base (PTBGene)*. Public Health Genomics.13:514-523

Doss MX, Wagh V, Schulz H, Kull M, Kolde R, Pfannkuche K, Nolden T, Himmelbauer H, Vilo J, Hescheler H, Sachinidis S (2010). *Global transcriptomic analysis of murine embryonic stem cell-derived brachyury (T) cells*. Genes to Cells 15:209-228

Duitama J, Huebsch T, McEwen G, Suk E-K, Hoehe MR (2010). *ReFHap: A reliable and fast algorithm for single individual haplotyping*. In: Proceedings of the First ACM International Conference on Bioinformatics and Computational Biology, pp. 160-169. ACM, Niagara Falls, New York

Ebner B, Panopoulou G, Vinogradov SN, Kiger L, Marden MC, Burmester T, Hankeln T (2010). *The globin gene family of the cephalochordate amphioxus: implications for chordate globin evolution*. BMC Evol Biol 10:370

Emde AK, Grunert M, Weese D, Reinert K, Sperling SR (2010). *MicroRazerS: rapid alignment of small RNA reads*. Bioinformatics 26:123–124

Galiveti CR, Rozhdestvensky TS, Brosius J, Lehrach H, Konthur Z (2010). *Application of housekeeping npcRNAs for quantitative expression analysis of human transcriptome by real-time PCR*. RNA 16:450-461

Giedraitis V, Glaser A, Sarajärvi T, et al. (2010). *CALHM1 P86L polymorphism does not alter amyloid-beta or tau in cerebrospinal fluid*. Neurosci Lett469(2):265–267

Glökler J, Schütze T, Konthur Z. (2010). *Automation in the High-Throughput Selection of Random Combinatorial Libraries — Different Approaches for Respective Applications*. Molecules 15:2478-2490

Gloriam DE, Orchard S, Bertinetti D, Björling E, Bongcam-Rudloff E, Bourbeillon J, Bradbury AR, de Daruvar A, Dübel S, Frank R, Gibson TJ, Haslam N, Herberg FW, Hiltker T, Hoheisel JD, Kerrien S, Koegel M, Konthur Z, Korn B, Landegren U, van der Maarel S, Montecchi-Palazzi L,

Palcy S, Rodriguez H, Schweinsberg S, Sievert V, Stoevesandt O, Taussig MJ, Uhlén M, Wingren C, Woollard P, Sherman DJ, Hermjakob H (2010). *A community standard format for the representation of protein affinity reagents*. Mol Cell Proteomics 9:1-10

Groth D, Hartmann S, Friemel M, Hill N, Muller S, Poustka AJ, Panopoulou G (2010). *Data integration using scanners with SQL output--the bioscanners project at sourceforge*. J Integr Bioinform 7

Grüning NM, Lehrach H, Ralser M (2010). *Regulatory crosstalk of the metabolic network.* Trends in Biochemical Sciences 35(4):220-7

Haapasalo A, Viswanathan J, Bertram L, et al. (2010). *Emerging role of Alzheimer's disease-associated ubiquilin-1 in protein aggregation.* Biochem Soc Trans 38(Pt 1):150–155

Habermann K, Lange BMH (2010). *Centrosomes: Methods for Preparation.* In: *Encyclopedia of Life Sciences* 2010, John Wiley & Sons, Ltd: Chichester

Harhausen D, Khojasteh U, Stahel PF, Morgan BP, Nietfeld W, Dirnagl U, Trendelenburg G (2010). *Membrane attack complex inhibitor CD59a protects against Focal cerebral ischemia in mice*. J Neuroinflammation 7:15

Harhausen D, Prinz V, Ziegler G, Gertz K, Endres M, Lehrach H, Gasque P, Botto M, Stahel PF, Dirnagl U, Nietfeld W, Trendelenburg G (2010). *CD93/AA4.1: a novel regulator of inflammation in murine focal cerebral ischemia*. J Immunol 184:6407-6417

Hu Y, Lehrach H, Janitz M (2010). *Apoptosis screening of human chromosome 21 proteins reveals novel cell death regulators*. Mol Biol Rep 37(7):3381-7

Jozefczuk J, Stachelscheid H, Chavez L, Herwig R, Lehrach H, Zeilinger K, Gerlach JC, Adjaye J (2010). *Molecular Characterization of Cultured Adult Human Liver Progenitor Cells.* Tissue Eng Part C Methods 16(5):821-834

Jung M, Peterson H, Chavez L, Kahlem P, Lehrach H, Vilo J, Adjaye J (2010). *A Data Integration Approach to Mapping OCT4 Gene Regulatory Networks Operative in Embryonic Stem Cells and Embryonal Carcinoma Cells*. PLoS One 5(5):e10709

Jürchott K, Kuban RJ, Krech T, Blüthgen N, Stein U, Walther W, Friese C, Kielbasa SM, Ungethüm U, Lund P, Knösel T, Kemmner W, Morkel M, Fritzmann J, Schlag PM, Birchmeier W, Krueger T, Sperling SR, Sers C, Royer HD, Herzel H, Schäfer R (2010). *Identification of Y-Box Binding Protein 1 As a Core Regulator of MEK/ERK Pathway-Dependent Gene Signatures in Colorectal Cancer Cells.* PLoS Genet 6(12):e1001231

Kerick M, Timmermann B, Schweiger MR (2010). *High-throughput sequencing of frozen and paraffin-embedded tumor and normal tissue.* Pathologe 31 Suppl 2:255-257

Konthur Z, Wilde J (2010) *Evaluation of recombinant antibodies on protein microarrays applying the multiple spotting technique.* In Antibody Engineering Vol. 2, 2nd Edition, Springer-Verlag (R. Kontermann & S. Dübel, Eds.), ISBN: 978-3-642-01146-7, pp. 447-460

Konthur Z, Wilde J, Lim TS (2010). *Semi-automated magnetic bead-based antibody selection from phage display libraries.* In Antibody Engineering

Vol. 1, 2nd Edition, Springer-Verlag (R. Kontermann & S. Dübel, Eds.), ISBN: 978-3-642-01143-6, pp. 265-285

Krawitz PM ϴ, Schweiger MRϴ, Rödelsperger C, Isau M, Fischer A, Dahl A, Jonske de Condor B, Kölsch U, Meisel C, Kinoshita T, Murakami Z, Hecht J, Brunner H, Meinecke P, Horn D, Mundlos S, Robinson PN (2010). *Identity-by-Descent Filtering of Exome Sequence Data identifies PIGV mutations in Hyperphosphatasia Mental Retardation (HPMR) syndrome*. Nat Genetics 42(10):827-9
ϴ shared co-first authors

Lambert JC, Sleegers K, González-Pérez A, et al. Bertram L (2010). *The CALHM1 P86L polymorphism is a genetic modifier of age at onset in Alzheimer's disease: a meta-analysis study*. J Alzheimers Dis 22(1):247–255

Lamina C, Küchenhoff H, Chang-Claude J, Paulweber B, Wichmann HE, Illig T, Hoehe MR, Kronenberg F, Heid IM (2010). *Haplotype Misclassification Resulting from Statistical Reconstruction and Genotype Error, and Its Impact on Association Estimates.* Annals of Human Genetics 00:1–11

Lange BM, Schweiger MR, Lehrach H (2010). *Global molecular and cellular measurement technologies*. In: Cancer Systems Biology, Bioinformatics and Medicine: Research and Clinical Applications. A., M.F.a.C., editor. Springer

Lange BMH (2010). *Protein complexes and their function in cell division pathways in Health and disease.* Cell News 36:28-31

Lange C, Mittermayr L, Dohm JC, Holtgräwe D, Weisshaar B, Himmelbauer H (2010). *High-throughput identification of genetic markers using representational oligonucleotide microarray analysis*. Theoretical and Applied Genetics 121:549-565

Laumet G, Chouraki V, Grenier-Boley B, et al. Bertram L (2010). *Systematic analysis of candidate genes for Alzheimer's disease in a French, genome-wide association study.* J Alzheimers Dis 20(4):1181–1188

Lill CM, Schjeide BMM, Roehr JT, et al. (2010). *Correspondence to Sand et Al. "Critical reappraisal of a catechol-o-methyltransferase transversion variant in schizophrenia."* Biol Psychiatry 67(7):e45–48

Lim TS, Mollova S, Rubelt F, Sievert V, Dübel S, Lehrach H, Konthur Z (2010). *V–gene amplification revisited – An optimised procedure for human antibody isotype and idiotype amplification*. N Biotechnol 27:108-117

Lunshof JE, Bobe J, Aach J, Angrist M, Thakuria JV, Vorhaus DB, Hoehe MR, Church GM (2010). *Personal genomes in progress: from the Human Genome Project to the Personal Genome Project*. Dialogues Clin Neurosci 12(1):47-60

Makrantonaki E, Schonknecht P, Hossini AM, Kaiser E, Katsouli MM, Adjaye J, Schroder J, Zouboulis CC (2010). *Skin and brain age together: The role of hormones in the ageing process*. Exp Gerontol 45:801-813

Mann K, Poustka AJ, Mann M (2010). *Phosphoproteomes of Strongylocentrotus purpuratus shell and tooth matrix: identification of a major acidic sea urchin tooth phosphoprotein, phosphodontin*. Proteome Science 8(1):6

Mann K, Wilt FH, Poustka AJ (2010). *Proteomic analysis of sea urchin*

(Strongylocentrotus purpuratus) spicule matrix. Proteome Science 8(1):33

Marklein B, Konthur Z, Cope A, Shlomchik M, Burmester G, Skriner K (2010). *JKTBP (HNRNPDL) is a novel toll7/9 dependent target in humans and animal models of inflammatory rheumatic diseases*. Ann Rheum Dis 69:A4

Marklein B, Konthur Z, Häupl T, Cope A, Shlomchik M, Steiner G, Burmester G, Skriner K (2010). *Toll-like receptor (TLR-7/9), MYOD88, TIR8, dependent and independent autoantigens*. Ann Rheum Dis 69:39

McMeekin TA, Hill C, Wagner M, Dahl A, Ross T (2010). *Ecophysiology of food borne pathogens: Essential knowledge to improve food safety*. Int J Food Microbiol 139 Suppl 1:S64-78

Meinel T, Mueller MS, Ahmed J, Yildirimman R, Dunkel M, Herwig R, Preissner R (2010). *SOAP/WSDL-Based Web Services for Biomedicine: Demonstrating the Technique with the CancerResource*. Proceedings of the MEDICON 2010 29(5):835-838

Mertes F, Biens K, Lehrach H, Wagner M, Dahl A (2010). *High-throughput Universal Probe Salmonella Serotyping (UPSS) by nanoPCR*. J Microbiol Methods 83(2):217-23

Müller H, Schmidt D, Steinbrink S, Mirgorodskaya E, Lehmann V, Habermann K, Dreher F, Gustavsson N, Kessler T, Lehrach H, Herwig R, Gobom J, Ploubidou A, Boutros M, Lange BMH (2010). *Proteomic and functional analysis of the mitotic Drosophila centrosome*. EMBO Journal 29:3344-3357

Pentchev K, Ono K, Herwig R, Ideker T, Kamburov A (2010). *Evidence mining and novelty assessment of protein-protein interactions with the ConsensusPathDB plugin for Cytoscape*. Bioinformatics 39(1):712-717

Pihlajamäki M, O' Keefe K, Bertram L, et al. (2010). *Evidence of altered posteromedial cortical FMRI activity in subjects at risk for Alzheimer disease.* Alzheimer Dis Assoc Disord 24(1):28–36

Prigione A, Adjaye J (2010). *Modulation of mitochondrial biogenesis and bioenergetic Metabolism upon in vitro and in vivo differentiation of human ES and iPS cells*. Int J Dev Biol 54(11-12):1729-41

Prigione A, Fauler B, Lurz R, Lehrach H, Adjaye J (2010). *The senescence-related mitochondrial/oxidative stress pathway is repressed in human induced pluripotent stem cells.* Stem Cells 28:721-733

Rasche A, Yildirimman R, Herwig R (2010). *Handbook of Systems Toxicology ch. Integrative Analysis of microarray data - a path for systems toxicology*. In Handbook of Systems Toxicology (Eds. D. A. Casciano and S. C. Sahu). John Wiley & Sons Ltd ISBN-13: 978-0-470-68401-6

Ravasi T, Suzuki H, Cannistraci CV, Katayama S, Bajic VB, Tan K, Akalin A, Schmeier S, Kanamori-Katayama M, Bertin N, Carninci P, Daub CO, Forrest AR, Gough J, Grimmond S, Han JH, Hashimoto T, Hide W, Hofmann O, Kamburov A, Kaur M, Kawaji H, Kubosaki A, Lassmann T, van Nimwegen E, MacPherson CR, Ogawa C, Radovanovic A, Schwartz A, Teasdale RD, Tegner J, Lenhard B, Teichmann SA, Arakawa T, Ninomiya N, Murakami K, Tagami M, Fukuda S, Imamura K, Kai C, Ishihara R, Kitazume Y, Kawai J, Hume DA, Ideker T, Hayashizaki Y (2010). *An*

atlas of combinatorial transcriptional regulation in mouse and man. Cell 140:744-752

Richard H, Schulz M, Sultan M, Nürnberger A, Schrinner S, Balzereit D, Dagand E, Rasche A, Lehrach H, Vingron M, Haas SA, Yaspo ML (2010). *Prediction of Alternative Isoforms from Exon Expression Levels in RNA-Seq Experiments.* Nucleic Acids Research 38(10):e112

Saykin AJ, Shen L, Foroud TM, et al. Bertram L (2010). *Alzheimer's Disease Neuroimaging Initiative biomarkers as quantitative phenotypes: Genetics core aims, progress, and plans.* Alzheimers Dement 6(3):265–273

Schlesinger J*, Schueler M*, Grunert M*, Fischer JJ*, Zhang Q, Krueger T, Lange M, Tönjes M, Dunkel I, and Sperling SR (*shared authorship) (2010) *The Effect of Micrococcal Nuclease Digestion on Nucleosome Positioning Data.* PLoS ONE 5(12):e15754

Schlesinger J, Tönjes M, Schueler M, Zhang Q, Dunkel I, Sperling SR (2010). *Evaluation of the LightCycler© 1536 Instrument for high-throughput quantitative real-time PCR.* Methods 50(4):S19-S22

Scholz AK, Klebl BM, Morkel M, Lehrach H, Dahl A, Lange BMH (2010). *A flexible multi-well format for immunofluorescence screening microscopy of small molecule inhibitors.* ASSAY and Drug Development Technologies 8:571-80

Schütze T, Arndt PF, Menger M, Wochner A, Vingron M, Erdmann VA, Lehrach H, Kaps C, Glökler J (2010). *A calibrated diversity assay for nucleic acid libraries using DiS-tRO--a Diversity Standard of Random Oligonucleotides*. Nucleic Acids Res 38(4):e23

Schütze T, Rubelt F, Repkow J, Greiner N, Erdmann VA, Lehrach H, Konthur Z, Glökler J (2010). *A streamlined protocol for emulsion polymerase chain reaction and subsequent purification*. Anal Biochem 410:155-157

Schwibbert K, Marin-Sanguino A, Bagyan I, Heidrich G, Lentzen G, Seitz H, Rampp M, Schuster SC, Klenk HP, Pfeiffer F, Oesterheld D, Kunte HJ (2010). *A blueprint of ectoine metabolism from the genome of the industrial producer Halomonas elongata DSM 2581 (T)*. Environ Microbiol 13(8):1973-94

Sleegers K, Lambert JC, Bertram L, Cruts M, Amouyel P, Van Broeckhoven C (2010). *The pursuit of susceptibility genes for Alzheimer's disease: progress and prospects*. Trends Genet 26:84-93

Sleegers K, Lambert JC, Bertram L, et al. (2010). *The pursuit of susceptibility genes for Alzheimer's disease: progress and prospects.* Trends Genet 26(2):84–93

Sudmant PH, Kitzman JO, Antonacci F, Alkan C, Malig M, Tsalenko A, Sampas N, Bruhn L, Shendure J; 1000 Genomes Project*, Eichler EE (2010). *Diversity of human copy number variation and multicopy genes.* Science 330(6004):641-6. *MPIMG contributors: Lehrach H, Sudbrak R (Project Leader), Borodina T, Davydov A, Marquardt P, Mertes F, Nietfeld W, Soldatov A, Timmermann B, Tolzmann M, Albrecht M, Amstislavskiy V, Herwig R, Parkhomchuk D

Tesfaye D, Regassa A, Rings F, Ghanem N, Phatsara C, Tholen E, Herwig R, Un C, Schellander K, Hoelker M (2010). *Suppression of the transcription factor MSX1 gene delays bovine preimplantation embryo*

development in vitro. Reproduction 139(5):857-870

The 1000 Genomes Project Consortium* (2010). *A map of human genome variation from population scale sequencing*. Nature 467(7319):1061-1073 *MPIMG contributors: Lehrach H, Sudbrak R (Project Leader), Borodina T, Davydov A, Marquardt P, Mertes F, Nietfeld W, Soldatov A, Timmermann B, Tolzmann M, Albrecht M, Amstislavskiy V, Herwig R, Parkhomchuk D

The FANTOM Consortium*. (2010). *An atlas of combinatorial transcriptional regulation in mouse and man*. Cell 140(5):744-752 *MPIMG contributor: Kamburov A

The International Cancer Genome Consortium* (2010). *International network of cancer genome projects.* Nature 464:993-998. *MPIMG contributors: Lehrach H, Himmelbauer H

Timmermann B, Jarolim S, Russmayer H, Kerick M, Michel S, Kruger A, Bluemlein K, Laun P, Grillari J, Lehrach H, Breitenbach M, Ralser M (2010). *A new dominant peroxiredoxin allele identified by whole-genome re-sequencing of random mutagenized yeast causes oxidant-resistance and premature aging*. Aging 2:475-486

Timmermann B[Θ], Kerick M[Θ], Röhr C, Fischer A. Isau M, Boerno S, Wunderlich A, Barmeyer C, Seemann P, Koenig J, Lappe M, Kuss AW, Garshasbi M, Bertram L, Trappe K, Werber M, Herrmann BG, Zatloukal K, Lehrach H, SchweigerMR (2010). *Somatic mutation profiles of MSI and MSS colorectal cancer identified by whole exome next generation sequencing and bioinformatics analysis.* PLoS One 5(12):e15661
[Θ] shared co-first authors

Ungethuem U, Haeupl T, Witt H, Koczan D, Krenn V, Huber H, von Helversen TM, Drungowski M, Seyfert C, Zacher J, Pruss A, Neidel J, Lehrach H, Thiesen HJ, Ruiz P, Blaess S (2010). *Molecular signatures and new candidates to target the Pathogenesis of rheumatoid arthritis*. Physiol Genomics 42A(4):267-82

Vivancos AP*, Güell M*, Dohm JC*, Serrano L, Himmelbauer H (2010). *Strand-specific sequencing of the transcriptome. Genome Research* 20:989-999 *equal contribution

von Kampen O, Lipinski S, Till A, Martin SJ, Nietfeld W, Lehrach H, Schreiber S, Rosenstiel P (2010). *The caspase recruitment domain-containing protein 8 (CARD8) negatively regulates NOD2-mediated signaling*. J Biol Chem 285(26):19921-6

Wamelink MM, Gruning NM, Jansen EE, Bluemlein K, Lehrach H, Jakobs C, Ralser M (2010). *The difference between rare and exceptionally rare: molecular characterization of ribose 5-phosphate isomerase deficiency*. J Mol Med 88:931-939

Warnatz HJ, Querfurth R, Guerasimova A, Cheng X, Haas SA, Hufton AL, Manke T, Vanhecke D, Nietfeld W, Vingron M, Janitz M, Lehrach H, Yaspo ML (2010). *Functional analysis and identification of cis-regulatory elements of human chromosome 21 gene promoters*. Nucleic Acids Res 38(18):6112-23

Warren WC, Clayton DF, Ellegren H, Arnold AP, Hillier LW, Künstner A, Searle S, White S, Vilella AJ, Fairley S, Heger A, Kong L, Ponting CP, Jarvis ED, Mello CV, Minx P, Lovell P, Velho TA, Ferris M, Balakrishnan CN, Sinha S, Blatti C, London SE, Li Y, Lin YC, George J, Sweedler J, Southey B, Gunaratne P, Watson

M, Nam K, Backström N, Smeds L, Nabholz B, Itoh Y, Whitney O, Pfenning AR, Howard J, Völker M, Skinner BM, Griffin DK, Ye L, McLaren WM, Flicek P, Quesada V, Velasco G, Lopez-Otin C, Puente XS, Olender T, Lancet D, Smit AF, Hubley R, Konkel MK, Walker JA, Batzer MA, Gu W, Pollock DD, Chen L, Cheng Z, Eichler EE, Stapley J, Slate J, Ekblom R, Birkhead T, Burke T, Burt D, Scharff C, Adam I, Richard H, Sultan M, Soldatov A, Lehrach H, Edwards SV, Yang SP, Li X, Graves T, Fulton L, Nelson J, Chinwalla A, Hou S, Mardis ER, Wilson RK (2010). *The genome of a songbird.* Nature 464(7289):757-62

Welte Y, Adjaye J, Lehrach HR, Regenbrecht CR (2010). *Cancer stem cells in solid tumors: elusive or illusive?* Cell Commun Signal 8:6

Wild R, P.S., Popovic M, Zappi M, Dufreche S, Bajpai R (2010). *Lipids from Lipomyces starkeyi*. Food Technol Biotechnol 48 (3):329-335

Wolfrum K, Wang Y, Prigione A, Sperling K, Lehrach H, Adjaye J (2010). *The LARGE principle of cellular reprogramming: lost, acquired and retained gene expression in foreskin and amniotic fluid-derived human iPS cells.* PLoS ONE 5:e13703

Wunderlich A, Krobitsch S, Konthur Z, Lange B, Glökler J, Hussong M, Howley P, Lehrach H, Schweiger M (2010). *The genome maintenance and transcriptional regulation mechanisms of human high risk papillomaviruses as target points for antiviral treatment and cervical cancer prevention*. Onkologie 33(suppl 2):115-128

Ziegler G, Freyer D, Harhausen D, Khojasteh U, Nietfeld W, Trendelenburg G (2010). *Blocking TLR2 in vivo protects against accumulation of inflammatory cells and neuronal injury in experimental stroke*. J Cereb Blood Flow Metab 31(2): 757–766

2009

Aebersold R, Auffray C, Baney E, Barillot E, Brazma A, Brett C, Brunak S, Butte A, Califano A, Celis J, Cufer T, Ferrell J, Galas D, Gallahan D, Gatenby R, Goldbeter A, Hace N, Henney A, Hood L, Iyengar R, Jackson V, Kallioniemi O, Klingmüller U, Kolar P, Kolch W, Kyriakopoulou C, Laplace F, Lehrach H, Marcus F, Matrisian L, Nolan G, Pelkmans L, Potti A, Sander C, Seljak M, Singer D, Sorger P, Stunnenberg H, Superti-Furga G, Uhlen M, Vidal M, Weinstein J, Wigle D, Williams M, Wolkenhauer O, Zhivotovsky B, Zinovyev A, Zupan B. *Report on EU-USA workshop: how systems biology can advance cancer research*. Mol Oncol 3(1):9-17

Baek YS, Haas S, Hackstein H, Bein G, Hernandez-Santana M, Lehrach H, Sauer S, Seitz H (2009). *Identification of novel transcriptional regulators involved in macrophage differentiation and activation in U937 cells*. BMC Immunol 10:18.E

Bertram L (2009). *Alzheimer's disease genetics current status and future perspectives*. Int Rev Neurobiol 84:167–184

Bertram L, Tanzi RE (2009). *Genome-wide association studies in Alzheimer's disease*. Hum Mol Genet 18(R2):R137–145

Bíliková K, Mirgorodskaya E, Bukovská G, Gobom J, Lehrach H, Simúth J (2009). *Towards functional proteomics of minority component of honeybee royal jelly: The effect of post-translational modifications on the antimicrobial activity of apalbumin2*. Proteomics 9(8):2131-8

Bradham C, Oikonomou, Kühn A, Core AB, Modell JW, McClay DR, Poustka AJ (2009). *Chordin is required for neural but not axial development in sea urchin embryos*. Current Opinion in Genetics & Development 19:1–7

Brink TC, Demetrius L, Lehrach H, Adjaye J (2009). *Age-related transcriptional changes in gene expression in different organs of mice support the metabolic stability theory of aging*. Biogerontology 10(5):549-64

Brink TC, Regenbrecht C, Demetrius L, Lehrach H, Adjaye J (2009). *Activation of the immune response is a key feature of aging in mice*. Biogerontology 10(6): 721–734

Chavez L, Bais AS, Vingron M, Lehrach H, Adjaye J, Herwig R (2009). *In silico identification of a core regulatory network of OCT4 in human embryonic stem cells using an integrated approach*. BMC Genomics 10:314

Daskalaki A, Wierling C, Herwig R (2009). *Computational tools and resources for systems biology approaches in cancer*. In Computational Biology - Issues and Applications in Oncology, Series: Applied Bioinformatics and Biostatistics in Cancer Research, Pham, Tuan (Ed.), Springer, New York Dordrecht Heidelberg London. 2009:227-242. Springer Press

Dohm JC, Lange C, Reinhardt R, Himmelbauer H (2009). Haplotype divergence in *Beta vulgaris* and microsynteny with sequenced plant genomes. *Plant Journal* 57:14-26

Dreja T, Jovanovic Z, Rasche A, Kluge R, Herwig R, Tung YCL, Joost HG, Yeo GSH, Al Hasani H (2009). *Diet-induced gene expression of isolated pancreatic islets from a polygenic mouse model for the metabolic syndrome*. Diabetologia 53(2):309-20

Fathi A, Pakzad M, Taei A, Brink TC, Pirhaji L, Ruiz G, Sharif Tabe Bordbar M, Gourabi H, Adjaye J, Baharvand H, Salekdeh GH (2009). *Comparative proteome and transcriptome analyses of embryonic stem cells during embryoid body-based differentiation*. Proteomics 9(21):4859-70

Gang Z, Schweiger MR, Martinez-Noel G, Zheng L, Smith JA, Harper JW, Howley PM (2009). *Brd4 Regulation of Papillomavirus E2 Protein Stability.* J Virol 83(17):8683-92

Gebert C, Wrenzycki C, Herrmann D, Gröger D, Thiel J, Reinhardt R, Lehrach H, Hajkova P, Lucas-Hahn A, Carnwath JW, Niemann H (2009). *DNA methylation in the IGF2 intragenic DMR is re-established in a sex-specific manner in bovine blastocysts after somatic cloning* Genomics 94(1):63-9

Hache H, Lehrach H, Herwig R (2009). *Reverse Engineering of Gene Regulatory Networks: A Comparative Study*. EURASIP J Bioinform Syst Biol 617281

Hache H, Wierling C, Lehrach H, Herwig R (2009). *GeNGe: Systematic Generation of Gene Regulatory Networks*. Bioinformatics 25(9):1205-1207

Heid IM, Huth C, Loos RJF, Kronenberg F, Adamkova V, Anand SS, Ardlie K, Biebermann H, Bjerregaard P, Boeing H, Bouchard C, Ciullo M, Cooper JA, Corella D, Dina C, Engert JC, Fisher E, Frances F, Froguel P, Hebebrand J, Hegele RA, Hinney A, Hoehe M, Hu F, Hubacek JA, Humphries SE, Hunt SC, Illig T, Kollerits B, Krude H, Kumar J, Lange LA, Langer B, Li S, Lyon

HN, Meyre D, Mohlke KL, Mooser V, Nebel A, Nguyen TT, Paulweber B, Perusse L, Qi L, Rankinen T, Rosskopf D, Schreiber S, Sengupta S, Sorice R, Suk A, Thorleifsson G, Thorsteinsdottir U, Völzke H, Wareham NJ, Waterworth D, Yusuf S, Lange C, Laird N, Hirschhorn JN, Wichmann HE (2009) *Meta-Analysis of the INSIG2 Association with Obesity Including 74,345 Individuals: Does Heterogeneity of Estimates Relate to Study Design?* PLoS Genet 5(10):e1000694

Hirsch-Kauffmann M, Schweiger M, Schweiger MR (2009). *Biologie und molekulare Medizin für Mediziner und Naturwissenschaftler*: Thieme Verlag, 7. Aufl. 2009

Hormeño S, Ibarra B, Chichón FJ, Habermann K, Lange BM, Valpuesta JM, Carrascosa JL, Arias-Gonzalez JR (2009). *Single centrosome manipulation reveals its electric charge and associated dynamic structure*. Biophys J 97(4):1022-30

Hu Y, Lehrach H, Janitz M (2009). *Comparative analysis of an experimental subcellular protein localization assay and in silico prediction methods*. J Mol Histol 40(5-6):343-52

Hufton AL, Mathia S, Braun H, Georgi U, Lehrach H, Vingron M, Poustka AJ, Panopoulou G (2009). *Deeply conserved chordate noncoding sequences preserve genome synteny but do not drive gene duplicate retention*. Genome Res 19(11):2036-51

Hufton AL, Panopoulou G (2009). *Polyploidy and genome restructuring: a variety of outcomes*. Curr Opin Genet Dev 19(6):600-6

Jung M, Mollenkopf HJ, Grimm C, Wagner I, Albrecht M, Waller T, Pilarsky C, Johannsen M, Stephan C, Lehrach H, Nietfeld W, Rudel T, Jung K, Kristiansen G (2009). *MicroRNA profiling of clear cell renal cell cancer identifies a robust signature to define renal malignancy*. J Cell Mol Med 13(9B):3918-28

Kamburov A, Wierling C, Lehrach H, Herwig R (2009). *ConsensusPathDB - a database for integrating human functional interaction networks*. Nucleic Acids Research 37(Database issue):D623-8

Klipp E, Liebermeister W, Wierling C, Kowald A, Lehrach H, Herwig R (2009). *Systems Biology - A Textbook*, Wiley-VCH, Weinheim.

Kühn C, Wierling C, Kühn A, Klipp E, Panopoulou G, Lehrach H, Poustka AJ (2009). *Monte-Carlo analysis of an ODE Model of the Sea Urchin Endomesoderm Network*. BMC Systems Biology 3:83

Kuhn SA, Mueller U, Hanisch UK, Regenbrecht CR, Schoenwald I, Brodhun M, Kosmehl H, Ewald C, Kalff R, Reichart R (2009). *Glioblastoma cells express functional cell membrane receptors activated by daily used medical drugs*. J Cancer Res Clin Oncol 135(12):1729-45.

Martinez-Morales MJ, Rembold M, Greger K, Simpson JC, Brown KE, Quiring R, Pepperkok R, Martin-Bermudo MD, Himmelbauer H, Wittbrodt J (2009). *Ojoplano-mediated basal constriction is essential for optic cup morphogenesis*. Development 136:2165-2175

Meinel T, Herwig R (2009). *SOAP/WSDL-based Web Services for Biomedicine*. Athina Lazakidou (Ed.): Web-Based Applications in Healthcare and Biomedicine – Springer Press ISBN-13: 978-1-4419-1273-2

Mertes F, Martinez-Morales JR, Nolden T, Spörle R, Wittbrodt J, Lehrach H, Himmelbauer H (2009). *Cloning of mouse ojoplano, a reticular cytoplasmic protein expressed during embryonic development*. Gene Expr Patterns 9:562-567

Nebrich G, Herrmann M, Hartl D, Diedrich M, Kreitler T, Wierling C, Klose J, Giavalisco P, Zabel C, Mao L (2009). *PROTEOMER: A workflow-optimized laboratory information management system for 2-D electrophoresis-centered proteomics*. Proteomics; 9(7):1795-1808

Ottinger M, Smith JA, Schweiger MR, Robbins D, Powell ML, You J, Howley PM (2009). *Cell-type specific transcriptional activities among different papillomavirus long control regions and their regulation by E2*. Virology 395(2):161-71.

Parkhomchuk D, Amstislavskiy V, Soldatov A, Ogryzko V (2009) *Use of high throughput sequencing to observe genome dynamics at a single cell level.* Proc Natl Acad Sci USA 106(49):20830-5

Parkhomchuk D, Borodina T, Amstislavskiy V, Banaru M, Hallen L, Krobitsch S, Lehrach H, Soldatov A (2009). *Transcriptome analysis by strand-specific sequencing of complementary DNA.* Nucleic Acids Res 37(18):e123

Postma L, Lehrach H, Ralser M (2009). *Surviving in the cold: yeast mutants with extended hibernating lifespan are oxidant sensitive*. Aging 1:957-960

Ralser M, Lehrach H (2009). *Building a new bridge between metabolism, free radicals and longevity*. Aging 1:836-838

Ralser M, Wamelink MMC, Latkolik S, Jansen EEW, Lehrach H, Jakobs C (2009). *Metabolic reconfiguration precedes transcriptional regulation in the antioxidant response.* Nat Biotech 27(7):604 – 605

Ralser M, Zeidler U, Lehrach H (2009). *Interfering with glycolysis causes Sir2-dependent hyper-recombination of Saccharomyces cerevisiae plasmids*. PLoS One 4:e5376

Rasche A, Herwig R (2009). *ARH: Predicting Splice Variants from Genome-wide Data with Modified Entropy.* Bioinformatics 26(1):84-90

Schulz H, et al., Himmelbauer H, et al., Hatzopolous AK (2009). *The FunGenES database: a genomics resource for mouse embryonic stem cell differentiation.* PLoS ONE 4:e6804

Schweiger MR, Kerick M, Timmermann B, Albrecht MW, Borodina T, Parkhomchuk D, Zatloukal K, Lehrach H (2009). *Genome-Wide Massively Parallel Sequencing of Formaldehyde Fixed-Paraffin Embedded (FFPE) Tumor Tissues for Copy-Number- and Mutation-Analysis.* PLoS One:4(5): e5548

Solomun T, Seitz H, Sturm H (2009). *DNA damage by low-energy electron impact: dependence on Guanine content*. J Phys Chem B 113:11557-9

Trendelenburg G, Ziegler G, Prinz V, Albrecht MW, Harhausen D, Khojasteh U, Nacken W, Endres M, Nietfeld W, Dirnagel U (2009). *Mrp-8 & -14 mediate CNS injury in focal cerebral ischemia*. BBA - Mol Basis of Disease 1792(12):1198-1204

Tsend-Ayush A, Dodge N, Mohr J, Casey A, Himmelbauer H, Kremitzki

CL, Schatzkammer K, Graves T, Warren W, Grützner F (2009). *Higher-order genome organization in platypus and chicken sperm and repositioning of sex chromosomes during mammalian evolution. Chromosoma* 118:53-69

Wamelink M, Jansen E, Struys E, Lehrach H, Jakobs C, Ralser M (2009). *Quantification of Saccharomyces cerevisiae pentose-phosphate pathway intermediates by LC-MS/MS.* Nature protocol Exchange 10.1038/nprot.2009.140

Zerck A, Nordhoff E, Resemann A, Mirgorodskaya E, Suckau D, Reinert K, Lehrach H, Gobom J (2009). *An Iterative Strategy for Precursor Ion Selection for LC-MS/MS Based Shotgun Proteomics*. J Proteome Res 8(7):3239-51

Zheng G, Schweiger MR, Martinez-Noel G, Zheng L, Smith JA, Harper JW, Howley PM *Brd4 regulation of papillomavirus protein E2 stability*. J Virol 83(17):8683-92

Ziegler G, Prinz V, Albrecht MW, Harhausen D, Khojasteh U, Nacken W, Endres M, Dirnagl U, Nietfeld W, Trendelenburg G (2009). *Mrp-8 and -14 mediate CNS injury in focal cerebral ischemia*. Biochim Biophys Acta 1792(12):1198-204

Zuccotti M, Merico V, Redi CA, Bellazzi R, Adjaye J, Garagna S (2009). *Role of Oct-4 during acquisition of developmental competence in mouse oocyte*. Reprod Biomed Online 19 Suppl 3:57-62

Zuccotti M, Merico V, Sacchi L, Bellone M, Brink TC, Stefanelli M, Redi CA, Bellazzi R, Adjaye J, Garagna S (2009). *Oct-4 regulates the expression of Stella and Foxj2 at the Nanog locus: implications for the developmental competence of mouse oocytes.* Hum Reprod 24(9):2225-37

Invited plenary lectures (Hans Lehrach)

ITFoM – IT Future of Medicine. fet11 - The European Future Technologies Conference and Exhibition, Belgium, 31.05.12

The Future of Medicine, Session 7: Diagnostics and Clinical Sequenzing. The International Conference on Genomics in Europe (ICG-Europe), BGI, Denmark, 25.05.12

ITFoM – IT Future of Medicine. Intel Leadership Conference on HPC for Life Sciences, Hungary, 05.05.12

Systems Genomics – the future of Medicine. Robert B Church Lecture in Biotechnology, University of Calgary, Canada, 26.03.12

The Big Six-Spotlight on EU Flagship: Medicine in the 21st Century: IT as a "Magic Bullet"? / Talk: Integrating data on genetic information, Metabolic processes, behavior and Environment to build up the virtual patient. Schweiz. Akademie der Wissenschaften, Swiss National Science Foundation (SNSF), Switzerland, 16.03.12

Genetics, Genomics & Systems Biology. 2nd Heidelberg Forum for Young Scientists, Germany, 24.02.12

Presentation ITFoM. National Research Foundation (NRF), Dubai, United Arab Emirates, 02.02.12

The IT Future of Medicine. Spanish National Cancer Research Centre, CNIO, Spain, 20.01.12

Environment and Genetics: what can we conclude, what should go on? EPH Environment and Public Health in modern society; FocusI: reports from basic research; mechanisms of disease origin & interdisciplinary research, Potsdam, Germany, 07.11.11

Forward Look on personalized Medicine. ESF Workshop, Disease Summit, Netherlands, 18.-20.10.11

Systems Patientomics: The Future of Medicine. Molecular Diagnostics Summit Europe, Germany, 12.10.11

The Future of Medicine. European Health Forum Gastein, Austria, 06.10.11

The IT Futue of Medcine - a flagship initiative to revolutionize our health care system. NGFN Jahrestagung, Symposium III, From Genomics to Application, Germany, 27.09.11

Patient Stratification based on individual full genome sequencing and disease modelling. Mühldorfer Conference Science to Market 2011, Germany, 27.09.11

An in silico reference model of humans as goal for personalised medicine. European Science Foundation (ESF) Technology Workshop, UK, 19.09.11

Der Virtuelle Patient-Systembiologie als Chance für eine individualisierte Medizin. 18. Dresdner Palais-Gespräch, Germany, 02.09.11

From Biobanks, clinical and molecular phenotypes towards systems biology (IT Future of Medicine - initiative). European Congress of Pathology, Session: Biobanks - the power of many, Finnland, 28.08.11

Genetics, Patient Modelling and IT. CARS 2011, Computer Assisted Radiology and surgery, 25th International Congress and Exhibition, Germany, 22.06.11

Other Childhood Cancers - Current Molecular Approaches to address Environmental Risk Factors. International Conference on Non-Ionizing Radiation and Children´s Health, Slovenien, 18.05.11

The Future of Medicine. International Prevention Research Institute (IPRI) "Meeting of National Cancer Institute Directors", France, 16.05.11

Genetics, Genomics & Systems Biology. VIB-Future and history of Genomics, Belgium, 20.04.11

The 1010$ model for all and then? FEBSX-SysBio2011: From Molecules to Function, Austria, 26.02.11

Latest developments in genomics research. Medizinische Universität Graz, Opening Event, Institute of Pathology, Austria, 15.02.11

Krankheiten verstehen und vorbeugen - Neue Ansätze der Systembiologie (Keynote Lecture). 6. Wissenschaft trifft Wirtschaft (WtW), Uni Konstanz, Germany, 17.12.10

Genetics, Genomics & Systems Biology, 6th Genetic Workshop, Germany, 24.11.10

Genetics, Genomics and Systems Biology: the Path to an Individualized Treatment of Cancer Patients (Keynote Lecture). NCSB Netherlands Consortium for Systems Biology Symposium 2010, Netherlands, 21.10.10

Deep sequencing and systems biology: steps on the way to an individualised treatment of cancer patients. MLSB - MachineLearning in SystemsBiology 2010, School of Informatics, University of Edinburgh, UK, Scotland, 15.10.10

Next Generation Sequencing. DGHO Jahrestagung 2010, Deutsche Gesellschaft für Hämatologie und Onkologie, ICC Berlin, Germany, 05.10.10

Genetics, Genomics and Systems Biology: the Path to an Individualized Treatment of Cancer Patients. International Society of Oncology and BioMarkers - ISOBM 2010, Germany, 04.09.10

Genomforschung und die Zukunft der Krebsmedizin. Alpbacher Technologiegespräche 2010, Österreich, 27.08.10

Genetics and individualizing therapy: learning from cancer and gambling. Heart Failure Congress 2010, ICC Berlin, Germany, 30.05.10

Genetics, Genomics & Systems Biology. Symposium in Honour of Professors Sir Kenneth and Noreen Murray, UK, 04.05.10

Systembiologie. Heinrich-Warner-Symposium, Translationale Forschung beim Prostatakarzinom, Hamburg, Germany, 25.02.10

Scientific and technical expertise required for the future. From Biobanks to Expert Centres, BBMRI, Paris, France, 16.12.09

Genetics, Genomics & Systems Biology. Baltic Summer School 2009 - Genetic Basis of Medicine, Kiel, Germany, 10.09.09

Omics in the present and future of human medicine (Keynote Lecture). Workshop: Genomics in Cancer Risk Assessment, San Servolo, Italy, 28.08.09

Technological Advances & Systems Biology. 25th Anniversary Symposium on Genomics and Medicine, Fondation Jean Dausset - CEPH, Paris, France, 23.-24.03.09

Awards/scientific honours

Yuliya Georgieva: *3rd prize, 5th Speed Lecture Award,* 10th BIONNALE, BioTOP Berlin-Brandenburg and vfa bio, 2012

Nana-Maria Grüning: *Nachwuchswissenschaftlerinnen-Preis*, Forschungsverbund Berlin, 2012

Markus Ralser: *European Molecular Biology organisation (EMBO) Young Investigator Award*, 2012

Alexander Kühn: *Winner „Gesundheit 2050 – Deine Ideen für die Zukunft der Gesundheitsforschung*“, Bundesministerium für Bildung und Forschung (BMBF) and WELT-Gruppe, 2011

Markus Ralser: *Wellcome-Beit Prize,* awarded to top four basic biomedical or clinical intermediate fellows of the Wellcome Trust in the UK, 2011

Markus Ralser: *Wellcome Trust Research Career development fellowship*, Wellcome Trust, UK, 2011

Markus Ralser: *ERC starting grant,* European Research Council, 2011

Lars Bertram: *Special-Award of the Hans-und-Ilse-Breuer Foundation for Research in Alzheimer's*, 2010

Stephan Klatt: *Posterprize*, Peptalk – 9th Annual Protein Science Week, Cambride-Healthtech Institute, 2010

Silke Sperling: *W3 Heisenberg professorship for Cardiovascular Genetics,* DFG, 2010

Lars Bertram: *Recipient of the 2009 Independent Investigator Award*, NARSAD, 2009

Thore Brink: *PhD prize*, Berliner Wissenschaftliche Gesellschaft e.V. and TSB Technologie Stiftung Berlin, 2009

Appointments of former members of the department

James Adjaye: *Professorship and Director* of the Institute for Transplantations Diagnostics and Cell Therapies (ITZ), Heinrich-Heine-University of Düsseldorf, Germany, 2012

Andreas Dahl: *Head of Deep Sequencing Group,* Technische Universität Dresden, Germany, 2010

Jörn Glökler: *Group leader*, Alacris Theranostics GmbH, Berlin, Germany, 2011

Zoltán Konthur: Dahlem Centre for Genome Research and Medical Systems Biology (DCGMS) Berlin, 2012

Bodo Lange: *Managing Director*, Alacris Theranostics GmbH, Berlin, Germany, 2012

Eckhard Nordhoff: *Group Leader*, Medizinische Proteom-Center, University Bochum, Germany, 2009

Georgia Panopoulou: *Lecturer,* Institute of Biochemistry and Biology, University of Potsdam, Germany, 2010

Markus Ralser: *Group leader*, Cambridge Systems Biology Centre and Dept. of Biochemistry, Cambridge, UK, 2012

Harald Seitz: *Group leader*, Fraunhofer Institute for Biomedical Engineering, Potsdam, Germany, 2012

Alexey Soldatov: Dahlem Centre for Genome Research and Medical Systems Biology (DCGMS), Berlin, 2012

Silke Sperling: *Professorship,* Department Cardiovascular Genetics at Experimental Clinical Research Center (ECRC), joint institute between Charité - Universitätsmedizin Berlin and Max Delbrück Center for Molecular Medicine. Co-appointed at Faculty of Biology, Chemistry and Pharmacy at Freie Universität Berlin

Habilitation / State doctorate

Silke Sperling: *Discovering the transcriptional networks for cardiac development, function and disease with a systems biology approach*. Venia legendi at Charité, Molecular Biology and Bioinformatics, 2009

PhD theses

Katharina Drews: *Generation and characterization of induced pluripotent stem cells from human amniotic fluid cells*. Freie Universität Berlin, 2012

Nana-Maria Grüning: *Pyruvate kinase triggers a metabolic feedback loop that controls redox metabolism in respiring cells*. Freie Universität Berlin, 2012

Christina M. Lill: *Comprehensive research synopsis and systematic meta-analyses in Parkinson's disease genetics: The PDGene database*. University of Münster, 2012

Guifré Ruiz: *Molecular mechanisms involved in the induction of pluripotency*. Freie Universität Berlin, 2012

Markus Schüler: *Bioinformatics Analysis of Cardiac Transcription Networks*. Freie Universität Berlin, 2012

Tatjana Schütze: *SELEX Technologie-entwicklung im Hochdurchsatz*. Freie Universität Berlin, 2012

Yvonne Welte: *Identification and characterization of cancer stem cells in cutaneous malignant melanoma*. Freie Universität Berlin, 2012

Jenny Schlesinger: *Regulation of Cardiac Gene Expression by Transcriptional and Epigenetik Mechanisms and Identification of a Novel Chromatin Remodeling Factor*. Freie Universität Berlin, 2011

Anne-Kathrin Scholz: *Identification of Aurora A interacting proteins and characterization of the Aurora A Interacting protein PP6*. Freie Universität Berlin, 2011

Anja Berger: *Molecular analysis of the Oryzias latipes (Medaka) transcriptome*. Freie Universität Berlin, 2010

Lukas Chavez: *Multivariate statistical analysis of epigenetic regulation with application to the analysis of human embryonic stem cells*. Freie Universität Berlin, 2010

Xi Cheng: *High-throughput functional analysis of human gene promoters using transfected cell arrays*. Freie Universität Berlin, 2010

Karin Habermann: *A molecular and functional analysis of the Drosophila melanogaster centrosome*. Freie Universität Berlin, 2010

Daniela Köster: *Rolling Circle Amplification auf Biochips*. Freie Universität Berlin, 2010

Cornelia Lange: *Generation and application of genomic tools as important prerequisites for sugar beet genome analyses.* Freie Universität Berlin, 2010

Tobias Nolden: *Genomweite Expressionsanalyse von Sox17 positiven, endodermalen Vorläufer-zellen der Maus*. Freie Universität Berlin, 2010

Axel Rasche: *Information Theoretical Prediction of Alternative Splicing with Application to Type-2 Diabetes mellitus*. Freie Universität Berlin, 2010

Smita Sudheer: *Differential modulation of BMP Signaling by Activin/Nodal and FGF pathways in lineage specification of Human Embryonic Stem Cells*. Freie Universität Berlin, 2010

Gina Ziegler: *Funktionelle Charakterisierung von differentiell exprimierten Genen in einem Maus-modell für die zerebrale Ischämie*. Technische Universität Berlin, 2010

Young-Sook Baek: *Gene Expression Analysis of Differentiating U937 Cells*. Freie Universität Berlin, 2009

Thore Brink: *Transcriptional and signaling analysis as a means of investigating the complexity of aging processes in human and mouse*. Freie Universität Berlin, 2009

Hendrik Hache: *Computational Analysis of Gene Regulatory Networks*. Freie Universität Berlin, 2009

Justyna Jozefczuk: *Analysis of dynamic regulatory events during human embyronic stem cell differentiation into hepatocytes: new insights on downstream targets*. Freie Universität Berlin, 2009

Marc Jung: *A data integration approach to mapping OCT4 gene regulatory networks operative in embryonic stem cells and embryonal carcinoma cells*. Freie Universität Berlin, 2009

Alexander Kühn: *Functional analysis of dickkopf in the seaurchin Strongylocentrotus Purpuratus*. Freie Universität Berlin, 2009

Robert Querfurth: *Functional and phylogenetic analyses of chromosome 21 promoters and hominid-specific transcription factor binding sites*. Freie Universität Berlin, 2009

Theam Soon Lim: *Parameters affecting phage display library design for improved generation of human antibodies*. Freie Universität Berlin, 2009

Martje Tönjes: *Transcription networks in heart development and disease with detailed analysis of TBX20 and DPF3*. Freie Universität Berlin, 2009

Hans-Jörg Warnatz: *Systematic cloning and functional analysis of the proteins encoded on human chromosome 21*. Freie Universität Berlin, 2009

Christoph Wierling: *Theoretical Biology: Modeling and Simulation of Biological Systems and Laboratory Methods*. Freie Universität Berlin, 2009

Student theses

Polixeni Burazi: *A comparative transcriptome analysis of the OCT4 isoforms A and B after over expression in human dermal fibroblast cells*. Master Thesis, Freie Universität Berlin, 2012

Grischa Fuge: *Affintiy maturation of anti-CD30 scFvs applying phage display and light chain shuffling*. Diploma Thesis, Freie Universität Berlin, 2012

Michell Houssong: *Regulatory Mechanisms between Bromodomain Protein 4 and Heme Oxigenase 1*. Master Thesis, Freie Universität Berlin, 2012

Miriam Baradari: *Localization studies for recombinant proteins in Leishmania tarentolae under the influence of Gateway recombination sequence*. Master Thesis, Freie Universität Berlin, 2011

Friedericke Braig: *A novel approach for the generation of full Ig-repertoire-based antibody phage display libraries*. Master Thesis, Universität Potsdam, 2011

Magdalena Ciurkiewicz: *Lokalisierung der embryonalen Expression von Orthologen der bei Autismus auf Chromosom 16p11.2 deletierten bzw. duplizierten Gene im Zebrafisch*. Diploma Thesis, Technical University Darmstadt, 2011

Ole Eigenbrod: *Integrative data analysis of cancer tissues using next generation sequencing*. Bachelor Thesis, Freie Universität Berlin, 2011

Cornelius Fischer: *Genome-wide liver X receptor alpha binding and chromatin acessibility in different macrophage models*. Master Thesis, Freie Universität Berlin, 2011

Sunniva Förster: *Optimizing cDNA phage display for high throughput biomarker discovery*. Diploma Thesis, Humboldt University of Berlin, 2011

Hanna Galincka: *Integriertes Konzept zur Bestimmung von Homologie zwischen verschiedenen Spezies*. Master Thesis, Freie Universität Berlin, 2011

Jonas Ibn-Salem: *Analysis of whole exome sequencing data from a family affected by autism spectrum disorder*. Bachelor Thesis, Freie Universität Berlin, 2011

Katja Lebrecht: *Posttranslationale Modifikation und die CREB-DNA Bindung*. Diploma Thesis, Technische Universität Berlin, 2011

Matthias Lienhard: *Analysis of RNA-seq experiments.* Master Thesis, Freie Universität Berlin, 2011

Matthias Linser: *Biotin-basiertes de-enrichment von low-copy Vektorsequenzen in humanen genomischen Fosmidbanken für die Next Generation Sequenzierung.* Bachelor Thesis, Humboldt University of Berlin, 2011

Barbara Mlody: *Inducing Pluripotency in Somatic Cells by Modulating ROS and p53levels.* Diploma Thesis, Universität Potsdam, 2011

Maja Olszewska: *Spleißanalyse des MeCP2-Gens mit einer intronischen Mutation bei einem Patienten mit Autismus-Spektrum-Störung (ASD).* Diploma Thesis, Julius-Maxilimians-Universität Würzburg, 2011

Bastian Otto: *Vergleichende Untersuchungen zur Simulation biologischer Systeme mit Hilfe Monte Carlo basierter Strategien und analytischer Verfahren*. Diploma Thesis, Freie Universität Berlin, 2011

Annika Pucknat: *Untersuchung und Modellierung der Wirt-Parasit-Interaktion bei Leischmania major*. Bachelor Thesis, Universität Potsdam, 2011

Johannes T. Röhr: *Haplotype reconstruction with partially assembled haplotypes of unrelated individuals*. Master Thesis, Freie Universität Berlin, 2011

Marcel Schilling: *In silico assessment of the effects of single nucleotide polymorphisms on miRNA-mRNA interactions*. Bachelor Thesis, Freie Universität Berlin, 2011

Dennis-Paul Schmoldt: *Laufzeitoptimierung von komplexen mixture models von Phylobayes für phylogenetische Bäume auf der GPU*. Bachelor Thesis, Freie Universität Berlin, 2011

Maria Schulz: *Verifizierung und funktionale Analyse von Mutationen bei Autisten anhand von Exom-sequenzierten-Daten*. Diploma Thesis, Beuth Hochschule für Technik, Berlin, 2011

Kristina Blank: *Die Optimierung der Aptamerselektion gegen kleine Moleküle*. Diploma Thesis, Freie Universität Berlin, 2010

Mirjam Blatter: *Single cell transcriptome analysis using next generation sequencing.* Master Thesis, Humboldt University of Berlin, 2010

Jörn Dietrich: *Statistische genomweite Analyse von MeDIP-Seq Daten zur Identifizierung differentiell methylierter Regionen*. Bachelor Thesis, Technische Hochschule Wildau, 2010

Dejan Gagoski: (2010) *Recombinant expression and biochemical characterization of a thermostable motor protein*. Diploma Thesis, University of Kassel, 2010

Michelle Hussong: *Das Zusammenspiel von Brd4 mit der zellulären transkriptionellen Regulation*. Bachelor Thesis, Fachhochschule Kaiserslautern, 2010

Jan-Martin Josten: *Identification of SNPs associated to autism via the analysis of next-generation sequencing data*. Bachelor Thesis, Freie Universität Berlin, 2010

Svetlana Mareva: *Integration and Visualization of Time Series Expression Data of Gene Regulatory Networks*. Diploma Thesis, Freie Universität Berlin, 2010

Robert Martin: *Funktionelle Analyse von konservierten nicht-codierenden Elementen mittels Zebrafisch als Modelorganismus*. Bachelor Thesis, Beuth Hochschule Berlin, 2010

Matthias Mark Megges: *Cellular reprogramming of human mesenchymal stem cells derived from young and old individuals using viral and non-viral approaches*. Master Thesis, Beuth Hochschule Berlin, 2010

Steve Michel: *Identification of yeast longevity factors via competitive chronological aging experiments*. Diploma Thesis, Universität Mainz, 2010

Kerstin Neubert: *Detection of copy number variants in sequencing data*. Master Thesis, Freie Universität Berlin, 2010

Konstantin Pentchev: *Identification of functional modules in human protein interaction networks*. Master Thesis, Freie Universität Berlin, 2010

Julia Repkow: *Ligase mediated tagged sequencing.* Diploma Thesis, Freie Universität Berlin, 2010

Sophia Schade: *Molecular and functional characterization of the two cancer-relevant proteins MAGED2 and NME7*. Diploma Thesis, Universität Potsdam, 2010

Sandra Schmökel: *Establishment of a purification method for the parathyroid hormone from human tissue extracts and phage display analysis of diagnostically important targets.* Master Thesis, Universität Potsdam, 2010

Stephan Starick: *DNA-based Detection of Proteins.* Diploma Thesis, Universität Kassel, 2010

Thomas Bergmann: *Development of a Multi-Wavelength Fluorescence Reader for Nanoliter PCR Applications*. Diploma Thesis, Technische Universität Berlin, Germany

Felix Bormann: *Development of a Multiplex Assay for robust Determination of epigenetic Changes in genomic Regions, associated with the Formation of Colorectal Cancer*. Diploma Thesis, Freie Universität Berlin, 2009

Helene Braun: *Functional Analysis of conserved non-coding elements in zebrafish.* Diploma Thesis, Technical University Darmstadt, 2009

Cornelia Dorn: *Charakterisierung der Interaktion des epigenetischen Transkriptionsfaktors DPF3 mit dem Baf-Chromatin-Remodelingkomplex*. Diploma Thesis, Humboldt University of Berlin, 2009

Udo Georgi: *Comparison of the co-injection and Tol2 transposon based injection methods for the identification of regulatory elements in zebrafish*. Diploma Thesis, Freie Universität Berlin, 2009

Nicole Hallung: *Die Rolle des Origin Recognition Complex in Zellzyklus, Zentrosomenzyklus und Mikrotubuliorganisation in Drosophila SL2 Zellen*. Diploma Thesis, Freie Universität Berlin, 2009

Oliver Herrmann: *Funktionelle Analyse der konservierten, nicht kodierenden Elemente des sonic hedgehog Gens in Fugu (Takifugu*

rubripes), Stickelback (Gasterosteus aculeatus) und Medaka (Oryzias latipes). Bachelor Thesis, Fachhochschule Zittau/Görlitz, 2009

Michalina Mankowska: *Untersuchung der Homo- und Heterodimerbildung des epigenetischen Transkriptionsfaktors DPF3*. Diploma Thesis, Freie Universität Berlin, 2009

Lina Milbrand: *Untersuchung zur zellulären Funktion von Ataxin-2*. Diploma Thesis, University Greifswald, 2009

Lukas Mittermayr: *Entwicklung einer array-basierten genetischen Karte für das Zuckerrübengenom*. Master Thesis, University Salzburg, Austria, 2009

Svetlana Mollova: *Untersuchungen zur Verwendung der Next Generation Sequenziertechnology für die Analyse von Antikörpergenen bei Gesunden und Autoimmunpatienten.* Diploma Thesis, Technische Universität Braunschweig, 2009

Lisa Scheunemann: *Development of a 'Quadro-TAG' Library Preparation for Gene Expression Profiling on Next-Generation Sequencing Platforms*. Diploma Thesis, Freie Universität Berlin, 2009

Constanze Schlachter: *Untersuchung der ligandenabhängigen strukturellen Dynamik repetitiver DNA-Elemente*. Master Thesis, Technische Hochschule Wildau, 2009

Jana Tänczyk: *Cloning and comparative analysis of the sea anemone (Nematostella vectensis) neural genes with their seaurchin orthologs*. Diploma Thesis, Freie Universität Berlin, 2009

Teaching activities

Since 1998, the department holds the lecture "From functional genomics to systems biology" at Freie Universität Berlin (each winter term, 2hrs/week).

Guest scientists

Huanhuan Cui, Northwest A & F University, Yangling, China, 09.05.11-31.10.12

Karin Mölling, Institut für med. Virologie und Mikrobiologie, Zürich, Switzerland, 01.03.08-31.08.12

Amir Hossini, Städtisches Klinikum Dessau, 21.09.10-30.04.12

Geertrui Tavernier, Ghent University, Belgium, 21.02.11-31.03.12

Gizem Tutku Sengül, Istanbul Technical University, Turkey, 23.01.-17.02.12

Lio Blondel, Université Montpellier, France,10.09.11-17.02.12

Patrick Stumpf, University of Southampton, UK, 02.11.11-31.01.12

Karl Skriner, Charité Berlin, 20.12.10-31.12.11

Floriana Capuano, Universita Di Roma La Sapienza, Italy, 24.10.-04.11.11

Kate Campbell, University of Cambridge, UK, 24.10.-04.11.11

Jong Seto, Max Planck Institute of Colloids and Interfaces, Potsdam, 01.09.10-31.08.11

Asli Pinar Zorba, Istanbul University, Faculty of Science, Turkey, 11.07.-26.08.11

Elizabeth Weingartner (DAAD RISE Stipend), DePauw University, Indiana, USA, 21.05.-30.07.11

Tim Ellermann, Maastricht University, Netherlands, 30.05.-10.07.11

Melanie Busse, imaGenes GmbH, Berlin, 17.02.10-30.04.11

Dmitri Deichel, DZNE Deutsches Zentrum für Neurodegenerative Erkrankungen, Bonn, 14.03.-01.04.11

Desponia Dogka, University of Athen, Greece, 01.04.10-28.02.11

Enrica Bertin, Department of Pedratics, Italy,08.-19.11.10

Zhang Qin, College of Animal Science and Technology, China, 01.10.08-31.10.10

Shiping Wang (DFG fellow), Central South University, Xiangya Medical School, China, 09.08.-30.09.10

Susana Castro-Obregon (Humboldt-Fellow), Institute of Biotechnology, National University of Mexico, 01.09.09-31.08.10

Xu Jing, Uppsala University, Sweden, 16.06.-23.08.10

Damaris Anell Rendon, Institute de Biotechnologia, Mexiko, 04.01.-28.02.10

Eneida Franco Vencio, Universidade Federal De Goias, Faculdade De Odontologia, Goiânia, Goiás, Brasil, 23.03.-31.07.09

Thomas Sander, Universität Köln, 01.07.09-30.06.10

Valko Petrov, Institute of Mechanics and Biomechanics, Bulgarian Academy of Science, Bulgaria, 01.08.-30.11.09

Lucia-Suzanne Postma, VU University Medical Center, Amsterdam, Netherlands, 13.07.-10.10.09

Rachel Cavill, Imperial College London, UK, 09.-24.03.09

Inventions, Patents and Licences

Identifizierung von Ceruloplasmin als potentieller Marker für die Diagnose von aggressivem Prostatakrebs, z.B. in Urin, durch in silico Modellierung auf Basis von experimentellen, Patienten-spezifischen Hochdurch-satzdaten. Hans Lehrach, Michal-Ruth Schweiger, Christoph Wierling, Ralf Herwig, Alexander Kühn, Georg Schäfer, Huaje Bu, Helmut Klocher, Thomas Ringer, Matthias Rainer, Fabio Palato, Günther Bonn. Erfindungsmeldung / patent application in progress, 2012

Prostate Cancer Markers. Hans Lehrach, Michal-Ruth Schweiger, Stefan Börno, Holger Sültmann, Guido Sauter, Thorsten Schlomm. EP 11162979.6; PCT/EP2012/05722, 2011

Methods for nucleic acid sequencing. Jörn Glökler, Hans Lehrach. EP 10168625.1, 2010

Diagnostic Prediction of Rheumatoid Arthritis and Systemic Lupus Erythematosus. Zoltán Konthur, Karl Skriner, Hans Lehrach. WO/2010/072673, 2010

Synthesis of chemical libraries. Jörn Glökler, Hans Lehrach. EP 10168638.8, 2010

Strand-specific cDNA sequencing. Aleksey Soldatov, Tatiana Borodina, Hans Lehrach. EP 09008808.9, 2009

SELEX based on isothermal amplification of nucleic acids in

emulsions. Jörn Glökler, Hans Lehrach, Volker A. Erdmann, Tatjana Schütze. EP 09179046.9, 2009

Biomarker for the prediction of responsiveness to an anti-Tumour Necrosis Factor alpha (TNFα) Treatment. Zoltán Konthur, Karl Skriner, Hans Lehrach. WO/2009/056633, 2009

Spin-offs

Global Action 4 Health Institute Limited, Dublin 2, Ireland, 2012

Dahlem Centre for Genome Research and Medical Systems Biology gGmbH, Berlin, 2010

Organization of scientific events

ITFoM Industry Day, Berlin, 25.04.2012

Triple Event of the year "*From Population Health to Personal Health*". The Genome-based Research and Population Health International Network (GRaPH-Int), the European Science Foundation "Forward Look on Personalised Medicine for the European citizen" (ESF) and the Public Heath Genomics European Network (PHGEN). Back-to-back conferences in association with the European Flagship Pilot for Future Emerging Technologies "IT Future of Medicine" (ITFoM) ,Rome, Italy, 17.-20.04.2012

ITFoM Stakeholder Meeting, Berlin, 13.03.2012

Organisation of two parallel forums: session on "*The future of medicine – developing an infrastructure for personalized medicine*" at the European Health Forum Gastein, Bad Hofgastein, Austria, 05.-08.10.2011

3rd Annual NGFN Meeting, Berlin, 25.-27.11.2010 (co-organizer: H. Lehrach, W. Nietfeld)

2nd Annual NGFN Meeting, Berlin, 26.-28.11.2009 (co-organizer: H. Lehrach, W. Nietfeld)

Genomic variations underlying common neuropsychiatric diseases and cognitive traits in different human populations, EU-project meeting, Berlin, 05.-06.10.2009

EMBO Practical Course: Next generation sequencing: ChIP-seq and RNA-seq., Berlin, 01.-13.02.2009

Seminars and lectures given by external speakers in the department

2012

Sonja Sievers, COMAS - Compound Management and Screening Center, MPI of Molecular Physiology, 28.03.2012. *The Compound Management and Screening Center*

Paul J. McEwan, KAPA Biosystems Inc., 29.02.2012. *Improvements in Next Generation Sequencing Workflow.*

Radek Szklarczyk, Radboud University Nijmegen Medical Centre, 20.01.2012. *Omics approaches for the discovery of mitochondrial disease proteins*

David Juncker, McGill University, Montreal, Canada, 19.01.2012. *Novel microarray and microfluidic technologies for antibody-based biomarker discovery, validation, and point-of-care diagnostics*

2011

Jia Song, Department of Biological Sciences, University of Delaware, USA, 20.12.2011. *Select microRNAs are essential for early development in the sea urchin*

Richard Imrich, Center for Molecular Medicine, Bratislava, Slovakia, 06.12.2011. *Implementation of Personalized Medicine Approaches to Electronic Medical Records Infrastructure*

Charles R. Cantor, SEQUENOM, Inc., San Diego, USA, 07.11.2011. *Quantitative and Sensitive Analysis of Nucleic Acid Biomakers in Cancer and Medical Indications*

Tarso Ledur Kist, Instituto de Biociencias, 12.10.2011. *Recent Advances in DNA Sequencing Techniques*

Bhavna Chanana, University of Cambridge, 07.10.2011. *Role of the Hippo tumour suppressor pathway in the female and male germline of Drosophila melanogaster*

Qiao Pan, Bejing Genome Institute (BGI) Europe, 30.09.2011. *Application of Next Generation Sequencing in human disease and animal Research*

Dennis Friedrich, Broad Institute, Cambridge, USA, 17.06.2011. *High-Throughput Targeted Capture at The Broad Institute*

Arbel Tadmor, California Institute of Technology, Pasadena, USA, 28.03.2011. *New Lessons on the Nature of Single Cells annd Their Viruses*

Peter Boyle, International Prevention Research Institute, Lyon, Frankreich, 27.01.2011. *Global Health and Epidemiology*

Nicolas Lartillot, Canadian Institute for Advanced Research, Canada, 21.01.2011. *Bayesian inference in phylogenomics: reconstructing the evolutionary history of animals*

Mihaela Ulieru, The University of New Brunswick, Canada, 06.01.2011. *Emergent Engineering for the Management of Complex Situations*

2010

Peter J. Hunter, Auckland Bioengineering Institute (ABI), University of Auckland, New Zealand, 09.12.2010. *VPH/Physiome Project and Computational Physiology of the Heart*

Stefan Kubick, Fraunhofer Institut für Biomedizinische Technik, Potsdam-Golm, 08.12.2010. *Cell-Free Synthesis of Posttranslationally Modified Membrane Proteins*

Markus Stoffel, ETH Zürich, 29.11.2010. *MicroRNAs: Regulators of metabolism and therapeutic opportunities*

Ralph Gräf, Universität Potsdam, 23.11.2010. *Functional and microscopic analysis of centrosome and nuclear envelope-associated proteins in Dictyostelium amoebae*

Nancy Hynes, FMI - Friedrich Miescher Institute Basel, 20.11.2010. *Targeting receptor tyrosine kinases and downstream effector proteins in breast cancer*

Sabine Klauck, DKFZ Heidelberg, 18.11.2010. *Challenge Autism - The Genetic Perspective*

Duncan Odom, University of Cambridge, Cancer Research UK, 07.10.2010. *Evolution of Transcriptional Regulation in Mammals*

Kalim Mir, The Wellcome Trust Centre for Human Genetics, Oxford, UK, 21.09.2010. *Nanoscience, Single Molecules and Genomics*

Tomas Marques-Bonet, University of Washington, Seattle, Inst. De Biologia Evolutiva, Barcelona, 06.09.2010. *The evolution of segmental duplications*

Axel Bethke, Boyce-Thompson-Institute, Cornell University, NY, USA, 26.08.2010. *Nuclear Hormone Receptor Regulation of microRNAs controls Developmental Progression*

Ion Mandoiu, University of Connecticut, USA, 19.07.2010. *Estimation of alternative splicing isoform frequencies from RNA-Seq data*

Friedrich Rippmann, Merck Serono, Darmstadt, 06.07.2010. *Industrial Bio- and Chemoinformatics*

David E. Housman, Massachusetts Institute of Technology, Cambridge, USA, 17.06.2010. *Huntington's Disease and Myotonic Dystrophy: RNA targets and Biological Assays*

Christian Ottmann, Chemical Genomics Centre der Max-Planck-Gesellschaft, Dortmund, 14.06.2010. *Small molecule stabilizers of 14-3-3 protein-protein interactions*

Anatoly Ruvinsky, University of New England, 08.06.2010. *Studies of exon-intron structure of eukaryotic genes*

Kai Lao, Applied Biosystems, 28.05.2010. *Tracing the Derivation of Embryonic Stem Cells from the Inner Cell Mass by Single-Cell RNA-Seq Analysis*

Thomas Mathias Gress, Universitätsklinikum Gießen und Marburg GmbH, Marburg, 27.05.2010. *Translational and functional genome analysis in pancreatic cancer*

Evgeny I. Rogaev, Head of 3 laboratories: Brudnick Neuropsychiatric Research Institute, University of Massachusetts Medical School, in the National Research Center of Mental Health, Russian Academy of Medical Sciences and in the Institute of General Genetics of Russian Academy of Sciences, 28.04.2010. *Molecular mechanisms of Alzheimer's disease: from genes to treatment*

Alistair Pagnamenta, Wellcome Trust Centre for Humann Genetics, University of Oxford, 15.04.2010. *Autism genetics: Strategies for use of Next Generation Sequencing*

Kelly Frazer, Moores UCSD Cancer Center, California, 22.03.2010. *Population sequencing and functional annotation of targeted genomic intervals*

Roman Thomas, MPI für Neurologische Forschung, Köln, 08.03.2010. *Functional Cancer Genomics*

Robert Penchovsky, Biotech company consulter, 22.02.10. *Design and Screening of Riboswitches and Their Applications as Biosensors to Exogenous Control of Gene Expression, Molecular Computing, and Drug Development*

David Gurwitz, Tel-Aviv University, 18.02.2010. *Antidepressant Response: Lessons from Human Lymphoblastoid Cell Lines*

Andreas Papassotiropoulos, University of Basel, 16.02.2010. *Genetics of human memory: understanding complexity*

Francois Lapraz, CNRS, University Pierre et Marie Curie, 01.02.2010. *TGF-beta signaling and axial specification in the sea urchin P. lividus*

2009

Kevin Nagel, European Bioinformatics Institute (EBI), Cambridge, UK, 11.12.2009. *Automatic functional annotation of predicted active sites: combining PDB and literature mining*

Alexander Chetverin, Institute of Protein Research, Russian Academy of Science, Moscow, 09.11.2009. *Molecular Colonies*

Joseph Lehar, Boston University, USA, 01.10.2009. *Systems Biology from Chemical Combinations*

Sergey Yazynin, Biosaving GmbH, 22.09.2009. *Barnase, Ribonuklease aus Bacillus amyloliqufaciens und ihr Inhibitor, Barstar, in der Proteinforschung und Biomedizin*

Alex Keshet, ENI, Georg August University, Göttingen, 18.09.2009. *A novel protein splicing process in E.coli is regulated by the phase of growth*

Pascal Braun, CCSB, Dana-Farber Cancer Institute, Boston, 15.09.2009. *High quality interactome mapping to study human disease*

Kathrin Plath, UCLA School of Medicine, USA, 13.07.2009. *Deciphering the mechanism of transcription factor-induced reprogramming and defining the iPS cell state*

Arne Pfeufer, Helmholtz Zentrum München, 12.06.2009. *Complex Genetics of EKG Signatures: GWAS guided identification of ion channel genes - and many surprises*

Gerd Illing, Jürgen Helfmann, Laser- und Medizin-Technologie GmbH, Berlin, 30.04.2009. *Biomedical Optics at LMTB: Molecules, Cells, Tissue - What would you like to see?*

Igor Grigoriev, US Department of Energy Joint Genome Institute, USA, 20.04.2009. *Genome Annotation and Analysis for Bioenergy*

Gottfried Schatz, Universität Basel, 02.03.2009. *What it Takes to Succeed in Science*

Katja Nowick, University of Illinois, USA, 24.02.09. *Rapid Sequence and Expression Pattern Evolution in Primate "Krüppel"-type Zinc Finger Transcription Factors*

Jörg Rademann, Leibniz Institute of Molecular Pharmacology (FMP), 11.02.2009. *Chemical tools for the validation of protein targets generated by dynamic ligation screening*

Stephan Sigrist, Institute for Genetics, FU Berlin, 29.01.2009. O*rganisation and function of the pre- und post-synaptic zones and the use of STED microscopy for supra-molecular structures*

Elke Binder, INSERM Institute Francois Magendie, Bordeaux, France, 29.01.2009. *Behavioural phenotyping, drugs and genetics*

Thomas Thum, Universität Würzburg, 19.01.2009. *Development of microRNA-based therapeutic strategies in cardiovascular medicine*

Department of Human Molecular Genetics

(Established: 07/1995)

Head

Prof. Dr. Hans-Hilger Ropers
Phone: +49 (0)30 8413-1240
Fax: +49 (0)30 8413-1383
Email: ropers@molgen.mpg.de

Secretary

Gabriele Eder, M.A.
Phone: +49 (0)30 8413-1241
Fax: +49 (0)30 8413-1383
Email: eder@molgen.mpg.de

Senior Scientific Staff

Prof. Dr. Thomas F. Wienker (since 05/11)
Dr. Luciana Musante* (since 01/11)
Dr. Hao Hu* (since 05/09)
PD Dr. Reinhard Ullmann (since 02/04)
Dr. Vera Kalscheuer (since 07/95)
Dr. Diego Walther (02/03-07/12)
Prof. Dr. Tim Hucho (09/05-01/12)
Dr. Andreas Tzschach (03/05-06/11)
Prof. Dr. Andreas Kuss (06/05-10/10)

Heads of associated groups

apl. Prof. Dr. Harry Scherthan (since 01/04)
Dr. Wei Chen (01/09-12/11)
Prof. Dr. Susann Schweiger (04/05 – 09/11)

Scientists and postdocs

Dr. Sybille Krauss* (05/05-2010)
Dr. Masoud Garshasbi (05/08-10/11)
Dr. Nils Paulmann (04/08-08/11)
Dr. Lars Riff Jensen (04/02-03/11)

PhD students

Daniel Mehnert (since 03/11)
Melanie Hambrock (since 12/09)
Silke Stahlberg (since 07/08)
Agnes Zecha (since 07/08)
Lucia Püttmann (since 06/08)
Roxana Karimi-Nejad* (at times at MPIMG until 07/12)
Hyung-Goo Kim* (at times at MPIMG until 07/12)
Jakob Vowinckel (07/06-10/11)
Lia Abbasi Moheb (10/05-09/11)
Julia Hoffer (09/10-08/11)
Robert Weißmann (09/09-04/11)
Sahar Esmaeeli-Nieh* (11/06-10/10)
Na Li (09/07-09/10)
Hui Kang (09/07-09/10)
Eva Kickstein* (03/06-03/10)
Nils Rademacher (02/05-03/10)
Joanna Walczak Stulpa (10/05-02/10)
Stella-Amrei Kunde (01/05-06/09)

Technicians

Melanie Bienek (since 06/09)
Susanne Freier (since 10/05)
Bettina Lipkowitz (since 10/05)
Sabine Otto (since 10/05)
Astrid Grimme (since 02/04, part time)
Ute Fischer (since 08/95, part time)
Marianne Schlicht (10/05-06/12)
Corinna Jensen (01/07-05/11)
Nadine Nowak (01/08-02/10)

* externally funded

Introduction

In search of clinically relevant genetic risk factors for common disorders, genome-wide association studies (GWAS) have been performed since more than 15 years, with meagre results. Now even the most prominent advocate of this research direction has thrown the sponge[1] and world-wide, rare disorders have come into focus of genome research again[2]. We were among the first to point out the inherent difficulties of GWAS in complex diseases, including their genetic heterogeneity, and to stress the importance of studying single gene disorders as a better alternative. Consequently, the research of our department has revolved around monogenic disorders since its inception[3].

During the past decade, we have increasingly focused on intellectual disability (ID) and related disorders. ID, also called mental retardation or early-onset cognitive impairment, is the biggest unsolved problem of clinical genetics and a far heavier socio-economic burden than, e.g., cancer[4]. Of the several thousand gene defects that may give rise to ID, only a few hundred have been identified so far, since many forms of ID are clinically indistinguishable and because in Western Societies, families tend to be small. We circumvented this problem by forging international cooperations, e.g. with a potent group in Iran. This enabled us to study familial forms of ID in a systematic manner, and put us in an excellent position for scaling up this research when high-throughput sequencing techniques became available, as outlined below.

This research would not have been possible without substantial financial support from the Max Planck Society and additional funding from the German Federal Ministry of Education and Research. Both came to an end in October 2011, when the Head of the Department reached his official retirement age, but was re-installed as Acting Director for a period of three years. Since then, a deconstruction plan is in place, which entails a progressive reduction of the structural budget until the end of October 2014, when the department will be closed. However, additional financial support has been obtained through an EU grant (Genetic and Epigenetic Networks in COgnitive DISorders [GENCODYS], FP7 reference no. 241995, 05/2010 – 04/2015) which will partially compensate the diminishing structural resources.

In accordance with the deconstruction plan, the groups of Tim Hucho (Signal Transduction in Pain and Mental Retardation) and Diego Walther (Monoamine Signalling and Disease, Mouse Lab.) were discontinued when the contracts of their leaders expired. Andreas Tzschach, our former clinical geneticist, left for the University of Tuebingen in 2011, and his position was filled by Thomas Wienker, a human geneticist and retired professor of biostatistics from the University of Bonn. Andreas Kuss was appointed as a professor (W2) by the University of Greifswald, and since 2011, Luciana Musante, a former post-doctoral fellow in the group of Vera Kalscheuer, is now in charge of our research into recessive ID. While we are still maintaining close ties with Wei Chen, now at the Max Delbrück Center in Berlin, his part-time appointment at the MPIMG was discontinued in 2011, and his former post-doctoral fellow Hao Hu took over his task as the

1 Zuk O, Hechter E, Sunyaev S R & Lander E S. (2012). *The mystery of missing heritability: Genetic interactions create phantom heritability*. Proc Natl Acad Sci U S A 109: 1193-1198

2 Ropers HH (2010). *Single gene disorders come into focus--again*. Dialogues Clin Neurosci 12: 95-102

3 for more details, see Research Report 2009

4 Smith K (2011). *Trillion-dollar brain drain*. Nature 478:15

bioinformatician of our department. Finally, the part-time appointment of Susann Schweiger (University of Dundee and future Head of Human Genetics at the University of Mainz) has also been terminated in 2011.
In view of the reduced size of the department, which is also due to the fact that we can no longer hire PhD students, most former groups have lost their critical mass. At the same time, the research of the remaining scientists converged and their collaboration intensified. Therefore, their scientific achievements are no longer presented separately, except for Tim Hucho, recently appointed as W2 professor at the University of Cologne, and Reinhard Ullmann's group with its gradually diverging orientation and own DFG support.

Scientific methods and findings

Next generation sequencing revolutionizes the identification of ID genes

(Wei Chen, Hao Hu, Vera Kalscheuer, Reinhard Ullmann, Andreas Tzschach, Andreas Kuss)
To elucidate the genetic defects underlying ID and related disorders, we have employed a wide spectrum of approaches, including breakpoint mapping in patients with disease-associated balanced chromosome rearrangements (DBCRs), screening for disease-associated copy number variants by array CGH, and linkage mapping in families and mutation screening of candidate genes, as outlined previously[3]. While array CGH-based screening for pathogenic copy number variants, the study of DBCRs and linkage mapping in patients and families remain useful strategies for the elucidation of genetic disorders, as illustrated by several recent publications of our group, our decision paid off to invest early into genomic enrichment techniques, high throughput sequencing and the storage, handling and interpretation of next generation sequencing data.

Development of mutation detection pipelines

(Hao Hu; together with Stefan Haas and co-workers, Dept. Computational Molecular Biology)
Various members of the Department of Computational Molecular Biology (Head: Martin Vingron) contributed to this effort by developing a bioinformatic mutation detection pipeline, which was first used to look for *de novo* mutations on the X-chromosome in patient-parent trios with a suspected X-linked dominant disorders (Chen W. et al, unpublished observations). Later on, this pipeline was instrumental in our comprehensive collaborative effort to identify the molecular defects underlying X-linked mental retardation (see below).
Independently, Hao Hu developed another algorithm for identifying pathogenic changes in whole genome and whole exome sequences. This algorithm has been employed successfully to look for mutations in consanguineous families with autosomal recessive ID and has been described in several publications[5]; a more comprehensive description is in preparation. Since 2010, these methods have become the mainstay of our research into the genetic causes of ID and related disorders.

5 See e.g. Najmabadi H, Hu H, Garshasbi M, […]Zecha A, Mohseni M, Puttmann L, Vahid L N, Jensen C, Moheb L A, Bienek M, Larti F, Mueller I, Weissmann R, Darvish H, Wrogemann K, Hadavi V, Lipkowitz B, Esmaeeli-Nieh S, […] Hoffer J, Falah M, Musante L, Kalscheuer V, Ullmann R, Kuss AW, Tzschach A, Kahrizi K, Ropers HH (2011). *Deep sequencing reveals 50 novel genes for recessive cognitive disorders*. Nature 478: 57-63

X-linked ID genes: draining the pond

(Vera Kalscheuer, Hao Hu, Chen Wei, Thomas Wienker; in cooperation with Stefan Haas, Tomasz Zemojtel, Martin Vingron, Dept. Computational Molecular Biology)

Employing a custom-made hybrid capture kit to enrich 7591 X-chromosomal exons, or 875 genes, we have performed targeted exon sequencing in 248 European families with X-linked forms of ID. In the vast majority of these families, X-linkage was virtually certain because of affected males in separate sibships that were connected through healthy females. Apparently deleterious DNA variants were identified in 13 genes that had not been implicated in ID before, and their identity as novel genes for X-linked ID (XLID) was corroborated in various ways. This study raises the number of known XLID genes to 110.

Using the same parameters to distinguish pathogenic from clinically irrelevant sequence variants, we have also reanalyzed the results of a previous study encompassing 208 Caucasian families, which had been screened for mutations by large-scale Sanger sequencing[6]. Under the (plausible) assumption that the cohorts analyzed by the two studies are part of the same population, this enabled us to estimate the total number of XLID genes as 123 (95% confidence limits: 91-155). This estimate is lower than expected and cannot be reconciled easily with our finding that mutations in the known 110 genes account for at most 71% of the XLID families. One possible explanation for this discrepancy is that most of the missing mutations may reside in non-coding, e.g. intronic sequences which were not analyzed in these studies (Kalscheuer et al, unpublished).

A plethora of novel genes for autosomal recessive forms of ID (ARID)

(Andreas Kuss, Andreas Tzschach, Hao Hu, Masoud Garshasbi, Luciana Musante, Thomas Wienker)

Following up on a previous, pioneering study to identify novel ARID loci and to assess the genetic heterogeneity of this condition[3], we have performed array-based SNP typing and linkage mapping in 300 consanguineous Iranian and German families. In 27 of these families, a single homozygous interval was observed, and at least 14 novel ARID loci could be identified[7]. Starting in 2006, when only three ARID genes had been reported,[3] mutation screening of all genes located single linkage intervals has revealed numerous novel genes for syndromic or non-syndromic forms of ID (see Table 1).

In 2010, by combining targeted exon enrichment and next generation sequencing, we extended these studies to 136 consanguineous ARID families with more than one linkage interval. In 78 of these, a single plausible causative mutation could be identified, involving 22 known and 50 novel candidate genes. This study, published in 2011,[5] quadruplicated the number of (candidate) genes for non-syndromic forms of ARID, which were found to be more common than syndromic forms. For the vast majority of these genes, pathogenic mutations were only seen in a single family. This corroborates our previous observations and suggests that none of these gene defects can account for more than a few percent of all forms of

6 Tarpey PS, […] Ropers HH, et al. (2009). A systematic, large-scale resequencing screen of X-chromosome coding exons in mental retardation. Nat Genet 41: 535-543

7 Kuss AW, Garshasbi M, Kahrizi K, Tzschach A, Behjati F, Darvish H, Abbasi-Moheb L, Puettmann L, Zecha A, Weissmann R, Hu H, […] Ullmann R, […] Ropers HH, Najmabadi H (2011). *Autosomal recessive mental retardation: homozygosity mapping identifies 27 single linkage intervals, at least 14 novel loci and several mutation hotspots*. Hum Genet 129:141-148

ID in Iran – but it does not exclude founder mutations for ARID in one or several of the 7 or 8 Iranian sub-populations.

Gene	Location	Function	Ethnicity	Reference
GRIK2	**6q16.3**	Involved in the transmission of light signals from the retina to the hypothalamus, Involved in the maturation of microcircuits and network formation in brain areas	Iranian	Motazacker MM et al., Am J Hum Genet 2007; 81: 792–798
TUSC3	**8p22**	Putative Mg2+ transporter, required for cellular Mg2+ uptake. Indispensable for normal vertebrate embryonic development.	Iranian, French	Garshasbi M et al., Am J Hum Genet 2008; 82: 1158–1164
VLDLR	**9p24**	Part of the reelin signaling pathway, which is involved in neuroblast migration in the cerebral cortex and cerebellum	Iranian, Canadian, Turkish	Abbasi Moheb L et al., Euro J Hum Genet 2008; 16: 270–273
TRAPPC9	**8q24.3**	Enhancer of the cytokine-induced NF-(kappa)B signaling pathway, having an essential function in post mitotic neurons as opposed to neural progenitors	Iranian, Pakistani, Tunisian, Israeli	Mir A et al., Am J Hum Genet 2009; 85: 909-915
SRD5A3	**4q12**	Polyprenol reductase with a crucial role in N-linked protein glycosylation that is required for converting polyprenol to dolichol.	Iranian, Emirati, Turkish, Polish	Kahrizi K et al., Euro J Hum Genet 2011; 19:115–117
ZC3H14	**14q31.3**	May contribute to control of gene expression in human cells through binding poly(A) RNA	Iranian	Pak CH et al., PNAS 2011; 108:12390-95
ST3GAL3	**1p34.1**	Transfers sialic acid to terminal positions on the carbohydrate groups of glycoproteins and glycolipids that are key determinants for a variety of cellular recognition processes	Iranian	Hu H et al, Am J Hum Genet 2011; 89:407-414
NSUN2	**5p15.31**	RNA methyltransferase that methylates tRNAs, and possibly RNA polymerase III transcripts. May act downstream of Myc to regulate epidermal cell growth and proliferation	Iranian	Abbasi Moheb L et al., Am J Hum Genet 2012; 90: 847-55
ZNF526	**19q13.2**	Involved in transcriptional regulation, role in regulation of translation	Iranian	Abbasi Moheb L et al., ESHG meeting 2011

Table 1: Novel molecular defects underlying syndromic and non-syndromic ARID.

Currently, we reanalyze families with more than one plausible mutation, e.g. by studying knock-down fly models for the relevant gene defects (see below), as well as families where no single homozygous mutation has been found. In some of these families, the ID may be due to compound heterozygosity, i.e., it may be unrelated to the parental consanguinity, or the causative mutations may hide in introns or other non-coding sequences that have not been investigated so far. To unveil these missing mutations, whole genome sequencing has been performed in 11 of these families, and analysis of the results is ongoing.

In parallel, we have embarked on a second, even larger study encompassing almost 150 consanguineous families from Iran and Germany, including all remaining large families collected by our Iranian partner in the course of this long-standing collaboration. In most of these families, SNP typing revealed multiple homozygous linkage intervals, which renders targeted exon enrichment with custom-made arrays costly and tedious. Therefore, and in order to detect compound heterozygosity, we have turned to whole exome sequencing instead,

which is being performed in collaboration with Wei Chen, now as Principal Investigator in charge of high-throughput sequencing at the Max Delbrück Center for Molecular Medicine in Berlin.

Functional studies

(Vera Kalscheuer, Susann Schweiger, Luciana Musante, Andreas Kuss; in cooperation with Thomasz Zemojtel, Dept. Computational Molecular Biology)

Given the large number of our newly discovered ID genes and the many different pathomechanisms that can give rise to ID, we realized early on that we would not have the manpower to study the function of all gene products on our own. Therefore, except for PQBP1 and CDKL5[8], the functions of which have been studied by V. Kalscheuer and co-workers and Susann Schweiger's equally remarkable research into the function of the MID1 gene,[3] we have collaborated with groups with complementary (e.g. biochemical) expertise. Such collaborations led to a series of articles in high-impact journals. For several of the many novel ID genes identified by our group since 2009, knockout mouse models are being generated by Yann Hérault and co-workers (IGBMC, Illkirch, F) within the framework of the EU-FP7 project GENCODYS. Prior to this project, KO mouse models had been developed by D. Walther and co-workers for the microcephalin 1 (MCPH1) gene and for FTSJ1, an XLID gene identified by our group in 2004. Behavioural testing of the C57BL-backcrossed FTSJ1 KO mice is being performed at the German Mouse Clinic in Munich, in collaboration with Lars Jensen and Andreas Kuss.

For ~70% of the recently reported 50 novel ARID candidate genes,[5] Krystyna Keleman and Cornelia Oppitz at the IMP in Vienna have identified orthologs in Drosophila, and in 14 of the corresponding RNAi knockdown fly models, repeated behavioural testing revealed long term or short term memory defects (unpublished observations). Other behavioural tests are being conducted in these fly models by Annette Schenck and co-workers (University of Nijmegen, NL),

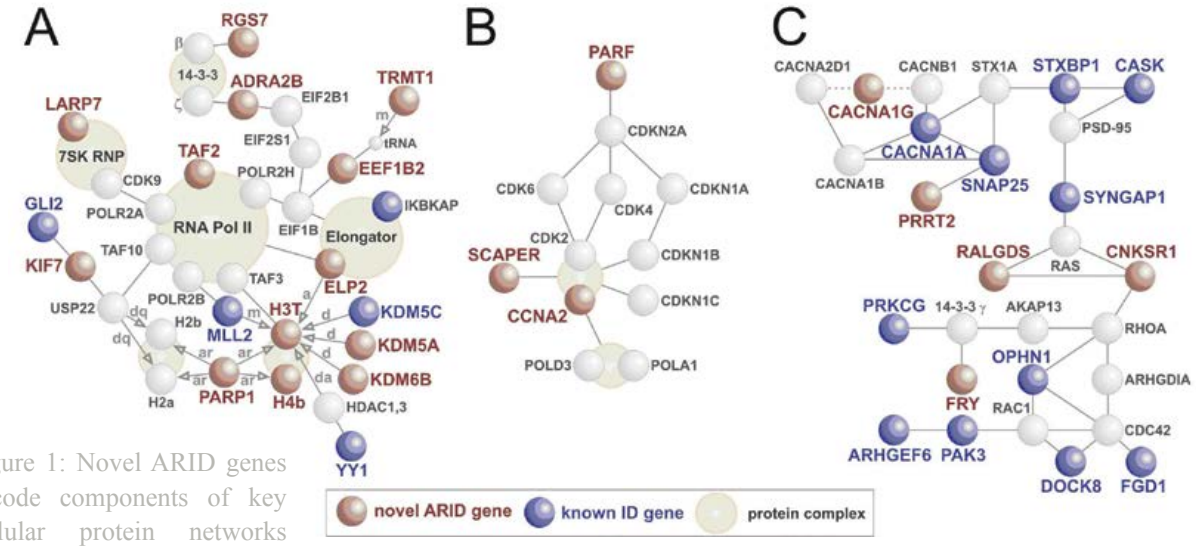


Figure 1: Novel ARID genes encode components of key cellular protein networks (from Najmabadi et al., 2011).

8 Ricciardi S, Ungaro F, Hambrock M, Rademacher N, […] Kalscheuer VM, Broccoli V (2012). *CDKL5 ensures excitatory synapse stability by reinforcing NGL-1-PSD95 interaction in the postsynaptic compartment and is impaired in patient iPS cell-derived neurons.* Nat Cell Biol, in press

also under the umbrella of GENCODYS. So far, the identity of more than half of the 50 novel ARID genes has been confirmed, either by animal studies or by finding additional families with deleterious mutations of the respective genes.
As an even faster approach to study the function of novel ID genes, we had previously initiated a collaboration with Matthias Mann (MPIMB, Martinsried) and Anthony Hyman (MPICBG, Dresden) who analyze the interaction of all human proteins in HeLa cells[3]. This project is being pursued by A. Kuss as part of our collaboration agreement. Before long, however, empirical data of this kind may no longer be required, as illustrated by the fact that many of the recently described novel ARID genes or their gene products interacted directly with known ID genes, thereby forming functional clusters or pathways, as predicted three years ago (Figure 1).

Diagnostic aspects

(Thomas Wienker, Wei Chen, Hao Hu; Thomasz Zemojtel, Dept. Computational Molecular Biology)
High-throughput sequencing techniques have not only revolutionized the elucidation of single gene disorders, but also provided the basis for comprehensive diagnostic tests to rule out mutations in all known disease genes. In collaboration with Stephen Kingsmore (Children's Mercy Hospital, Kansas City, USA), the Pediatric Department of the Berlin University Hospital Charité and Wei Chen at the MDC Berlin, we have developed a clinical entry test for children with severe ID and/or unexplained developmental delay. This test, baptized 'MPIMG1', encompasses numerous severe childhood disorders, all published ID genes as well as the many novel ones identified by our group. In principle, this test can also be employed for non-invasive preconception carrier detection to rule out elevated parental risks for children with severe recessive disorders. This application renders it particularly useful for consanguineous parents and for countries from the so-called 'Consanguinity Belt' that extends from the Maghreb to India.

Outlook

Until our department will be officially closed in late fall 2014 and the EU-FP7 project will expire in April 2015, most of our remaining resources will be used to successfully conclude three ongoing, closely related projects.
First, the long-standing collaboration with our Iranian partner will reach its natural end when all ARID families collected since 2004 have been analyzed and funding of both groups by the afore-mentioned EU project will be discontinued. We expect that until then, our research into autosomal and X-linked forms of ID will remain internationally competitive.
Secondly, provided that residual administrative hurdles can be overcome, we will implement our novel MPIMG1 test at the Charité – Universitätsmedizin Berlin, the Children's Mercy Hospital in Kansas City and elsewhere as a first-line diagnostic tool for all known genetic defects that cannot be readily diagnosed by clinical examination alone. Thereafter, we intend to offer this diagnostic and carrier test to improve genetic health care in selected countries with frequent parental consanguinity, intellectual disability and developmental delay.
Finally, in collaboration with a Dutch group, we are about to shed light on the old question whether ID genes also have a role as determinants of the normal IQ

distribution, as postulated by Lehrke[9] 40 years ago. Targeted exon enrichment and next generation sequencing in ~ 170 selected ID genes will be performed to test the hypothesis that there is an inverse relationship between the IQ of the proband and the number of subtle mutations in these genes. While the analysis of these data may turn out to be quite difficult and time-consuming, we expect that this investigation will be completed well within the available time-frame, even if it should turn out that follow-up studies will be required to answer this question in full.

9 Lehrke R. (1972). *Theory of X-linkage of major intellectual traits*. Am J Ment Defic 76: 611-619

Molecular Cytogenetics Group

(since 02/2004)

Head
PD Dr. Reinhard Ullmann
Phone: +49 (0)30 8413-1251
Fax: +49 (0)30 8413-1383
Email: ullmann@molgen.mpg.de

Bioinformatician
Katrin Tebel* (since 04/09)

PhD students
Grit Ebert (since 10/09)
Anne Steininger (since 07/08)
Vivien Boldt (since 02/08, currently on maternity leave)

Technicians
Melanie Heßler* (since 03/12)
Ines Müller* (03/04 – 12/11)
Alisho Ahmed* (06/08-03/11)
Hannelore Madle (11/07-08/09)

Student
Jana Grune (03/11-09/11)

Scientific overview

Structural variants of the genome, in particular those that are associated with loss or gain of DNA, represent the greatest portion of genomic variability in humans. They are important forces of evolution, but at the same time, they pose a considerable threat to individual health. Structural variants are the underlying cause of many congenital disorders and their accumulation in somatic cells is a hallmark of tumor development.

Predicting the biological consequences of structural variants is challenging in many cases. Importantly, a specific genetic change that appears to have no effect at all in one individual may cause severe health problems in another one. These individual differences in response to structural variants result from the complex interplay of diverse regulatory mechanisms at the genetic and epigenetic level. Our current research is dedicated to improving our understanding of these influences.

We have identified and published numerous rare genomic variants as the underlying genetic cause of various diseases, including autism, schizophrenia, amyotrophic lateral sclerosis, ADHD and congenital malformations of the heart, thyroid and brain. Furthermore, we have investigated the patterns of somatic DNA copy number changes in tumors of the breast and the hematopoietic system. We have developed strategies using the gene content of rare structural variants identified in a cohort of patients suffering of brain malformations to learn more about the regulatory networks involved in brain development. Currently, we are applying the same bioinformatics approach to somatic mutations identified in T cell-lymphomas. Proceeding on the assumption that the tumor-specific patterns of chromosomal aberrations reflect the selective pressure favoring the observed

* externally funded

combination of mutations, we have employed network analysis to identify recurrently mutated genes that have a high likelihood to act in concert to promote T-cell lymphoma progression. The relevance of the predicted interactions will be tested *in vitro* by combinatorial deregulation of gene dosage either by knock down or over-expression in a project funded by the Wilhelm Sander foundation. For a more comprehensive view on the causes of structural variants and their variable phenotypic consequences, we have started to integrate information on higher order chromatin organization and the three dimensional structure of the nucleus into our analysis. In this context, we are developing appropriate software tools (funded by the DFG) and are in the process of establishing a technique called Hi-C, which is a sequencing-based method that has been developed to elucidate the three-dimensional organization of the nucleus with unprecedented resolution.

Co-operations within the department

My group has been part of the department´s initiative to decipher the genetic causes of autosomal recessive intellectual disability. We have established the techniques and performed the experiments that have allowed the region specific sequencing of all the selected genomic intervals of interest. Furthermore we are continuing our fruitful cooperation with the group of Vera Kalscheuer to fine-map the breakpoints of balanced chromosomal translocations

Selected publications

Kariminejad R, Lind-Thomsen A, Tümer Z, Erdogan F, Ropers HH, Tommerup N*, Ullmann R*, Møller RS (2011). *High frequency of rare copy number variants affecting functionally related genes in patients with structural brain malformations.* Hum Mutat 32(12):1427-35 * shared corresponding authors

Steininger A, Möbs M, Ullmann R, Köchert K, Kreher S, Lamprecht B, Anagnostopoulos I, Hummel M, Richter J, Beyer M, Janz M, Klemke CD, Stein H, Dörken B, Sterry W, Schrock E, Mathas S, Assaf C (2011). *Genomic loss of the putative tumor suppressor gene E2A in human lymphoma.* J Exp Med 208(8):1585-93.

Lesch KP, Selch S, Renner TJ, Jacob C, Nguyen TT, Hahn T, Romanos M, Walitza S, Shoichet S, Dempfle A, Heine M, Boreatti-Hümmer A, Romanos J, Gross-Lesch S, Zerlaut H, Wultsch T, Heinzel S, Fassnacht M, Fallgatter A, Allolio B, Schäfer H, Warnke A, Reif A, Ropers HH, Ullmann R (2011). *Genome-wide copy number variation analysis in attention-deficit/hyperactivity disorder: association with neuropeptide Y gene dosage in an extended pedigree.* Mol Psychiatry 16(5):491-503

Former Group of the Department

Signal Transduction in Pain and Mental Retardation

(10/2005-01/2012)

Head
Prof. Dr. Tim Hucho
(10/2005-01/2012)
now at University Hospital
of Cologne Experimental
Anaesthesiology and Pain Research
Phone: +49 (0)221 478-97760
Fax: +49 (0)221 478-6093
Email: tim.hucho@uk-koeln.de

Scientists and postdocs
Dr. Anibal Garza Carbajal[#] (08/11-02/12)
Dr. Jörg Isensee*[#] (04/09-01/12)
Dr. Chandan Goswami (01/06-07/09)

PhD students
Juliane Schreier*[#] (01/09-01/12)
René Buschow[#] (12/08-01/12)
Christine Andres (07/08-10/11)
Julia Kuhn* (10/06-06/11)

Students
Lukas Pluska (01/12-02/12)
Johanna Meyer (11/11-12/11)
Josephine Pott (09/11-10/11)
Carsten Wenzel (01/11-06/11)
Stefan Schüler (03/11-05/11)
Sheila Hoffmann (09/10-10/10)

Technician
Vanessa Suckow (since 09/05)

Scientific overview

Research concept

The knowledge about signaling networks in areas like cancer, immune regulation, and neuronal memory is enormous. In contrast, knowledge about signaling networks underlying pain sensitization in the peripheral "pain" neuron, the nociceptive neuron, is lagging behind greatly. Reasons are that nociceptive neurons are small in numbers, very difficult to transfect, and extremely heterogenous in functionality. A cell line model system is not available. The focus of my group is to develop signaling based mechanism inspiring novel mechanism based therapeutic approaches.

Scientific methods and findings

In 2005-2009 we made great progress by investigating the regulation of sensitization signaling involving the protein kinase C epsilon (PKCe). Among others and as reported in reports covering the years 2005-2009, we established

* externally funded

[#] Since the beginning of 2012, Jörg Isensee, Anibal Garza Carbajal, René Buschow and Juliane Schreier are employed by the University Hospital of Cologne.

interconnectivity between the most classical signaling pathways cAMP and PKC, sparked the research of estrogen signaling in nociceptive neurons, and identified the interaction of microtubules and TRPV1.
During the last three years, many of the projects mentioned in the report of 2006-2009 bore fruit. Beyond the molecular and cellular details referred to in the last report, three projects with great potential for the clinical therapy may be highlighted:
1) We showed *via* biochemical, cell biological and behavioural experiments that the interaction of the cytoskeleton with TRPV1 is regulated by the pain-sensitization signaling component PKCe. TRPV1 regulated microtubuli by a mechanism independent of the channel opening of TRPV1. Abolishing the interaction of TRPV1 and microtubules in the animal by downregulation of TRPV1 switched the microtubule stabilizing comonent and potent anti-cancer drug, Taxol, from a strong pain sensitizer into a sensitization blocking substance. Many aspects of this interaction might be relevant also to a second ion channel involved in pain sensitization, TRPV4.
2) We found that PKCe is a convergence point or "nociceptive module", a concept introduced by us to the pain community in our review Hucho & Levine, Neuron 2007. At least bradykinin, isoproterenol and estrogen signaling converged onto PKCe. This opened the way to novel questions such as, how are multiple sensitization stimuli signaling to the same signaling component integrated in nociceptive neurons. We found that a first sensitization signaling results in the activation of a "switch". As a result, a subsequent second stimulation of the same pathway then does not result in further sensitization but in desensitization. We identified components of the cellular "switch". Thereby we now can "reprogram" the signal computation of nociceptive neurons also by non-sensitizing pharmacological means. These results were established in cell biological, electrophysiological, behavioral, and first human experiments and could provide a fully novel concept of therapy.
3) We established that also components central to the wound healing process such as EGF can act on nociceptive neurons directly. In contrast to other growth factors, we found EGF not to induce but to block sensitization.
Much of this work was based on the investigation of PKCe signaling. One difficulty of monitoring endogenous PKCe activity is that it is extremely laborious. The investigator has to visually judge large numbers of neurons if a partial and often nearly invisible translocation of the cytoplasmic PKCe to the plasma membrane has occurred. While robust in its results, the time to invest was enormous.
Over the last 4 years we established as the first group in pain research an unbiased, fully quantitative, software based approach. This was only possible, as the Max Planck Society has generously contributed to a "High Content Screening (HCS) Microscope", bought by the institute. HCS microscopes are fully automated and photograph major parts of multi-well cell cultures. Neurons are then identified by image analysis and quantified for immunfluorescence intensity. As large numbers of cells can be investigated resulting in fully graded quantification of single neurons, for the first time a highly sensitive true population analysis of this highly heterogeneous system could be performed. Thereby, we are now able to investigate expression and kinetics (!) of endogenous (!) signaling components in highly heterogeneous primary nociceptive neurons.
To establish the HCS-microscopy approach was a great challenge. There were no techniques available to quantitatively analyze heterogeneous populations defined by sparse data sets if compared to flow cytometry. About 500-few thousand neurons are quantified per condition. Accordingly, in cooperation with Jan Hasenauer of the group of Frank Allgöwer at the University Stuttgart, we developed a Kernel

Density Estimation Substraction approach. The advantage of this approach is, that no (!) parameter estimates are necessary allowing a fully unbiased quantification. The disadvantage is, that only a lower bounds of e.g. "responding cells" could be established. We now just succeeded with the development of an alternative mixture modeling approache based in part on the methods established by the Vingron department for the determination of the full size of subgroups being based only on few assumptions such as population size and behaviour of control subgroups.
Based on our technical advances we currently finish studies showing that the TRPV1 responses in primary neurons are of very high heterogeneity. The investigation of large numbers of single cell life cell traces in combination with expression studies of the very same cells was a great challenge so far not approached by others in the field. The result of this study is, that intracellular signaling only in part controls the heterogeneity. Also, we developed the first primary neuron based screening approach being now able to screen dozens of compounds per day for their sensitization potential of nociceptive neurons. This has led to the identification of novel endogenous compounds like FGF and PDGF. Like EGF, which we described as a novel endogenous inhibitory system, FGF and PDGF again are components of high importance for the wound healing process. This shifts the focus of the pain field from the investigation of inflammatory mediators to mediators of wound healing. Also, surprisingly, a first screen on cytokines revealed, that the ones tested so far, do not sensitize nociceptive neurons directly contradicting many expectations.
The fully quantitative nature and the scale of our studies was the base of the first consortium using systems biological methods for the study of pain, which I initiated and coordinated. In the highly competitive recent e:BIO call by the BMBF we succeeded again and will get under my coordination funding for additional three years. We are hopefull thereby to establish, that systems biological analysis of signaling can be powerful tool to elucidate novel therapeutic concepts. The relevance of our approaches and our results of the pain switch eluded to above resulted was granted by an invitation for a full talk at the international meeting of „Systems Biology in Human Disease" in Havard in 2011.
And last but not least, I am proud to have accomplished to combine on a molecular level the studies of pain with the ongoing studies of mental retardation in the department. We found TRPV1 to be an important component regulating synapse formation. Also, we were crucial in establishing and coordinating the collaboration of the department with the Group of Rita Gerardy-Schahn. Thereby the genetic finding, that ST3GAL3 mutations impair higher brain functions, was complemented with strong molecular and cellular mechanistic evidence.

Selected publications

Goswami C, Kuhn J, Dina OA, Fernández-Ballester G, Levine JD, Ferrer-Montiel A, Hucho T (2011). *Estrogen destabilizes microtubules through an ion-conductivity-independent TRPV1 pathway*. J Neurochem 117(6):995-1008

Hu H, Eggers K, Chen W, Garshasbi M, Motazacker MM, Wrogemann K, Kahrizi K, Tzschach A, Hosseini M, Bahman I, Hucho T, Mühlenhoff M, Gerardy-Schahn R, Najmabadi H, Ropers HH, Kuss AW (2011). *ST3GAL3 mutations impair the development of higher cognitive functions*. Am J Hum Genet 89(3):407-14

Andres C, Meyer S, Dina OA, Levine JD, Hucho T (2010). *Quantitative automated microscopy (QuAM) elucidates growth factor specific signalling in pain sensitization.* Mol Pain 6:98

General information about the whole Department

Complete list of publications (2009-2012)

2012

Abbasi-Moheb L, Mertel S, Gonsior M, Nouri-Vahid L, Kahrizi K, Cirak S, Wieczorek D, Motazacker MM, Esmaeeli-Nieh S, Cremer K, Weissmann R, Tzschach A, Garshasbi M, Abedini SS, Najmabadi H, Ropers HH, Sigrist SJ, Kuss AW (2012). *Mutations in NSUN2 cause autosomal-recessive intellectual disability*. Am J Hum Genet 90: 847

Andres C, Hasenauer J, Allgower F, Hucho T (2012). *Threshold-free population analysis identifies larger DRG neurons to respond stronger to NGF stimulation*. PLoS One 7: e34257

Arnold H, Wienker TF, Hoffmann MM, Scheuerbrandt G, Kemp K, Bugert P (2012). *High levels of brain-type creatine kinase activity in human platelets and leukocytes: a genetic anomaly with autosomal dominant inheritance*. Blood Cells Mol Dis 48: 62

Azimzadeh O, Scherthan H, Yentrapalli R, Barjaktarovic Z, Ueffing M, Conrad M, Neff F, Calzada-Wack J, Aubele M, Buske C, Atkinson M J, Hauck SM, Tapio S (2012). *Label-free protein profiling of formalin-fixed paraffin-embedded (FFPE) heart tissue reveals immediate mitochondrial impairment after ionising radiation*. J Proteomics 75: 2384

Braunholz D, Hullings M, Gil-Rodriguez MC, Fincher CT, Mallozzi MB, Loy E, Albrecht M, Kaur M, Limon J, Rampuria A, Clark D, Kline A, Dalski A, Eckhold J, Tzschach A, Hennekam R, Gillessen-Kaesbach G, Wierzba J, Krantz ID, Deardorff MA, Kaiser FJ (2012). *Isolated NIBPL missense mutations that cause Cornelia de Lange syndrome alter MAU2 interaction*. Eur J Hum Genet 20: 271

Emde AK, Schulz MH, Weese D, Sun R, Vingron M, Kalscheuer VM, Haas SA, Reinert K (2012). *Detecting genomic indel variants with exact breakpoints in single- and paired-end sequencing data using SplazerS*. Bioinformatics 28: 619

Forcheron F, Agay D, Scherthan H, Riccobono D, Herodin F, Meineke V, Drouet M (2012). *Autologous adipocyte derived stem cells favour healing in a minipig model of cutaneous radiation syndrome*. PLoS One 7: e31694

Hockner M, Spreiz A, Fruhmesser A, Tzschach A, Dufke A, Rittinger O, Kalscheuer V, Singer S, Erdel M, Fauth C, Grossmann V, Utermann G, Zschocke J, Kotzot D (2012). *Parental Origin of de novo Cytogenetically Balanced Reciprocal Non-Robertsonian Translocations*. Cytogenet Genome Res 136: 242

Hoffer J, Fryssira H, Konstantinidou A, Ropers HH, Tzschach A (2012). *Novel WDR35 mutations in patients with cranioectodermal dysplasia (Sensenbrenner syndrome)*. Clin Genet [Epub ahead of print Apr. 9]

Huppke P, Brendel C, Kalscheuer V, Korenke GC, Marquardt I, Freisinger P, Christodoulou J, Hillebrand M, Pitelet G, Wilson C, Gruber-Sedlmayr U, Ullmann R, Haas S, Elpeleg O,

Nurnberg G, Nurnberg P, Dad S, Moller LB, Kaler SG, Gartner J (2012). *Mutations in SLC33A1 cause a lethal autosomal-recessive disorder with congenital cataracts, hearing loss, and low serum copper and ceruloplasmin*. Am J Hum Genet 90: 61

Kelly S, Pak C, Garshasbi M, Kuss A, Corbett AH, Moberg K (2012). *New kid on the ID block: Neural functions of the Nab2/ZC3H14 class of Cys3His tandem zinc-finger polyadenosine RNA binding proteins*. RNA Biol 9: [Epub ahead of print May 1]

Kim HG, Kim HT, Leach NT, Lan F, Ullmann R, Silahtaroglu A, Kurth I, Nowka A, Seong IS, Shen Y, Talkowski ME, Ruderfer D, Lee JH, Glotzbach C, Ha K, Kjaergaard S, Levin AV, Romeike BF, Kleefstra T, Bartsch O, Elsea SH, Jabs EW, Macdonald ME, Harris DJ, Quade BJ, Ropers HH, Shaffer LG, Kutsche K, Layman LC, Tommerup N, Kalscheuer VM, Shi Y, Morton CC, Kim CH, Gusella JF (2012). *Translocations Disrupting PHF21A in the Potocki-Shaffer-Syndrome Region Are Associated with Intellectual Disability and Craniofacial Anomalies*. Am J Hum Genet [Epub ahead of print Jul 3]

Mutschler J, Abbruzzese E, von der Goltz C, Dinter C, Mobascher A, Thiele H, Diaz-Lacava A, Dahmen N, Gallinat J, Majic T, Petrovsky N, Kornhuber J, Thuerauf N, Grunder G, Brinkmeyer J, Wienker T, Wagner M, Winterer G, Kiefer F (2012). *Genetic Variation in the Neuropeptide Y Gene Promoter Is Associated with Increased Risk of Tobacco Smoking*. Eur Addict Res 18:246

Quednow BB, Brinkmeyer J, Mobascher A, Nothnagel M, Musso F, Grunder G, Savary N, Petrovsky N, Frommann I, Lennertz L, Spreckelmeyer KN, Wienker TF, Dahmen N, Thuerauf N, Clepce M, Kiefer F, Majic T, Mossner R, Maier W, Gallinat J, Diaz-Lacava A, Toliat MR, Thiele H, Nurnberg P, Wagner M, Winterer G (2012). *Schizophrenia risk polymorphisms in the TCF4 gene interact with smoking in the modulation of auditory sensory gating*. Proc Natl Acad Sci US 109:6271

Ricciardi S, Ungaro F, Hambrock M, Rademacher N, Stefanelli G, Brambilla D, Sessa A, Magagnotti C, Bachi A, Giarda E, Verpelli C, Kilstrup-Nielsen C, Sala C, Kalscheuer VM, Broccoli V (2012). *CDKL5 ensures excitatory synapse stability by reinforcing NGL-1-PSD95 interaction in the postsynaptic compartment and is impaired in patient iPS cell-derived neurons*. Nat Cell Biol, in press

Ropers HH (2012). *On the future of genetic risk assessment*. J Community Genet [Epub ahead of print Apr 1]

Singer H, Walier M, Nusgen N, Meesters C, Schreiner F, Woelfle J, Fimmers R, Wienker T, Kalscheuer VM, Becker T, Schwaab R, Oldenburg J, El-Maarri O (2012). *Methylation of L1Hs promoters is lower on the inactive X, has a tendency of being higher on autosomes in smaller genomes and shows inter-individual variability at some loci*. Hum Mol Genet 21:219

Wagner M, Schulze-Rauschenbach S, Petrovsky N, Brinkmeyer J, Von Der Goltz C, Grunder G, Spreckelmeyer KN, Wienker T, Diaz-Lacava A, Mobascher A, Dahmen N, Clepce M, Thuerauf N, Kiefer F, De Millas JW, Gallinat J, Winterer G (2012). *Neurocognitive impairments in non-deprived smokers-results from a population-based multi-center study on smoking-related behavior*. Addict Biol [Epub ahead of print Feb 17]

Zanni G, Calì T, Kalscheuer VM, Ottolini D, Barresi S, Lebrun N, Montecchi-Palazzi L, Hu H, Chelly J, Bertini E, Brini M, Carafoli E (2012). *A mutation of plasma membrane Ca^{2+} ATPase isoform 3 in a family with X-linked congenital cerebellar ataxia impairs Ca^{2+} homeostasis*. Proc Natl Acad Sci USA, in press

2011

Adelfalk C, Ahmed EA, Scherthan H (2011). *Reproductive phenotypes of mouse models illuminate human infertility*. J Reprod Medicine & Endocrinol 8:376

Azimzadeh O, Scherthan H, Sarioglu H, Barjaktarovic Z, Conrad M, Vogt A, Calzada-Wack J, Neff F, Aubele M, Buske C, Atkinson MJ, Tapio S (2011). *Rapid proteomic remodeling of cardiac tissue caused by total body ionizing radiation*. Proteomics 11:3299

Bordlein A, Scherthan H, Nelkenbrecher C, Molter T, Bosl MR, Dippold C, Birke K, Kinkley S, Staege H, Will H, Winterpacht A (2011). *SPOC1 (PHF13) is required for spermatogonial stem cell differentiation and sustained spermatogenesis*. J Cell Sci 124:3137

Brinkmeyer J, Mobascher A, Musso F, Schmitz M, Wagner M, Frommann I, Grunder G, Spreckelmeyer K N, Wienker T, Diaz-Lacava A, Holler D, Dahmen N, Thuerauf N, Clepce M, Kiefer F, De Millas W, Gallinat J, Winterer G (2011). *P50 sensory gating and smoking in the general population*. Addict Biol 16:485

Buonincontri R, Bache I, Silahtaroglu A, Elbro C, Nielsen AM, Ullmann R, Arkesteijn G, Tommerup N (2011). *A cohort of balanced reciprocal translocations associated with dyslexia: identification of two putative candidate genes at DYX1*. Behav Genet 41:125

Cichon S, Muhleisen TW, Degenhardt FA, Mattheisen M, Miro X, Strohmaier J, Steffens M, Meesters C, Herms S, Weingarten M, Priebe L, Haenisch B, Alexander M, Vollmer J, Breuer R, Schmal C, Tessmann P, Moebus S, Wichmann HE, Schreiber S, Muller-Myhsok B, Lucae S, Jamain S, Leboyer M, Bellivier F, Etain B, Henry C, Kahn JP, Heath S, Hamshere M, O'donovan MC, Owen MJ, Craddock N, Schwarz M, Vedder H, Kammerer-Ciernioch J, Reif A, Sasse J, Bauer M, Hautzinger M, Wright A, Mitchell PB, Schofield PR, Montgomery GW, Medland SE, Gordon SD, Martin NG, Gustafsson O, Andreassen O, Djurovic S, Sigurdsson E, Steinberg S, Stefansson H, Stefansson K, Kapur-Pojskic L, Oruc L, Rivas F, Mayoral F, Chuchalin A, Babadjanova G, Tiganov AS, Pantelejeva G, Abramova LI, Grigoroiu-Serbanescu M, Diaconu CC, Czerski PM, Hauser J, Zimmer A, Lathrop M, Schulze TG, Wienker TF, Schumacher J, Maier W, Propping P, Rietschel M, Nothen MM (2011). *Genome-wide association study identifies genetic variation in neurocan as a susceptibility factor for bipolar disorder*. Am J Hum Genet 88:372

Cingoz S, Bache I, Bjerglund L, Ropers HH, Tommerup N, Jensen H, Brondum-Nielsen K, Tumer Z (2011). *Interstitial deletion of 14q24.3-q32.2 in a male patient with plagiocephaly, BPES features, developmental delay, and congenital heart defects*. Am J Med Genet A 155A:203

De Assis NA, Nowak S, Ludwig KU, Reutter H, Vollmer J, Heilmann S, Kluck N, Lauster C, Braumann B, Reich RH, Hemprich A, Knapp M, Wienker TF, Kramer FJ, Hoffmann P, Nothen MM, Mangold E (2011). *SUMO1 as a candidate gene for non-syndromic cleft lip with or without cleft palate: no evidence for the involvement of common or rare*

variants in Central European patients. Int J Pediatr Otorhinolaryngol 75:49

Delbeck M, Golz S, Vonk R, Janssen W, Hucho T, Isensee J, Schafer S, Otto C (2011). *Impaired left-ventricular cardiac function in male GPR30-deficient mice*. Mol Med Report 4:37

Diaz-Lacava A, Walier M, Penacino G, Wienker TF, Baur MP (2011). *Spatial assessment of Argentinean genetic admixture with geographical information systems*. Forensic Sci Int Genet 5:297

Diaz-Lacava A, Walier M, Willuweit S, Wienker TF, Fimmers R, Baur MP, Roewer L (2011). *Geostatistical inference of main Y-STR-haplotype groups in Europe*. Forensic Sci Int Genet 5:91

El-Maarri O, Walier M, Behne F, Van Uum J, Singer H, Diaz-Lacava A, Nusgen N, Niemann B, Watzka M, Reinsberg J, Van Der Ven H, Wienker T, Stoffel-Wagner B, Schwaab R, Oldenburg J (2011). *Methylation at global LINE-1 repeats in human blood are affected by gender but not by age or natural hormone cycles*. PLoS One 6:e16252

Fullston T, Gabb B, Callen D, Ullmann R, Woollatt E, Bain S, Ropers HH, Cooper M, Chandler D, Carter K, Jablensky A, Kalaydjieva L, Gecz J (2011). *Inherited balanced translocation t(9;17)(q33.2;q25.3) concomitant with a 16p13.1 duplication in a patient with schizophrenia*. Am J Med Genet B Neuropsychiatr Genet 156:204

Garshasbi M, Kahrizi K, Hosseini M, Nouri Vahid L, Falah M, Hemmati S, Hu H, Tzschach A, Ropers HH, Najmabadi H, Kuss AW (2011). *A novel nonsense mutation in TUSC3 is responsible for non-syndromic autosomal recessive mental retardation in a consanguineous Iranian family*. Am J Med Genet A 155A:1976

Gilling M, Lind-Thomsen A, Mang Y, Bak M, Moller M, Ullmann R, Kristoffersson U, Kalscheuer VM, Henriksen KF, Bugge M, Tumer Z, Tommerup N (2011). *Biparental inheritance of chromosomal abnormalities in male twins with non-syndromic mental retardation*. Eur J Med Genet 54:e383

Goswami C, Kuhn J, Dina OA, Fernandez-Ballester G, Levine JD, Ferrer-Montiel A, Hucho T (2011). *Estrogen destabilizes microtubules through an ion-conductivity-independent TRPV1 pathway*. J Neurochem 117:995

Gregor A, Albrecht B, Bader I, Bijlsma EK, Ekici AB, Engels H, Hackmann K, Horn D, Hoyer J, Klapecki J, Kohlhase J, Maystadt I, Nagl S, Prott E, Tinschert S, Ullmann R, Wohlleber E, Woods G, Reis A, Rauch A, Zweier C (2011). *Expanding the clinical spectrum associated with defects in CNTNAP2 and NRXN1*. BMC Med Genet 12:106

Hu H, Eggers K, Chen W, Garshasbi M, Motazacker MM, Wrogemann K, Kahrizi K, Tzschach A, Hosseini M, Bahman I, Hucho T, Muhlenhoff M, Gerardy-Schahn R, Najmabadi H, Ropers HH, Kuss AW (2011). *ST3GAL3 mutations impair the development of higher cognitive functions*. Am J Hum Genet 89:407

Jakobsen LP, Bugge M, Ullmann R, Schjerling CK, Borup R, Hansen L, Eiberg H, Tommerup N (2011). *500K SNP array analyses in blood and saliva showed no differences in a pair of monozygotic twins discordant for cleft lip*. Am J Med Genet A 155A:652

Jensen LR, Chen W, Moser B, Lipkowitz B, Schroeder C, Musante L, Tzschach A, Kalscheuer VM, Meloni I, Raynaud M, Van Esch H, Chelly J, De Brouwer AP, Hackett A, Van Der Haar S, Henn W, Gecz J, Riess O, Bonin M, Reinhardt R, Ropers HH, Kuss AW (2011). *Hybridisation-based resequencing of 17 X-linked intellectual disability genes in 135 patients reveals novel mutations in ATRX, SLC6A8 and PQBP1*. Eur J Hum Genet 19:717

Kahrizi K, Hu CH, Garshasbi M, Abedini SS, Ghadami S, Kariminejad R, Ullmann R, Chen W, Ropers HH, Kuss AW, Najmabadi H, Tzschach A (2011). *Next generation sequencing in a family with autosomal recessive Kahrizi syndrome (OMIM 612713) reveals a homozygous frameshift mutation in SRD5A3*. Eur J Hum Genet 19:115

Kariminejad R, Lind-Thomsen A, Tumer Z, Erdogan F, Ropers HH, Tommerup N, Ullmann R, Moller RS (2011). *High frequency of rare copy number variants affecting functionally related genes in patients with structural brain malformations*. Hum Mutat 32:1427

Koopmann A, Dinter C, Grosshans M, Von Der Goltz C, Hentschel R, Dahmen N, Gallinat J, Wagner M, Grunder G, Thurauf N, Wienker T, Brinkmeyer J, Mobascher A, Spreckelmeyer K N, Clepce M, De Millas W, Wiedemann K, Winterer G, Kiefer F (2011). *Psychological and hormonal features of smokers at risk to gain weight after smoking cessation--results of a multicenter study*. Horm Behav 60:58

Kunde SA, Musante L, Grimme A, Fischer U, Muller E, Wanker EE, Kalscheuer VM. (2011). *The X-chromosome-linked intellectual disability protein PQBP1 is a component of neuronal RNA granules and regulates the appearance of stress granules*. Hum Mol Genet 20:4916

Kuss AW, Garshasbi M, Kahrizi K, Tzschach A, Behjati F, Darvish H, Abbasi-Moheb L, Puettmann L, Zecha A, Weissmann R, Hu H, Mohseni M, Abedini SS, Rajab A, Hertzberg C, Wieczorek D, Ullmann R, Ghasemi-Firouzabadi S, Banihashemi S, Arzhangi S, Hadavi V, Bahrami-Monajemi G, Kasiri M, Falah M, Nikuei P, Dehghan A, Sobhani M, Jamali P, Ropers HH, Najmabadi H (2011). *Autosomal recessive mental retardation: homozygosity mapping identifies 27 single linkage intervals, at least 14 novel loci and several mutation hotspots*. Hum Genet 129:141

Lesch KP, Selch S, Renner TJ, Jacob C, Nguyen TT, Hahn T, Romanos M, Walitza S, Shoichet S, Dempfle A, Heine M, Boreatti-Hummer A, Romanos J, Gross-Lesch S, Zerlaut H, Wultsch T, Heinzel S, Fassnacht M, Fallgatter A, Allolio B, Schafer H, Warnke A, Reif A, Ropers HH, Ullmann R (2011). *Genome-wide copy number variation analysis in attention-deficit/hyperactivity disorder: association with neuropeptide Y gene dosage in an extended pedigree*. Mol Psychiatry 16:491

Lindenberg A, Brinkmeyer J, Dahmen N, Gallinat J, De Millas W, Mobascher A, Wagner M, Schulze-Rauschenbach S, Grunder G, Spreckelmeyer KN, Clepce M, Thurauf N, Von Der Goltz C, Kiefer F, Steffens M, Holler D, Diaz-Lacava A, Wienker T, Winterer G (2011). *The German multi-centre study on smoking-related behavior-description of a population-based case-control study*. Addict Biol 16:638

Mitter D, Ullmann R, Muradyan A, Klein-Hitpass L, Kanber D, Ounap K, Kaulisch M, Lohmann D (2011).

Genotype-phenotype correlations in patients with retinoblastoma and interstitial 13q deletions. Eur J Hum Genet 19:947

Muradyan A, Gilbertz K, Stabentheiner S, Klause S, Madle H, Meineke V, Ullmann R, Scherthan H. (2011). *Acute high-dose X-radiation-induced genomic changes in A549 cells*. Radiat Res 175:700

Najmabadi H, Hu H, Garshasbi M, Zemojtel T, Abedini S S, Chen W, Hosseini M, Behjati F, Haas S, Jamali P, Zecha A, Mohseni M, Puttmann L, Vahid LN, Jensen C, Moheb LA, Bienek M, Larti F, Mueller I, Weissmann R, Darvish H, Wrogemann K, Hadavi V, Lipkowitz B, Esmaeeli-Nieh S, Wieczorek D, Kariminejad R, Firouzabadi SG, Cohen M, Fattahi Z, Rost I, Mojahedi F, Hertzberg C, Dehghan A, Rajab A, Banavandi M J, Hoffer J, Falah M, Musante L, Kalscheuer V, Ullmann R, Kuss AW, Tzschach A, Kahrizi K, Ropers HH (2011). *Deep sequencing reveals 50 novel genes for recessive cognitive disorders*. Nature 478:57

Pabst S, Golebiewski M, Herms S, Karpushova A, Diaz-Lacava A, Walier M, Zimmer S, Cichon S, Wienker TF, Nothen MM, Nickenig G, Meyer R, Skowasch D, Grohe C (2011). *Caspase recruitment domain 15 gene haplotypes in sarcoidosis*. Tissue Antigens 77:333

Pagan C, Botros HG, Poirier K, Dumaine A, Jamain S, Moreno S, De Brouwer A, Van Esch H, Delorme R, Launay JM, Tzschach A, Kalscheuer V, Lacombe D, Briault S, Laumonnier F, Raynaud M, Van Bon BW, Willemsen MH, Leboyer M, Chelly J, Bourgeron T (2011). *Mutation screening of ASMT, the last enzyme of the melatonin pathway, in a large sample of patients with intellectual disability*. BMC Med Genet 12:17

Pak C, Garshasbi M, Kahrizi K, Gross C, Apponi LH, Noto JJ, Kelly SM, Leung SW, Tzschach A, Behjati F, Abedini SS, Mohseni M, Jensen LR, Hu H, Huang B, Stahley SN, Liu G, Williams KR, Burdick S, Feng Y, Sanyal S, Bassell GJ, Ropers HH, Najmabadi H, Corbett AH, Moberg KH, Kuss AW (2011). *Mutation of the conserved polyadenosine RNA binding protein, ZC3H14/dNab2, impairs neural function in Drosophila and humans*. Proc Natl Acad Sci U S A 108:12390

Rademacher N, Hambrock M, Fischer U, Moser B, Ceulemans B, Lieb W, Boor R, Stefanova I, Gillessen-Kaesbach G, Runge C, Korenke GC, Spranger S, Laccone F, Tzschach A, Kalscheuer VM. (2011). *Identification of a novel CDKL5 exon and pathogenic mutations in patients with severe mental retardation, early-onset seizures and Rett-like features*. Neurogenetics 12:165

Rafiq MA, Kuss AW, Puettmann L, Noor A, Ramiah A, Ali G, Hu H, Kerio N A, Xiang Y, Garshasbi M, Khan MA, Ishak GE, Weksberg R, Ullmann R, Tzschach A, Kahrizi K, Mahmood K, Naeem F, Ayub M, Moremen KW, Vincent JB, Ropers HH, Ansar M, Najmabadi H (2011). *Mutations in the alpha 1,2-mannosidase gene, MAN1B1, cause autosomal-recessive intellectual disability*. Am J Hum Genet 89:176

Rietschel M, Mattheisen M, Degenhardt F, Kahn RS, Linszen DH, Os JV, Wiersma D, Bruggeman R, Cahn W, De Haan L, Krabbendam L, Myin-Germeys I, Muhleisen TW, Kirsch P, Esslinger C, Herms S, Demontis D, Steffens M, Strohmaier J, Haenisch B, Breuer R, Czerski P M, Giegling I, Strengman E, Schmael C, Mors O, Mortensen P B, Hougaard D M, Orntoft T, Kapelski P, Priebe L, Basmanav F B, Forstner AJ, Hoffmann

P, Meier S, Nikitopoulos J, Moebus S, Alexander M, Mossner R, Wichmann HE, Schreiber S, Rivandeneira F, Hofman A, Uitterlinden AG, Wienker TF, Schumacher J, Hauser J, Maier W, Cantor RM, Erk S, Schulze TG, Stefansson H, Steinberg S, Gustafsson O, Sigurdsson E, Petursson H, Kong A, Stefansson K, Pietilainen OP, Tuulio-Henriksson A, Paunio T, Lonnqvist J, Suvisaari J, Peltonen L, Ruggeri M, Tosato S, Walshe M, Murray R, Collier DA, Clair DS, Hansen T, Ingason A, Jakobsen KD, Duong L, Werge T, Melle I, Andreassen OA, Djurovic S, Bitter I, Rethelyi JM, Abramova L, Kaleda V, Golimbet V, Jonsson EG, Terenius L, Agartz I, Winkel RV, Kenis G, Hert MD, Veldink J, Wiuf C, Didriksen M, Craddock N, Owen MJ, O'donovan MC, Borglum AD, Rujescu D, Walter H, Meyer-Lindenberg A, Nothen MM, Ophoff RA, Cichon S (2011). *Association between genetic variation in a region on chromosome 11 and schizophrenia in large samples from Europe*. Mol Psychiatry [Epub ahead of print July 12, 2011]

Rivera-Brugues N, Albrecht B, Wieczorek D, Schmidt H, Keller T, Gohring I, Ekici AB, Tzschach A, Garshasbi M, Franke K, Klopp N, Wichmann HE, Meitinger T, Strom TM, Hempel M (2011). *Cohen syndrome diagnosis using whole genome arrays*. J Med Genet 48:136

Roepcke S, Stahlberg S, Klein H, Schulz MH, Theobald L, Gohlke S, Vingron M, Walther DJ (2011). *A tandem sequence motif acts as a distance-dependent enhancer in a set of genes involved in translation by binding the proteins NonO and SFPQ*. BMC Genomics 12:624

Ropers F, Derivery E, Hu H, Garshasbi M, Karbasiyan M, Herold M, Nurnberg G, Ullmann R, Gautreau A, Sperling K, Varon R, Rajab A (2011). *Identification of a novel candidate gene for non-syndromic autosomal recessive intellectual disability: the WASH complex member SWIP*. Hum Mol Genet 20:2585

Santos-Reboucas CB, Fintelman-Rodrigues N, Jensen LR, Kuss AW, Ribeiro MG, Campos M, Jr., Santos JM, Pimentel MM (2011). *A novel nonsense mutation in KDM5C/JARID1C gene causing intellectual disability, short stature and speech delay*. Neurosci Lett 498:67

Scherthan H, Adelfalk C (2011). *Live cell imaging of meiotic chromosome dynamics in yeast*. Methods Mol Biol 745:537

Scherthan H, Sfeir A, De Lange T (2011). *Rap1-independent telomere attachment and bouquet formation in mammalian meiosis*. Chromosoma 120:151

Schraders M, Haas SA, Weegerink NJ, Oostrik J, Hu H, Hoefsloot LH, Kannan S, Huygen PL, Pennings RJ, Admiraal RJ, Kalscheuer VM, Kunst HP, Kremer H (2011). *Next-generation sequencing identifies mutations of SMPX, which encodes the small muscle protein, X-linked, as a cause of progressive hearing impairment*. Am J Hum Genet 88:628

Sklar P, Ripke S, Scott LJ, Andreassen OA, Cichon S, Craddock N, Edenberg HJ, Nurnberger JI, Jr., Rietschel M, Blackwood D, Corvin A, Flickinger M, Guan W, Mattingsdal M, Mcquillin A, Kwan P, Wienker TF, Daly M, Dudbridge F, Holmans PA, Lin D, Burmeister M, Greenwood TA, Hamshere ML, Muglia P, Smith EN, Zandi PP, Nievergelt CM, Mckinney R, Shilling PD, Schork NJ, Bloss CS, Foroud T, Koller DL, Gershon ES, Liu C, Badner JA, Scheftner WA, Lawson WB, Nwulia EA, Hipolito M, Coryell W, Rice J, Byerley W, Mcmahon FJ,

Schulze TG, Berrettini W, Lohoff FW, Potash JB, Mahon PB, Mcinnis MG, Zollner S, Zhang P, Craig DW, Szelinger S, Barrett TB, Breuer R, Meier S, Strohmaier J, Witt SH, Tozzi F, Farmer A, Mcguffin P, Strauss J, Xu W, Kennedy JL, Vincent JB, Matthews K, Day R, Ferreira MA, O'dushlaine C, Perlis R, Raychaudhuri S, Ruderfer D, Hyoun PL, Smoller JW, Li J, Absher D, Thompson RC, Meng FG, Schatzberg AF, Bunney WE, Barchas JD, Jones EG, Watson SJ, Myers RM, Akil H, Boehnke M, Chambert K, Moran J, Scolnick E, Djurovic S, Melle I, Morken G, Gill M, Morris D, Quinn E, Muhleisen TW, Degenhardt FA, Mattheisen M, Schumacher J, Maier W, Steffens M, Propping P, Nothen MM, Anjorin A, Bass N, Gurling H, Kandaswamy R, Lawrence J, Mcghee K, Mcintosh A, Mclean AW, Muir WJ, Pickard BS, Breen G, St Clair D, Caesar S, Gordon-Smith K, Jones L, Fraser C, Green EK, Grozeva D, Jones IR, Kirov G, Moskvina V, Nikolov I, O'donovan MC, Owen MJ, Collier DA, Elkin A, Williamson R, Young AH, Ferrier IN, Stefansson K, Stefansson H, Thornorgeirsson T, Steinberg S, Gustafsson O, Bergen SE, Nimgaonkar V, Hultman C, Landen M, Lichtenstein P, Sullivan P, Schalling M, Osby U, Backlund L, Frisen L, Langstrom N, Jamain S, Leboyer M, Etain B, Bellivier F, Petursson H, Sigur Sson E, Muller-Mysok B, Lucae S, Schwarz M, Schofield PR, Martin N, Montgomery GW, Lathrop M, Oskarsson H, Bauer M, Wright A, Mitchell PB, Hautzinger M, Reif A, Kelsoe JR, Purcell SM (2011). *Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4*. Nat Genet 43:977

Stacher E, Boldt V, Leibl S, Halbwedl I, Popper HH, Ullmann R, Tavassoli FA, Moinfar F (2011). *Chromosomal aberrations as detected by array comparative genomic hybridization in early low-grade intraepithelial neoplasias of the breast*. Histopathology 59: 549

Stanke F, Becker T, Kumar V, Hedtfeld S, Becker C, Cuppens H, Tamm S, Yarden J, Laabs U, Siebert B, Fernandez L, Macek M, Jr., Radojkovic D, Ballmann M, Greipel J, Cassiman JJ, Wienker TF, Tummler B (2011). *Genes that determine immunology and inflammation modify the basic defect of impaired ion conductance in cystic fibrosis epithelia*. J Med Genet 48:24

Steininger A, Mobs M, Ullmann R, Kochert K, Kreher S, Lamprecht B, Anagnostopoulos I, Hummel M, Richter J, Beyer M, Janz M, Klemke CD, Stein H, Dorken B, Sterry W, Schrock E, Mathas S, Assaf C (2011). *Genomic loss of the putative tumor suppressor gene E2A in human lymphoma*. J Exp Med 208:1585

Strobl-Wildemann G, Kalscheuer VM, Hu H, Wrogemann K, Ropers HH, Tzschach A (2011). *Novel GDI1 mutation in a large family with nonsyndromic X-linked intellectual disability*. Am J Med Genet A 155A:3067

Tzschach A, Ullmann R, Ahmed A, Martin T, Weber G, Decker-Schwering O, Pauly F, Shamdeen MG, Reith W, Oehl-Jaschkowitz B (2011). *Christianson syndrome in a patient with an interstitial Xq26.3 deletion*. Am J Med Genet A 155A:2771

Walther DJ, Stahlberg S, Vowinckel J (2011). *Novel roles for biogenic monoamines: from monoamines in transglutaminase-mediated post-translational protein modification to monoaminylation deregulation diseases*. FEBS J 278:4740

2010

Agay D, Scherthan H, Forcheron F, Grenier N, Herodin F, Meineke V, Drouet M (2010). *Multipotent mesenchymal stem cell grafting to treat cutaneous radiation syndrome: development of a new minipig model*. Exp Hematol 38:945

Andres C, Meyer S, Dina OA, Levine JD, Hucho T (2010). *Quantitative automated microscopy (QuAM) elucidates growth factor specific signalling in pain sensitization*. Mol Pain 6:98

Attarbaschi A, Pisecker M, Inthal A, Mann G, Janousek D, Dworzak M, Potschger U, Ullmann R, Schrappe M, Gadner H, Haas OA, Panzer-Grumayer R, Strehl S (2010). *Prognostic relevance of TLX3 (HOX11L2) expression in childhood T-cell acute lymphoblastic leukaemia treated with Berlin-Frankfurt-Munster (BFM) protocols containing early and late re-intensification elements*. Br J Haematol 148:293

Boldt V, Stacher E, Halbwedl I, Popper H, Hultschig C, Moinfar F, Ullmann R, Tavassoli FA (2010). *Positioning of necrotic lobular intraepithelial neoplasias (LIN, grade 3) within the sequence of breast carcinoma progression*. Genes Chromosomes Cancer 49:463

Budny B, Badura-Stronka M, Materna-Kiryluk A, Tzschach A, Raynaud M, Latos-Bielenska A, Ropers HH (2010). *Novel missense mutations in the ubiquitination-related gene UBE2A cause a recognizable X-linked mental retardation syndrome*. Clin Genet 77:541

Chen W, Ullmann R, Langnick C, Menzel C, Wotschofsky Z, Hu H, Doring A, Hu Y, Kang H, Tzschach A, Hoeltzenbein M, Neitzel H, Markus S, Wiedersberg E, Kistner G, Van Ravenswaaij-Arts CM, Kleefstra T, Kalscheuer VM, Ropers HH. (2010). *Breakpoint analysis of balanced chromosome rearrangements by next-generation paired-end sequencing*. Eur J Hum Genet 18:539

Cordova-Fletes C, Rademacher N, Muller I, Mundo-Ayala J N, Morales-Jeanhs E A, Garcia-Ortiz JE, Leon-Gil A, Rivera H, Dominguez MG, Kalscheuer VM (2010). *CDKL5 truncation due to a t(X;2) (p22.1;p25.3) in a girl with X-linked infantile spasm syndrome*. Clin Genet 77:92

Darvish H, Esmaeeli-Nieh S, Monajemi GB, Mohseni M, Ghasemi-Firouzabadi S, Abedini SS, Bahman I, Jamali P, Azimi S, Mojahedi F, Dehghan A, Shafeghati Y, Jankhah A, Falah M, Soltani Banavandi MJ, Ghani-Kakhi M, Garshasbi M, Rakhshani F, Naghavi A, Tzschach A, Neitzel H, Ropers HH, Kuss AW, Behjati F, Kahrizi K, Najmabadi H (2010). *A clinical and molecular genetic study of 112 Iranian families with primary microcephaly*. J Med Genet 47:823

Endele S, Rosenberger G, Geider K, Popp B, Tamer C, Stefanova I, Milh M, Kortum F, Fritsch A, Pientka FK, Hellenbroich Y, Kalscheuer VM, Kohlhase J, Moog U, Rappold G, Rauch A, Ropers HH, Von Spiczak S, Tonnies H, Villeneuve N, Villard L, Zabel B, Zenker M, Laube B, Reis A, Wieczorek D, Van Maldergem L, Kutsche K (2010). *Mutations in GRIN2A and GRIN2B encoding regulatory subunits of NMDA receptors cause variable neurodevelopmental phenotypes*. Nat Genet 42:1021

Giannandrea M, Bianchi V, Mignogna ML, Sirri A, Carrabino S, D'elia E, Vecellio M, Russo S, Cogliati F,

Larizza L, Ropers HH, Tzschach A, Kalscheuer V, Oehl-Jaschkowitz B, Skinner C, Schwartz CE, Gecz J, Van Esch H, Raynaud M, Chelly J, De Brouwer AP, Toniolo D, D'adamo P (2010). *Mutations in the small GTPase gene RAB39B are responsible for X-linked mental retardation associated with autism, epilepsy, and macrocephaly*. Am J Hum Genet 86:185

Goswami C, Kuhn J, Heppenstall PA, Hucho T (2010). *Importance of non-selective cation channel TRPV4 interaction with cytoskeleton and their reciprocal regulations in cultured cells*. PLoS One 5:e11654

Goswami C, Rademacher N, Smalla KH, Kalscheuer V, Ropers HH, Gundelfinger ED, Hucho T (2010). *TRPV1 acts as a synaptic protein and regulates vesicle recycling*. J Cell Sci 123:2045

Grabundzija I, Irgang M, Mates L, Belay E, Matrai J, Gogol-Doring A, Kawakami K, Chen W, Ruiz P, Chuah MK, Vandendriessche T, Izsvak Z, Ivics Z (2010). *Comparative analysis of transposable element vector systems in human cells*. Mol Ther 18:1200

Grohmann M, Hammer P, Walther M, Paulmann N, Buttner A, Eisenmenger W, Baghai TC, Schule C, Rupprecht R, Bader M, Bondy B, Zill P, Priller J, Walther DJ (2010). *Alternative splicing and extensive RNA editing of human TPH2 transcripts*. PLoS One 5:e8956

Jakubiczka S, Schroder C, Ullmann R, Volleth M, Ledig S, Gilberg E, Kroisel P, Wieacker P (2010). *Translocation and deletion around SOX9 in a patient with acampomelic campomelic dysplasia and sex reversal*. Sex Dev 4:143

Jensen LR, Bartenschlager H, Rujirabanjerd S, Tzschach A, Numann A, Janecke AR, Sporle R, Stricker S, Raynaud M, Nelson J, Hackett A, Fryns JP, Chelly J, De Brouwer AP, Hamel B, Gecz J, Ropers HH, Kuss AW (2010). *A distinctive gene expression fingerprint in mentally retarded male patients reflects disease-causing defects in the histone demethylase KDM5C*. Pathogenetics 3:2

Kariminejad A, Kariminejad R, Tzschach A, Najafi H, Ahmed A, Ullmann R, Ropers HH, Kariminejad MH (2010). *11q14.1-11q22.1 deletion in a 1-year-old male with minor dysmorphic features*. Am J Med Genet A 152A:2651

Kim GJ, Georg I, Scherthan H, Merkenschlager M, Guillou F, Scherer G, Barrionuevo F (2010). *Dicer is required for Sertoli cell function and survival*. Int J Dev Biol 54:867

Kim HG, Ahn JW, Kurth I, Ullmann R, Kim HT, Kulharya A, Ha KS, Itokawa Y, Meliciani I, Wenzel W, Lee D, Rosenberger G, Ozata M, Bick DP, Sherins RJ, Nagase T, Tekin M, Kim SH, Kim CH, Ropers HH, Gusella JF, Kalscheuer V, Choi CY, Layman LC (2010). *WDR11, a WD protein that interacts with transcription factor EMX1, is mutated in idiopathic hypogonadotropic hypogonadism and Kallmann syndrome*. Am J Hum Genet 87:465

Lassmann M, Hanscheid H, Gassen D, Biko J, Meineke V, Reiners C, Scherthan H (2010). *In vivo formation of gamma-H2AX and 53BP1 DNA repair foci in blood cells after radioiodine therapy of differentiated thyroid cancer*. J Nucl Med 51:1318

Laumonnier F, Shoubridge C, Antar C, Nguyen LS, Van Esch H, Kleefstra T, Briault S, Fryns JP, Hamel B, Chelly

J, Ropers HH, Ronce N, Blesson S, Moraine C, Gecz J, Raynaud M (2010). *Mutations of the UPF3B gene, which encodes a protein widely expressed in neurons, are associated with nonspecific mental retardation with or without autism*. Mol Psychiatry 15:767

Liu Y, Hu W, Wang H, Lu M, Shao C, Menzel C, Yan Z, Li Y, Zhao S, Khaitovich P, Liu M, Chen W, Barnes BM, Yan J (2010). *Genomic analysis of miRNAs in an extreme mammalian hibernator, the Arctic ground squirrel*. Physiol Genomics 42A:39

Lugtenberg D, Zangrande-Vieira L, Kirchhoff M, Whibley AC, Oudakker AR, Kjaergaard S, Vianna-Morgante AM, Kleefstra T, Ruiter M, Jehee FS, Ullmann R, Schwartz CE, Stratton M, Raymond FL, Veltman JA, Vrijenhoek T, Pfundt R, Schuurs-Hoeijmakers JH, Hehir-Kwa JY, Froyen G, Chelly J, Ropers HH, Moraine C, Gecz J, Knijnenburg J, Kant SG, Hamel BC, Rosenberg C, Van Bokhoven H, De Brouwer AP (2010). *Recurrent deletion of ZNF630 at Xp11.23 is not associated with mental retardation*. Am J Med Genet A 152A:638

Musante L, Kunde SA, Sulistio TO, Fischer U, Grimme A, Frints SG, Schwartz CE, Martinez F, Romano C, Ropers HH, Kalscheuer VM (2010). *Common pathological mutations in PQBP1 induce nonsense-mediated mRNA decay and enhance exclusion of the mutant exon*. Hum Mutat 31:90

Peleg S, Sananbenesi F, Zovoilis A, Burkhardt S, Bahari-Javan S, Agis-Balboa RC, Cota P, Wittnam JL, Gogol-Doering A, Opitz L, Salinas-Riester G, Dettenhofer M, Kang H, Farinelli L, Chen W, Fischer A (2010). *Altered histone acetylation is associated with age-dependent memory impairment in mice*. Science 328:753

Qiu W, Hu Y, Andersen TE, Jafari A, Li N, Chen W, Kassem M (2010). *Tumor necrosis factor receptor superfamily member 19 (TNFRSF19) regulates differentiation fate of human mesenchymal (stromal) stem cells through canonical Wnt signaling and C/EBP*. J Biol Chem 285:14438

Ropers HH (2010). *'Vom Mütterchen die Frohnatur…': neue Perspektiven für die Aufklärung der Funktion des menschlichen Genoms und Konsequenzen für die Krankenversorgung*. In: Gaterslebener Begegnungen XI, 133. Diskussion V. 148

Ropers HH (2010). *Genetics of early onset cognitive impairment*. Annu Rev Genomics Hum Genet 11:161

Ropers HH (2010). *Single gene disorders come into focus--again*. Dialogues Clin Neurosci 12:95

Salmina K, Jankevics E, Huna A, Perminov D, Radovica I, Klymenko T, Ivanov A, Jascenko E, Scherthan H, Cragg M, Erenpreisa J (2010). *Up-regulation of the embryonic self-renewal network through reversible polyploidy in irradiated p53-mutant tumour cells*. Exp Cell Res 316:2099

Scherthan H, Loidl J (2010). *FISH as a tool to investigate chromosome behavior in budding yeast*. Methods Mol Biol 659:363

Semerci CN, Cinbis M, Ullmann R, Steininger A, Bahce M, Yagci B, Ozden S, Sabir N, Gumus D, Tepeli E, Arteaga J, Mutchinick OM (2010). *Subtelomeric 6p monosomy and 12q trisomy in a patient with a 46,XX,der(6)t(6;12)(p25.3;q24.31) karyotype: Phenotypic overlap with Mutchinick syndrome*. Am J Med Genet A 152A:1724

Shafeghati Y, Kahrizi K, Najmabadi H, Kuss AW, Ropers HH, Tzschach A

(2010). *Brachyphalangy, polydactyly and tibial aplasia/hypoplasia syndrome (OMIM 609945): case report and review of the literature*. Eur J Pediatr 169:1535

Sharbati S, Friedlander MR, Sharbati J, Hoeke L, Chen W, Keller A, Stahler PF, Rajewsky N, Einspanier R (2010). *Deciphering the porcine intestinal microRNA transcriptome*. BMC Genomics 11:275

Storlazzi CT, Lonoce A, Guastadisegni MC, Trombetta D, D'addabbo P, Daniele G, L'abbate A, Macchia G, Surace C, Kok K, Ullmann R, Purgato S, Palumbo O, Carella M, Ambros PF, Rocchi M (2010). *Gene amplification as double minutes or homogeneously staining regions in solid tumors: origin and structure*. Genome Res 20:1198

Thorwarth A, Mueller I, Biebermann H, Ropers HH, Grueters A, Krude H, Ullmann R (2010). *Screening chromosomal aberrations by array comparative genomic hybridization in 80 patients with congenital hypothyroidism and thyroid dysgenesis*. J Clin Endocrinol Metab 95:3446

Timmermann B, Kerick M, Roehr C, Fischer A, Isau M, Boerno ST, Wunderlich A, Barmeyer C, Seemann P, Koenig J, Lappe M, Kuss AW, Garshasbi M, Bertram L, Trappe K, Werber M, Herrmann BG, Zatloukal K, Lehrach H, Schweiger MR (2010) *Somatic mutation profiles of MSI and MSS colorectal cancer identified by whole exome next generation sequencing and bioinformatics analysis*. PLoS One 5(12):e15661

Trimborn M, Ghani M, Walther DJ, Dopatka M, Dutrannoy V, Busche A, Meyer F, Nowak S, Nowak J, Zabel C, Klose J, Esquitino V, Garshasbi M, Kuss AW, Ropers HH, Mueller S, Poehlmann C, Gavvovidis I, Schindler D, Sperling K, Neitzel H (2010). *Establishment of a mouse model with misregulated chromosome condensation due to defective Mcph1 function*. PLoS One 5:e9242

Tzschach A, Bisgaard AM, Kirchhoff M, Graul-Neumann LM, Neitzel H, Page S, Ahmed A, Muller I, Erdogan F, Ropers HH, Kalscheuer VM, Ullmann R (2010). *Chromosome aberrations involving 10q22: report of three overlapping interstitial deletions and a balanced translocation disrupting C10orf11*. Eur J Hum Genet 18:291

Tzschach A, Menzel C, Erdogan F, Istifli ES, Rieger M, Ovens-Raeder A, Macke A, Ropers HH, Ullmann R, Kalscheuer V (2010). *Characterization of an interstitial 4q32 deletion in a patient with mental retardation and a complex chromosome rearrangement.* Am J Med Genet A 152A(4):1008

Urcan E, Haertel U, Styllou M, Hickel R, Scherthan H, Reichl FX (2010). *Real-time xCELLigence impedance analysis of the cytotoxicity of dental composite components on human gingival fibroblasts*. Dent Mater 26:51

Urcan E, Scherthan H, Styllou M, Haertel U, Hickel R, Reichl FX (2010). *Induction of DNA double-strand breaks in primary gingival fibroblasts by exposure to dental resin composites*. Biomaterials 31:2010

Walczak-Sztulpa J, Eggenschwiler J, Osborn D, Brown DA, Emma F, Klingenberg C, Hennekam RC, Torre G, Garshasbi M, Tzschach A, Szczepanska M, Krawczynski M, Zachwieja J, Zwolinska D, Beales PL, Ropers HH, Latos-Bielenska A, Kuss AW (2010). *Cranioectodermal Dysplasia, Sensenbrenner syndrome, is a ciliopathy caused by mutations in the IFT122 gene*. Am J Hum Genet 86:949

Zanni G, Van Esch H, Bensalem A, Saillour Y, Poirier K, Castelnau L, Ropers HH, De Brouwer AP, Laumonnier F, Fryns JP, Chelly J (2010). *A novel mutation in the DLG3 gene encoding the synapse-associated protein 102 (SAP102) causes non-syndromic mental retardation*. Neurogenetics 11:251

2009

Adelfalk C, Janschek J, Revenkova E, Blei C, Liebe B, Gob E, Alsheimer M, Benavente R, De Boer E, Novak I, Hoog C, Scherthan H, Jessberger R (2009). *Cohesin SMC1beta protects telomeres in meiocytes*. J Cell Biol 187:185

Barrionuevo F, Georg I, Scherthan H, Lecureuil C, Guillou F, Wegner M, Scherer G (2009). *Testis cord differentiation after the sex determination stage is independent of Sox9 but fails in the combined absence of Sox9 and Sox8*. Dev Biol 327:301

Bashiardes S, Kousoulidou L, Van Bokhoven H, Ropers HH, Chelly J, Moraine C, De Brouwer AP, Van Esch H, Froyen G, Patsalis PC (2009). *A new chromosome x exon-specific microarray platform for screening of patients with X-linked disorders*. J Mol Diagn 11:562

Erenpreisa J, Cragg MS, Salmina K, Hausmann M, Scherthan H (2009). *The role of meiotic cohesin REC8 in chromosome segregation in gamma irradiation-induced endopolyploid tumour cells*. Exp Cell Res 315:2593

Fu X, Fu N, Guo S, Yan Z, Xu Y, Hu H, Menzel C, Chen W, Li Y, Zeng R, Khaitovich P (2009). *Estimating accuracy of RNA-Seq and microarrays with proteomics*. BMC Genomics 10:161

Garcia-Cruz R, Roig I, Robles P, Scherthan H, Garcia Caldes M (2009). *ATR, BRCA1 and gammaH2AX localize to unsynapsed chromosomes at the pachytene stage in human oocytes*. Reprod Biomed Online 18:37

Graul-Neumann LM, Stieler KM, Blume-Peytavi U, Tzschach A (2009). *Autosomal dominant inheritance in a large family with focal facial dermal dysplasia (Brauer-Setleis syndrome)*. Am J Med Genet A 149A:746

Grohmann M, Paulmann N, Fleischhauer S, Vowinckel J, Priller J, Walther DJ (2009). *A mammalianized synthetic nitroreductase gene for high-level expression*. BMC Cancer 9:301

Haensel J, Kohlschmidt N, Pitz S, Keilmann A, Zenker M, Ullmann R, Haaf T, Bartsch O (2009). *Case report supporting that the Barber-Say and ablepharon macrostomia syndromes could represent one disorder*. Am J Med Genet A 149A: 2236

Hilhorst-Hofstee Y, Tumer Z, Born P, Knijnenburg J, Hansson K, Yatawara V, Steensberg J, Ullmann R, Arkesteijn G, Tommerup N, Larsen LA (2009). *Molecular characterization of two patients with de novo interstitial deletions in 4q22-q24*. Am J Med Genet A 149A:1830

Hu H, Wrogemann K, Kalscheuer V, Tzschach A, Richard H, Haas SA, Menzel C, Bienek M, Froyen G, Raynaud M, Van Bokhoven H, Chelly J, Ropers H, Chen W (2009). *Mutation screening in 86 known X-linked mental retardation genes by droplet-based multiplex PCR and massive parallel sequencing*. Hugo J 3:41 Erratum 3:83

Hu HY, Yan Z, Xu Y, Hu H, Menzel C, Zhou YH, Chen W, Khaitovich P (2009). *Sequence features associated with microRNA strand selection in humans and flies*. BMC Genomics 10:413

Kahrizi K, Najmabadi H, Kariminejad R, Jamali P, Malekpour M, Garshasbi M, Ropers HH, Kuss AW, Tzschach A (2009). *An autosomal recessive syndrome of severe mental retardation, cataract, coloboma and kyphosis maps to the pericentromeric region of chromosome 4*. Eur J Hum Genet 17:125

Kalscheuer VM, Musante L, Fang C, Hoffmann K, Fuchs C, Carta E, Deas E, Venkateswarlu K, Menzel C, Ullmann R, Tommerup N, Dalpra L, Tzschach A, Selicorni A, Luscher B, Ropers HH, Harvey K, Harvey RJ (2009). *A balanced chromosomal translocation disrupting ARHGEF9 is associated with epilepsy, anxiety, aggression, and mental retardation*. Hum Mutat 30:61

Kariminejad A, Kariminejad R, Tzschach A, Ullmann R, Ahmed A, Asghari-Roodsari A, Salehpour S, Afroozan F, Ropers HH, Kariminejad MH (2009). *Craniosynostosis in a patient with 2q37.3 deletion 5q34 duplication: association of extra copy of MSX2 with craniosynostosis*. Am J Med Genet A 149A:1544

Krauss S, So J, Hambrock M, Kohler A, Kunath M, Scharff C, Wessling M, Grzeschik KH, Schneider R, Schweiger S (2009). *Point mutations in GLI3 lead to misregulation of its subcellular localization*. PLoS One 4:e7471

Lugtenberg D, Kleefstra T, Oudakker AR, Nillesen WM, Yntema HG, Tzschach A, Raynaud M, Rating D, Journel H, Chelly J, Goizet C, Lacombe D, Pedespan JM, Echenne B, Tariverdian G, O'Rourke D, King MD, Green A, Van Kogelenberg M, Van Esch H, Gecz J, Hamel BC, Van Bokhoven H, De Brouwer AP (2009). *Structural variation in Xq28: MECP2 duplications in 1% of patients with unexplained XLMR and in 2% of male patients with severe encephalopathy*. Eur J Hum Genet 17:444

Mangerich A, Scherthan H, Diefenbach J, Kloz U, Van Der Hoeven F, Beneke S, Burkle A (2009). *A caveat in mouse genetic engineering: ectopic gene targeting in ES cells by bidirectional extension of the homology arms of a gene replacement vector carrying human PARP-1*. Transgenic Res 18:261

Mir A, Kaufman L, Noor A, Motazacker MM, Jamil T, Azam M, Kahrizi K, Rafiq MA, Weksberg R, Nasr T, Naeem F, Tzschach A, Kuss AW, Ishak GE, Doherty D, Ropers HH, Barkovich AJ, Najmabadi H, Ayub M, Vincent JB (2009). *Identification of mutations in TRAPPC9, which encodes the NIK- and IKK-beta-binding protein, in nonsyndromic autosomal-recessive mental retardation*. Am J Hum Genet 85:909

Neumann TE, Allanson J, Kavamura I, Kerr B, Neri G, Noonan J, Cordeddu V, Gibson K, Tzschach A, Kruger G, Hoeltzenbein M, Goecke TO, Kehl HG, Albrecht B, Luczak K, Sasiadek MM, Musante L, Laurie R, Peters H, Tartaglia M, Zenker M, Kalscheuer V (2009). *Multiple giant cell lesions in patients with Noonan syndrome and cardio-facio-cutaneous syndrome*. Eur J Hum Genet 17:420

Pouya AR, Abedini SS, Mansoorian N, Behjati F, Nikzat N, Mohseni M, Nieh SE, Abbasi Moheb L, Darvish H, Monajemi G B, Banihashemi S, Kahrizi K, Ropers HH, Najmabadi H (2009). *Fragile X syndrome screening of families with consanguineous and non-consanguineous parents in the Iranian population*. Eur J Med Genet 52:170

Raile K, Klopocki E, Holder M, Wessel T, Galler A, Deiss D, Muller

D, Riebel T, Horn D, Maringa M, Weber J, Ullmann R, Gruters A (2009). *Expanded clinical spectrum in hepatocyte nuclear factor 1b-maturity-onset diabetes of the young*. J Clin Endocrinol Metab 94:2658

Scherthan H (2009). *Analysis of telomere dynamics in mouse spermatogenesis*. Methods Mol Biol 558:383

Schmidt S, Hommel A, Gawlik V, Augustin R, Junicke N, Florian S, Richter M, Walther DJ, Montag D, Joost HG, Schurmann A (2009). *Essential role of glucose transporter GLUT3 for post-implantation embryonic development*. J Endocrinol 200:23

Seifert W, Holder-Espinasse M, Kuhnisch J, Kahrizi K, Tzschach A, Garshasbi M, Najmabadi H, Walter Kuss A, Kress W, Laureys G, Loeys B, Brilstra E, Mancini GM, Dollfus H, Dahan K, Apse K, Hennies HC, Horn D (2009). *Expanded mutational spectrum in Cohen syndrome, tissue expression, and transcript variants of COH1*. Hum Mutat 30:E404

Shoichet SA, Waibel S, Endruhn S, Sperfeld AD, Vorwerk B, Muller I, Erdogan F, Ludolph AC, Ropers HH, Ullmann R (2009). *Identification of candidate genes for sporadic amyotrophic lateral sclerosis by array comparative genomic hybridization*. Amyotroph Lateral Scler 10:162

Stoeckius M, Maaskola J, Colombo T, Rahn HP, Friedlander MR, Li N, Chen W, Piano F, Rajewsky N (2009). *Large-scale sorting of C. elegans embryos reveals the dynamics of small RNA expression*. Nat Methods 6:745

Tarpey PS, Smith R, Pleasance E, Whibley A, Edkins S, Hardy C, O'meara S, Latimer C, Dicks E, Menzies A, Stephens P, Blow M, Greenman C, Xue Y, Tyler-Smith C, Thompson D, Gray K, Andrews J, Barthorpe S, Buck G, Cole J, Dunmore R, Jones D, Maddison M, Mironenko T, Turner R, Turrell K, Varian J, West S, Widaa S, Wray P, Teague J, Butler A, Jenkinson A, Jia M, Richardson D, Shepherd R, Wooster R, Tejada MI, Martinez F, Carvill G, Goliath R, De Brouwer AP, Van Bokhoven H, Van Esch H, Chelly J, Raynaud M, Ropers HH, Abidi FE, Srivastava AK, Cox J, Luo Y, Mallya U, Moon J, Parnau J, Mohammed S, Tolmie JL, Shoubridge C, Corbett M, Gardner A, Haan E, Rujirabanjerd S, Shaw M, Vandeleur L, Fullston T, Easton DF, Boyle J, Partington M, Hackett A, Field M, Skinner C, Stevenson RE, Bobrow M, Turner G, Schwartz CE, Gecz J, Raymond FL, Futreal PA, Stratton MR (2009). *A systematic, large-scale resequencing screen of X-chromosome coding exons in mental retardation*. Nat Genet 41:535

Turkmen S, Guo G, Garshasbi M, Hoffmann K, Alshalah AJ, Mischung C, Kuss A, Humphrey N, Mundlos S, Robinson PN (2009). *CA8 mutations cause a novel syndrome characterized by ataxia and mild mental retardation with predisposition to quadrupedal gait*. PLoS Genet 5:e1000487

Tzschach A (2009). Genetik der nicht-syndromalen geistigen Behinderung. Medgen 21:231

Tzschach A, Bozorgmehr B, Hadavi V, Kahrizi K, Garshasbi M, Motazacker MM, Ropers HH, Kuss AW, Najmabadi H (2008). *Alopecia-mental retardation syndrome: clinical and molecular characterization of four patients*. Br J Dermatol 159:748

Tzschach A, Graul-Neumann LM, Konrat K, Richter R, Ebert G, Ullmann R, Neitzel H (2009). *Interstitial deletion 2p11.2-p12: report of a patient with mental retardation and review of the literature*. Am J Med Genet A 149A:242

Tzschach A, Ramel C, Kron A, Seipel B, Wuster C, Cordes U, Liehr T, Hoeltzenbein M, Menzel C, Ropers HH, Ullmann R, Kalscheuer V, Decker J, Steinberger D (2009). *Hypergonadotropic hypogonadism in a patient with inv ins (2;4)*. Int J Androl 32:226

Ullmann R(2009). *Array Comparative Genomic Hybridization in Pathology*, pp87-96. In: Basic concepts of Molecular Pathology, Eds: Cagle PT and Craig A. Springer Verlag Heidelberg [ISBN: 978-0-387-89625-0]

Zatkova A, Merk S, Wendehack M, Bilban M, Muzik EM, Muradyan A, Haferlach C, Haferlach T, Wimmer K, Fonatsch C, Ullmann R (2009). *AML/MDS with 11q/MLL amplification show characteristic gene expression signature and interplay of DNA copy number changes*. Genes Chromosomes Cancer 48:510

Zhang L, Tumer Z, Mollgard K, Barbi G, Rossier E, Bendsen E, Moller RS, Ullmann R, He J, Papadopoulos N, Tommerup N, Larsen LA (2009). *Characterization of a t(5;8)(q31;q21) translocation in a patient with mental retardation and congenital heart disease: implications for involvement of RUNX1T1 in human brain and heart development*. Eur J Hum Genet 17:1010

Invited plenary lectures (H.-Hilger Ropers)

Genetic analysis of cognitive disorders and related complex diseases: why GWAS failed. Hinxton Advanced Courses: Human Genome Analysis: Genetic Analysis of Multifactorial Diseases, Wellcome Trust, Hinxton, July 11-17, 2012

Predictive diagnosis of cognitive impairment: Where are we heading? Molecular Neurobiology Today and Tomorrow, Russian-German Symposium of the Russian Academy of Sciences and the Berlin-Brandenburgische Akademie der Wissenschaften, Moscow, April 25-29, 2012

Genetics of Cognitive Disorders and Intellectual Disabilities. Golden Helix Conference 2012, Genomic Medicine: Translating Genes into Health, Turin, April 19, 2012

Keynote: *The Sequencing Revolution: Implications for Genetic Research and Health Care.* ISHG 2012, International Conference on Genes, Genetics & Genomics, Today & Tomorrow –Human Concerns– and XXXVII Annual Conference of the Indian Society of Human Genetics, Panjab University, Chandigarh, India, March 03-05, 2012

Geistige Behinderung und verwandte genetische Störungen: neue Perspektiven für Forschung und Krankenversorgung. 1. DFG Winter School: Seltene Erkrankungen mit Beginn im Kindesalter, Göttingen, February 15-17, 2012

Hält die personalisierte Medizin, was sie verspricht? Handelsblatt Konferenz Pharma 2012, Frankfurt/Main, February 09, 2012

Re-Focusing Genome Research on Single Gene Disorders: an Opportunity – and an Obligation. Philippine Genome Center Manila, November 28, 2011

Genetics of Autism., 1st International and 5th National Congress of the Iranian Neurogenetic Society, Tehran, November 23-25, 2011

Next Generation Sequencing und Perspektiven für die Bioanalytik. 3. Potsdamer Bioanalytik Kolloquium, Golm, November 10, 2011

On the future of genetic risk assessment. Workshop Personalized Medicine, Leopoldina/Acatech, Berlin, November 09, 2011

Keynote: *High-Throughput sequencing in Intellectual Disability research and Health-Care: The Future is now.* 16th Annual Meeting of the German Society of Neurogenetics, Erlangen, Germany, October 27, 2011

High Throughput-Low Cost Sequencing: Implications for Genome Research and Health Care. Children's Mercy Hospital, Kansas City, October 04, 2011

Next Generation Sequencing als universaler Test für Screening und prädiktive Diagnostik. Öquasta Symposium, Igls/Innsbruck, September 29-30, 2011

Deep sequencing reveals 50 novel genes for recessive cognitive disorders. NGFN meeting 2011, Berlin, September 27, 2011

New Sequencing Technologies and Consequences for Genome Applications. EU-Workshop on 'Privacy Issues Arising from Next Generation Sequencing', Brussels, June 01, 2011

Unravelling the Genetic Heterogeneity of Cognitive Impairment by Next Generation Sequencing. Wilhelm Johannsen Centre, Copenhagen, March 22, 2011

The 1000 $ Genome: implications for genome research and health care. Emory University, Atlanta, Georgia, September 28, 2011

Comprehensive Mutation and Carrier Screening in X-linked disorders. 8th Annual Molecular Diagnostics Conference (Tri-Con), San Francisco, February 23-28, 2011

Recent Advances in the elucidation of mental retardation. 4th Iranian Annual Neurogenetic Congress (Advances in Neurogenetics), Tehran, November 23, 2010

Keynote: *Molecular dissection of intellectual disability: of chromosomes aberrations, linkage studies and high-throughput sequencing.* Mental retardation: from genes to synapses, functions and dysfunctions. CNRS-Conférences Jacques Monod, Roscoff, France, October 7-11, 2010

X-linked forms of mental retardation. International Conference on ‚Genetics and Neurobiology of Mental Retardation', and Biennial conference of the German Academy of Sciences Leopoldina/National Academy of Sciences, Erlangen, September 29-October 1, 2010

Next-Generation-Sequencing und die Zukunft der genetischen Diagnostik. Workshop Individualisierte Prävention der Deutschen Gesellschaft für Sozialmedizin und Prävention, DGSMP, Berlin, September 24, 2010

Das 1000 $ Genom. BBAW Tagung "Leben 3.0 und die Zukunft der Evolution", Berlin, September 17, 2010

Zur Zukunft der Genomforschung und prädiktiven genetischen Diagnostik. 9. Cadenabbia Gespräch Medizin – Ethik – Recht „Medizin nach Maß", Cadenabbia, September 09-12, 2010

Neueste Entwicklungen auf dem Gebiet der NGS: Implikationen für die humangenetische Diagnostik. AG Genomics und Bioinformatik DGKL, Evangelische Akademie Tutzing, May 07, 2010

Comprehensive Carrier Screening. International Symposium of the Academies „Predictive genetic diagnostics as an instrument for

disease prevention“ Leopoldina/ BBAW, Bonn, February 07-08, 2010

Appointments of former members of the department

Tim Hucho: *Professorship (W2) on Anaesthesiology and Pain Research,* University Hospital of Cologne, 2012

Diego Walther: *scientist,* Technische Fachhochschule Wildau, 2012

Andreas Tzschach: *scientist,* Institute of Human Genetics, University Hospital of Tübingen, 2011

Andreas Walter Kuss: *Professorship (W2) on Molecular Human Genetics*, Ernst-Moritz-Arndt Universität, Greifswald, 2010

Sybille Krauss: *Groupleader*, Deutsches Zentrum für Neurodegenerative Erkrankungen (DZNE), Bonn, 2010

Habilitationen / state doctorates

Reinhard Ullmann**:** *state doctorate (Habilitation) on Genetics and Cell Biology,* University of Salzburg, Austria, 2011

PhD theses

Lia Abbasi-Moheb: *Identification of three novel genes for autosomal recessive intellectual disability and molecular characterization of the causative defects,* Freie Universität Berlin, 2012

Roxana Kariminejad: *Copy Number Variations in Structural Brain Malformations*, Freie Universität Berlin, 2012

Hyung-Goo Kim: *Positional Cloning of New Disease Genes for Kallmann Syndrome and Potocki-shaffer Syndrome,* Freie Universität Berlin, 2012

Silke Stahlberg: Die Serotonylierung von Transkriptionsfaktoren als Mechanismus in der Pathogenese des Alkoholismus, Freie Universität Berlin, 2012

Christine Andres: *Growth Factor-Induced Subgroup Specific Signaling in Nociceptive Neurons,* Freie Universität Berlin, 2011

Sahar Esmaeeli-Nieh: *Identification and functional characterization of a genetic defect in the kinetochore protein BOD1 associated with autosomal recessive mental retardation and oligomenorrhea,* Freie Universität Berlin, 2010

Eva Kickstein: *Einfluß Mikrotubulus-assoziierter PP2A auf neurodegenerative Erkrankungen,* Freie Universität Berlin, 2010

Julia Kuhn: *Estrogen Signaling in Nociceptive Neurons*, Freie Universität Berlin, 2010

Jakob Vowinckel: *Die Protein-Monoaminylierung als regulatorischer Mechanismus in der Signaltransduktion,* Freie Universität Berlin, 2010

Masoud Garshasbi: *Identification of 31 genomic loci for autosomal recessive mental retardation and molecular genetic characterization of novel causative mutations in four genes*, Freie Universität Berlin, 2009

Stella Amrei Kunde: *Untersuchung zur Funktion des PQBP1-Komplexes*, Freie Universität Berlin, 2009

Artur Muradyan: *SPON2 and its implication in epithelial-mesenchymal transition*, Freie Universität Berlin, 2009

Nils Rademacher: *Untersuchungen zur Funktion der Kinase CDKL5/ STK9,* Freie Universität Berlin, 2009

Student theses

Katharina Frieß: *Synthese, Aufreinigung und biologische Testung von biotinylierten, monoaminergenen Neurotransmittern*. Diploma Thesis, Beuth Hochschule für Technik, Berlin, 2012

Werner Irrgang: *Charakterisierung von Aptameren zur Detektion monoaminylierter Proteine.* Diploma Thesis, Freie Universität Berlin, 2012

Claudia Strecke: *Untersuchung zur Serotonylierung in der zytotoxischen T-Zellreaktion von Jurkats und deren membrangebundenen Proteinen Rab3a und Rab27a.* Master Thesis, Hochschule Albstadt, Sigmaringen, 2012

Jana Grune: *Chromatindomänen und ihr Einfluss auf die DNA-Degradation während der Apoptose.* Master Thesis, Charité – Universitätsmedizin Berlin, 2011

Florian Hinze: *Einfluss biogener Monoamine auf die Leberregeneration.* Master Thesis, Universität zu Lübeck, 2011

Diana Karweina: *Die serotonerge Abhängigkeit bei Alkoholismus und anderen psychiatrischen Krankheiten*. Diploma Thesis, Technische Universität Berlin, 2011

Björn Samans: *Serotonylierung von Rab3a und Rab27a in der zytotoxischen T-Zell-Reaktion*. Bachelor Thesis, Beuth Hochschule für Technik Berlin, 2011

Andreas Schmidt: *Einfluss der Monoaminylierung auf die Aktivität von glutaminreichen Transkriptionsfaktoren.* Diploma Thesis, Freie Universität Berlin, 2011

Carsten Wenzel: Transciptome analysis of TRPV1-positive rat dorsal root ganglion neurons. Diploma Thesis, Technische Universität Berlin, 2011

Tilo Knape: *Der Einfluss von Phosphodiesterasen und Serotonin auf das Proliferationsverhalten von Prostatakarzinomzellen*. Diploma Thesis, Freie Universität Berlin, 2010

Alexander Reichenbach: *Molecular characterisation of gene defects in hereditary forms of intellectual disability*. Diploma Thesis, Freie Universität Berlin, 2010

Gunnar Seidel: *Charakterisierung der Monoaminylierung in der Adipozyten-Zelllinie 3T3-L1*. Diploma Thesis, Universität Potsdam, 2010

Franziska Sotzny: *Der Einfluss der Serotonylierung auf die Proliferation von Karzinomzellen.* Master Thesis, Universität Bayreuth, 2010

Caroline Zerbst: *Optimierung des NTR/CB1954-Systems für die Anwendung in Säugerzellen.* Diploma Thesis, Universität Potsdam, 2010

Christin Bäth: *Monoaminylierung von Transkriptionsfaktoren am Beispiel des CLOCK*. Bachelor Thesis, Freie Universität Berlin, 2009

B. Behrendt: *Modellierung der serotonininduzierten Ausschüttung dense-granulärer Botenstoffe aus Thrombozyten*. Bachelor Thesis, Technische Fachhochschule Wildau, 2009

M. Bienemann: *Modellierung der serotonininduzierten Ausschüttung des Insulins aus β-Zellen des Pankreas*.

Bachelor Thesis, Technische Fachhochschule Wildau, 2009

Grit Ebert: *Veränderungen der Genexpressionsmuster im Verlauf der Retinsäure-induzierten Zell-differenzierung.* Diploma Thesis, Humboldt Universität zu Berlin, 2009

A. Funke: *Identifikation von BRN2-regulierten Genkandidaten und Untersuchung der Auswirkung des SNPs rs16967794 auf die Bindungsaffinität von BRN2*. Bachelor Thesis, Technische Fachhochschule Wildau, 2009

Sören Mucha: *Studien zum Einfluss von Serotonin auf die Glucose-induzierte Insulinsekretion*. Bachelor Thesis, Technische Fachhochschule Wildau, 2009

Fabian Roske: *Die Veränderung der Aufnahme und Inkorporation von biogenen Monoaminen durch Adipogenese in 3T3-L1-Zellen.* Bachelor Thesis, Universität Bayreuth, 2009

B. Schoder: *Vergleich viraler Promotoren mit dem humanen ribosomalen RPL12-Promotor in eukaryotischen Expressionssystemen*. Bachelor Thesis, Technische Fachhochschule Wildau, 2009

Caroline Schwarzer: *Etablierung eines immunologischen Nachweissystems zur Erfassung des Aktivierungszustandes der kleinen GFTPase Cdc42*. Bachelor Thesis, Universität Bayreuth, 2009

Christine Technau: *Monoaminylierung von Transkriptionsfaktoren am Beispiel des CREB*. Bachelor Thesis, Freie Universität Berlin, 2009

Lars Theobald: *Untersuchung und Charakterisierung transaktivierender Eigenschaften monoaminylierter Transkriptionsfaktoren.* Diploma Thesis, Technische Fachhochschule Berlin, 2009

Zofia Wotschofsky: *Translocation breakpoint mapping by Illumina/ Solexa technology*. Diploma Thesis, Technische Universität Berlin, 2009

Teaching activities

Reinhard Ullmann: Lecture "Modern methods for the analysis of genetic and epigenetic changes in human genetics and tumor biology" at the University of Salzburg, Austria, 2009, 2010, 2011

Diego Walther, Technische Fachhochschule Wildau, Freie Universität Berlin

Guest scientists

Zohreh Fattahi, The Social Welfare and Rehabilitation Sciences University, Tehran, Iran, 10/11-01/12

Farzaneh Larti, The Social Welfare and Rehabilitation Sciences University, Tehran, Iran, 09/11-12/11

Mohammad Javad Soltani Banavandi, The Social Welfare and Rehabilitation Sciences University, Tehran, Iran, 03/10-05/10

Usha Dutta, Centre for DNA fingerprinting and diagnostics (CDFD), Hyderabad, India, 07/12-09/12

Organization of scientific events

The Department organized the *15th International Workshop on Fragile X and Mental Retardation* at the Harnack House Berlin, September 5-7, 2011.

Seminars and lectures given by external speakers in the department

Magdalena Skipper, Senior Editor Nature, London, UK, 15.01.2010. *Editorial processes at Nature – a look behind the scenes*

Torsten Hartwig, Fordham University, New York, 10.11.2010. *The Kv1.2 potassium channel: the position of an N-glycan on the extracellular linkers affects its protein expression and function*

Stephen Kingsmore, Chief Scientific Officer, National Center for Genome Resources, Santa Fe, New Mexico, 13.12.2010. *Sequencing Genomes and Exomes to understand the Genetics of Mendelian Diseases* (Dahlem Colloquium)

Anibal Garza-Carbajal, University Medical Center, Utrecht, NL, 09.06.2011. *GRK2 in the transition from acute to chronic pain*

Caleb Webber, Programme Leader, MRC Functional Genomics Unit, Dept. of Physiology, Anatomy & Genetics, Oxford University, UK, 18.01.2012. *Mouse functional genomics approaches to human disease*

Department of Computational Molecular Biology

(Established: 10/2000)

Head
Prof. Dr. Martin Vingron
Phone: +49 (0)30 8413-1150
Fax: +49 (0)30 8413-1152
Email: vingron@molgen.mpg.de

Secretary
Martina Lorse
Phone: +49 (0)30 8413-1151
Fax: +49 (0)30 8413-1152
Email: vinoffic@molgen.mpg.de

Scientific assistant
Patricia Marquardt
Phone: +49 (0)30 8413-1716
Fax: +49 (0)30 8413-1671
Email:
patricia.marquardt@molgen.mpg.de

IMPRS coordinator
Kirsten Kelleher
Phone: +49 (0)30 8413-1154
Fax: +49 (0)30 8413-1152
Email: kelleher@molgen.mpg.de

Computer systems administrator
Wilhelm Rüsing

Introduction

Computational biology studies biological questions with mathematical and computational methods. In the area of molecular biology and genomics, the possibility to apply such formal methods, of course, comes from the availability not only of genome sequences, but also of large amounts of functional data about biological processes. Computational molecular biology encompasses both the development and the adaptation of methods from mathematics, statistics, and computer science, as well as pursuing biological questions and applying these tools in close collaborations with experimentalists. The research interest of the Computational Molecular Biology Department lies in understanding of gene regulatory mechanisms in the context of the structure and evolution of eukaryotic genomes. To this end, mathematical, computational, and also experimental approaches are being developed and employed. At the MPIMG, computational approaches have also become an integral part of many research projects.

Structure and organization of the department

The department comprises scientists coming from various backgrounds, ranging from mathematics, statistics and computer science *via* physics to biology and genetics. They are organized in several project groups, the largest of which is the Transcriptional Regulation Group headed by Martin Vingron. The work of this group focuses on theoretical concepts in the prediction of cis-regulatory elements, gene regulatory networks and epigenetic aspects of regulation. Closely associated to Vingron's work are the project groups of Stefan Haas and Sebastiaan Meijsing. Stefan Haas heads the project "Gene Structure and Array Design". The focus of this group is on transcriptomics and gene structure. As of September 2009, Sebastiaan Meijsing has built up a new, experimental group that works on glucocorticoid receptors as a model transcription factor. Ho-Ryun Chung has been a senior postdoc in the Vingron group until mid 2011, at which point he was appointed as leader of a Max Planck Research Group. As such he and his new group are now part of the Otto Warburg Laboratory at the institute. Peter Arndt heads the Evolutionary Genomics Group, a research group within the department, which works on developing models how the DNA in humans and other primates has evolved. In January 2012, Ulf Ørom started as a project leader working on non-coding RNAs. Later in the year, Ørom received a Sofja Kovalevskaya Award of the Alexander von Humboldt Foundation, which will fund his own group for the coming five years. Having his own funds, and with the endorsement of the vice president of the MPG, Ørom has been appointed as leader of an independent research group by the Institute and transferred with his group to the Otto Warburg Laboratory of the MPIMG.

Scientific methods and findings

During the last couple of years, the efforts on prediction of transcription factor binding sites have been continued and the techniques have matured to a state, where we are actually offering a web server and software packages for transcription factor target prediction, motif enrichment, and ChIP-seq peak analysis. These methods have been widely applied, both within our projects and various collaborations. Most notably, in the context of an international collaboration, application of the target prediction methods has led to the discovery of a network of inflammation-related genes regulated by the transcription factor Irf7, with genes from this network contributing the Type I diabetes risk in humans.

The efforts on understanding the contribution of chromatin structure and epigenetic marks have led to a model for predicting gene expression from the histone modifications that are observed in promoter regions of genes. Remarkably, this analysis unravelled that transcription factors play only a minor role, which contradicts the traditional understanding. However, it reinforces the point that transcriptional control is exerted on many levels and those levels appear to be highly interrelated. This work has had considerable impact in the community and has given rise to many similar analyses.

Evolutionary analysis of genomes and of genetic regulation, in the Arndt and Vingron groups, respectively, has yielded new insights into forces shaping our genome and its regulatory networks. Many of these findings rest on Arndt's detailed study of the Cytosine methylation-deamination process as it can be extracted from the differences between complete, very similar genomes. This has led to models for evolution of transcription factor binding sites, insight into the plasticity of regulatory connections due to transposable elements, and to the observation

of particular statistical asymmetries around transcription starts sites, which are likely a reflection of the process of transcription itself.
In an intensive collaboration with the Human Molecular Genetics Department, Stefan Haas and co-workers have designed and implemented analysis pipelines for next generation sequencing data, in particular RNA-seq data and ChIP-seq data. The RNA-seq pipeline is particularly geared toward identification of mutations in sequenced genomes of patients with particular disease phenotypes. Within this collaboration, it was employed, among others, to uncover mutations involved in intellectual disability. The ChIP-seq pipeline serves primarily the analysis of gene regulatory networks.
As can be seen from this short summary, the methods employed in the department are, with the exception of the experimental work of Sebastiaan Meijsing, mostly theoretical. Mathematics, statistics, and computer science supply the basis for the bioinformatics analysis performed. Projects may either approach biological questions through theoretical analysis, or may be collaborative projects, where frequent interaction and feedback between experimentalists and theoreticians drive the work forward.

Material resources and equipment

The theoretical work of the Computational Molecular Biology Department relies heavily on powerful computers. The computer equipment of the department comprises nine compute servers with 48 or 64 processors and containing 256GB or 512 GB of RAM, respectively. This architecture serves the classical sequence analysis, the numerical calculations, and the analysis of next generation sequencing data. The computers are not accessed by the researchers directly, but through a queuing system. This allows for very efficient utilization of the entire infrastructure. Storage space on hard disks in the institute comprises approximately 3.8 PB and the department participates in this. The computer set-up is maintained by the department system administrator, Wilhelm Rüsing, in close cooperation with the institute's computing unit. Sebastiaan Meijsing, who does experimental work, has laboratory space in tower 4.

Cooperation with national and international research institutions

Department members contribute substantially to the bioinformatics curriculum at Free University of Berlin. They teach courses like *Algorithmic Bioinformatics* and *Population Genetics*. Students can do internships, practical courses, and thesis work with members of the department. This brings many bright, young students to the department and at the same time allows the university to show the students a much larger spectrum of bioinformatics than would normally be possible in the university framework. In cooperation with the university, the department has established the International Max Planck Research School for Computational Biology and Scientific Computing (IMPRS-CBSC, http://www.imprs-cbsc.mpg.de/), which is now in its second 6-year funding period. Starting with this second period, the IMPRS-CBSC has been extended to encompass a parallel graduate school at the CAS-MPG Partner Institute for Computational Biology (PICB) in Shanghai. Every second year, members of the department organize the International Otto Warburg Summer School that have now become an integral part of the IMPRS-CBSC and are alternately held in Berlin and in Shanghai. These summer schools bring together international lecturers with a selected group of national

and international students and combine lecture-style teaching and research seminars to give an overview of new important areas in computational biology.

During the reporting period, members of the department have been involved in a number of national and international projects and collaborations. On a national level, the department is part of the Collaborative Research Center for Theoretical Biology (Sonderforschungsbereich, SFB 618), the Transregional Collaborative Research Center "Innate Immunity of the Lung" (SFB-TR 84), and of a DFG funded graduate school, the Research Training Group "Computational Systems Biology" (DFG-Graduiertenkolleg 1772). From amongst the BMBF-funded projects the department participates in the German Epigenome Project DEEP and in a consortium CancerEpiSys, which studies epigenetic effects in cancer. On an international level, the department participates in the EU projects EURATRANS and BLUEPRINT.

Planned developments

Gene regulation and evolution are still the overarching questions of the department. With the increasing realization that there are also epigenetic aspects to gene regulation, we are now trying to understand both the individual levels of regulation as well as their interplay. All this is, of course, reflected in the architecture of the genome, which we are trying to understand from an evolutionary viewpoint as well as from a mechanistic viewpoint. This work relies on the plethora of data that is available these days – and which one needs to know how to analyze. Thus, much effort goes into finding or developing the best methods to extract answers to our biological questions from the genomic and functional data. At the same time, we are intensifying the collaborations between theoreticians and experimentalists in order to continuously test our hypotheses, and to come up with experimental procedures that produce the most information gain.

Transcriptional Regulation Group

(Established: 10/2000)

Head
Prof. Dr. Martin Vingron
Phone: +49 (0)30 8413-1150
Fax: +49 (0)30 8413-1152
Email: vingron@molgen.mpg.de

Scientists and Postdocs
Guofeng Meng (since 09/11)
Navodit Misra (since 01/11)
Matthias Heinig* (since 10/10)
Annalisa Marsico* (since 09/10)
Morgane Thomas-Chollier* (since 04/09)
Julia Lasserre* (since 04/08)
Roman Brinzanik (since 01/07)
Tomasz Zemojtel (since 03/04)
Roland Krause (01/05-05/12)
Ho-Ryun Chung (01/03-12/11)
Szymon Kielbasa* (03/05-06/11)
Christine Steinhoff** (08/05-02/11)
Anirban Banerjee* (11/07-10/10)
Sarah Behrens* (10/08-06/10)
Andrew Hufton* (11/06-01/10)

PhD students
Mahsa Ghanbari* (since 10/11)
Juliane Perner* (since 12/10, IMPRS)
Alena Mysickova (since 10/09, IMPRS)
Jonathan Göke (09/07-09/12, IMPRS)
Ewa Szczurek* (10/06-06/12, IMPRS)
Shengyu Ni (09/10-04/12)
Marta Luksza (10/09-06/11)
Akdes Serin (08/06-10/11, IMPRS)
Rosa Karlic (10/07-12/11, IMPRS)

* externally funded

Scientific overview

The field of transcriptional regulation has gone through a rapid development over the last couple of years. This is due to the plethora of whole-genome sequence data and the functional genomics data on gene expression and on DNA-binding proteins which have become available. The group works on exploiting this data for the purpose of gaining a better understanding of transcriptional regulation in eukaryotes. The main questions are the identification of the sequence motifs in promoters and enhancers, the interplay beween epigenetic marks and regulation, and the evolution of regulatory elements. Construction of gene regulatory networks is the ultimate goal to which all this information shall eventually contribute.

Enhancers & combinatorial regulation

(Jonathan Göke, Alena Mysickova)

The advent of the new sequencing technologies has revolutionized the study of transcription factor binding and of association of other proteins with DNA. These investigations are generally based on Chromatin Immunoprecipitation (ChIP) which yields the DNA elements to which a factor binds. The determination of the binding sequence or location used to be done through hybridization on arrays (ChIP-chip). This, however, implies that one has to know the candidate sequences – typically collections of promoters – or use tiling arrays, which are hard to produce with sufficient density. This is where new sequencing technologies came in by allowing the direct determination of the sequence that was associated to the binding protein. By subsequent mapping of the sequence to the genome, the binding locations can be determined. This routine is called ChIP-seq and has become the workhorse of DNA-binding and chromatin structure studies. Large amounts ChIP-seq data for many transcription factors and chromatin marks in many cell lines are now publicly available.

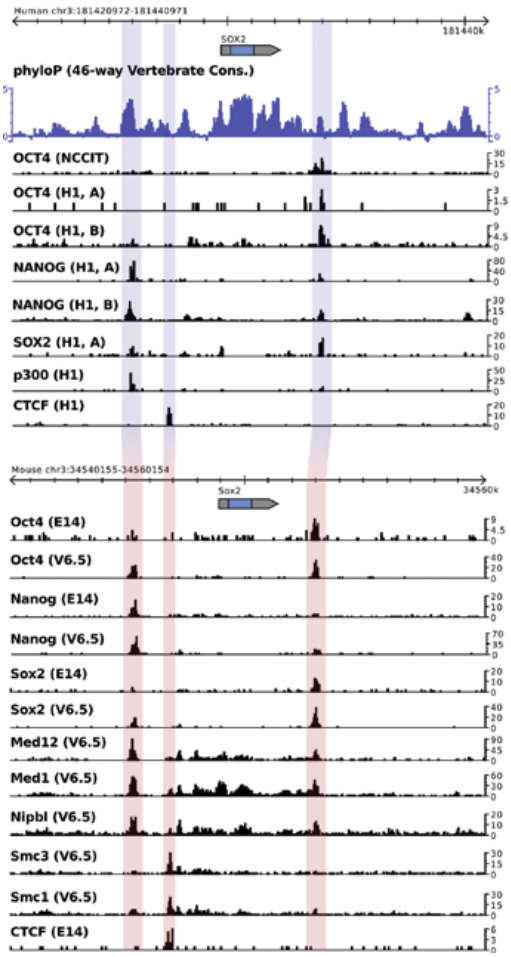
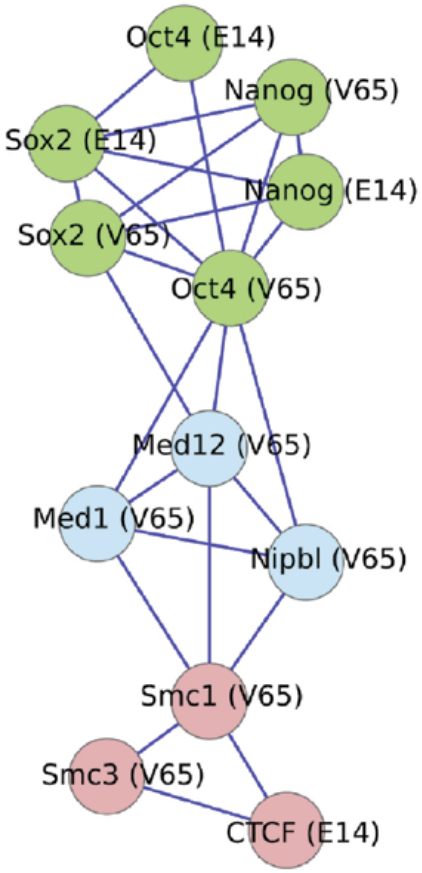


Figure 1: a) Public domain ChIP-seq data on binding of a number of transcription factors (OCT4, SOX2, NANOG) and other proteins interacting with DNA has been integrated into a coherent picture, allowing the delineation of likely enhancer regions. Genomic regions from man and mouse are aligned bases on a whole genome alignment. The picture shows tracks of binding along homologous genomic regions, which includes the SOX2 gene. The bumps in the tracks indicate where the respective protein binds. Enhancers are characterized by the simultaneous binding of the studied proteins, excluding the bottom three (two cohesin subunits and CTCF) which are probably involved in the three-dimensional folding of the chromosome. b) Based on the ChIP-seq data from a), factors that frequently bind to the same regions have been identified and, in the plot, connected by edges. The clusters suggest that these proteins interact with others in order to exert their regulatory role.



In the group, we have exploited these data for a number of questions. In collaboration with James Adjaye (Dept. of Vertebrate Genomics) we analyzed binding data for the stem cell transcription factors OCT4, SOX2, and NANOG, together with histone marks and footprints for the mediator complex and the CTCF insulator protein. While one does see binding of the transcription factors to promoters, there are a large number of distal regions where all three factors bind, apparently denoting enhancer regions. Thus, the ChIP-seq technology yields extensive information on tissue- or condition-specific enhancer elements. Jonathan Göke showed that the enhancer regions are characterized by the combinatorial binding of the transcription factors and that these combinatorial binding events tend to be more highly conserved in evolution than binding sites of individual transcription factors.
The essence of enhancer regions seems to be their contents in transcription factor binding sites. In his PhD thesis, Jonathan Göke has devised a new comparison method that rests not on the order of a sequence but on the contents of short DNA words. This new similarity measure falls under the class of alignment-free similarity measures and is based on an inner product of occurrence-counts of words. The key idea in the measure is that not only the identity of two words is accounted for, but also a possible similarity between words. Based on the new measure Göke has shown that one can cluster enhancer regions, classify enhancers for particular tissues, and detect new motifs in groups of enhancers.
Another aspect that concerns regulatory motifs both in promoters and enhancers is the combinatorial interaction between transcription factors. Here, Alena Mysickova developed a new method to detect pairs of transcription factors whose binding sites co-occur more frequently than expected. While this has been a research topic in the group for a long time, an advance now came through focussing on tissue-specific promoters. This in turn led to a new mathematical question which we formalized by introducing three-dimensional contingency tables. These then allow testing for significant pairs of transcription factors in particular tissues.

TRAP Web Tools : a portal for transcription factor binding predictions

(Morgane Thomas-Chollier, Andrew Hufton, Matthias Heinig, Sean O'Keeffe, Annalisa Marsico)

Between 2006 and 2009 we developed the TRAP method for Transcription Factor Affinity Prediction that calculates the affinity of transcription factors for DNA sequences on the basis of a biophysical model. This method has proven to be useful for several applications besides determining the putative target genes of a given factor. An accurate p-value computation for TRAP scores allows to determine which factors are the most likely to regulate a given target gene. This has also opened the way to the detection of regulatory SNPs and determination of over-represented transcription factor binding sites in sets of related sequences. This ap-

proach is particularly suited for the analysis of genome-wide binding profiles like ChIP-seq data. Morgane Thomas-Chollier has summarized all this in a Web portal to offer easy access to the TRAP method and its derivatives. This work builds upon all the previous efforts of the group, and provides integrated, free and easy access to the various applications of the TRAP method. To help usage of the Web portal and ensure correct interpretation of the results, we provide a protocol focused on the two most popular applications: regulatory SNPs and ChIP-seq data. Significant efforts have been put into extending the statistical TRAP normalization to various model organisms, thereby further expanding its applicability.
The methods have been widely applied, both within our own projects like the combinatorial regulation described above, and various external collaborations.

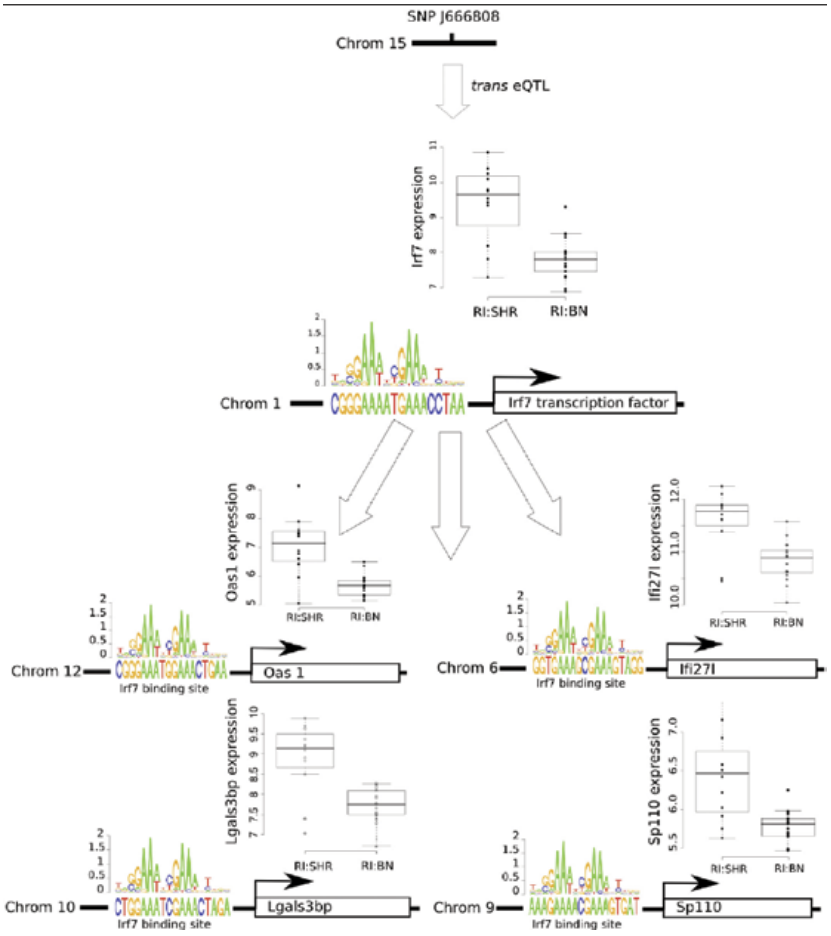


Figure 2: Identification of transcription factor driven gene networks from expression QTL (eQTL) data. Genetic data from rat were scanned for loci that are genetically linked to expression changes (so-called eQTLS) in transcription factors. We then tested for enrichment of transcription factor binding sites (TFBSs) in the putative promoter sequences using our TRAP and PASTAA approach. The strongest enrichment of transcription factor binding sites was observed for interferon regulatory transcription factor/Irf7./Irf7 TFBSs were predicted in the promoters of 23 genes, including Irf7 itself, that all mapped to a single eQTL on rat chromosome 15q25 in adrenal gland, kidney, heart and liver. A subset of the predicted Irf7 targets could be experimentally confirmed within the consortium by chromatin immunoprecipitation and quantitative PCR, which established direct interaction of Irf7 with the promoters of these genes. The boxplots show genotype dependent expression patterns in the heart for selected target genes. The predicted TFBS are shown as alignments of the Irf7 binding logo with the promoter sequence.

We work e.g. in a long-standing cooperation with Norbert Hübner from the Max Delbrück Centre (MDC) in Berlin who is also one of our partners in the EURATRANS network. Hübner is a rat geneticist and together we work on eQTLs in rat. Matthias Heinig, who divides his time between MDC and MPIMG, has applied our bioinformatics methods for gene regulation to explain observed expression changes between rat strains. This work has led to the identification of a network of inflammation-related genes regulated by the transcription factor Irf7, which Heinig originally identified in the rat based on expression-QTLs (see Figure 2). The theoretical prediction of this gene regulatory network suggested a role in Type I diabetes, which was then supported by experimental work in the EURATRANS consortium. This work eventually led to a Nature publication of the consortium with Heinig as first author.
The framework has also been extended to microRNA promoters, with the aim to discover which transcription factors might be involved in regulation of microRNAs. This work is rooted in a DFG SFB-Transregio project between Berlin and Gießen, where our main partner is Bernd Schmeck, formerly Charité and now University of Marburg. In this context we are also studying the role of microRNAs in inflammatory processes.

Epigenetic regulation

(Ho-Ryun Chung, Rosa Karlic, Julia Lasserre)
Transcription factors regulate gene expression in accordance with chromatin structure, histone modifications, DNA methylation patterns, and maybe other players like non-coding RNAs. We investigated several aspects of this network. Firstly, Ho-Ryun Chung studied the sequence dependence of nucleosome positioning. This issue is still under debate, with researchers arguing about the degree to which sequence features determine the location of nucleosomes. In collaboration with experimental partners we showed that published data on sequence preference of nucleosomes can also be explained as a consequence of the cutting preference of the enzyme employed in the experimental procedures. Chung, now heading a Max Planck Research Group at the Otto Warburg Laboratory of the MPIMG, is continuing this line of research and more information can be found in his section of this report.
The second line of research concerns the information encoded by histone modifications. We showed that it is possible to predict expression of a gene from only the histone modifications in its promoter region. This is not to say that histone modifications directly influence the level of transcription, but the interpretation of this mathematical relationship is rather that histone modifications reflect the transcriptional state of a promoter. In this context we could also show that the most informative modification patterns differ between those promoters that have a high content of CpG dinucleotides *vs*. those with a low CpG contents. In ongoing work, we are examining the correlation structure among histone modifications in promoters and enhancers. The work on histone modification now continues in close collaboration with the group of Ho-Ryun Chung.

Evolution of regulation

(Tomasz Zemojtel, Sarah Behrens, Morgane Thomas-Chollier)

Genome evolution, the question how the genome came to look the way it does, is one of the fundamental questions in biology. While the evolution of genes and species has been studied in depth, many aspects of genomic evolution are still poorly understood. The group of Peter Arndt concentrates entirely on genomic evolution and the respective work is described in his section. In the Vingron group the focus is on evolution of regulatory DNA and regulatory systems in general. In order to answer the question, how a set of transcription factor binding sites in, say, a promoter region came about, we model promoter sequence evolution by a Markov Model as developed by Arndt. This model aims at correctly modelling not only the transitions between nucleotides, but also describes the mutation of CpG dinucleotides, i.e. the methylation-deamination process. Based on this model, Sarah Behrens could compute the expected time it takes in evolution until a particular binding site occurs by chance in a promoter. While our naïve expectation was that it is very hard to generate a particular DNA motif, the computation shows that indeed it is feasible to assume that nature can play with placing a binding site somewhere in a promoter. In particular, the waiting time among different patterns varies strongly, and the ones that show up sooner are indeed the ones that are observed more frequently.

In another study, Tomasz Zemojtel found that transposable elements may harbour binding sites. This led to the hypothesis that transposable elements upon moving around in the genome can carry binding sites to different places and put genes under the control of a new transcription factor. Detailed searches among vertebrate genomes have then uncovered cases where indeed p53 binding sites are located within Alu-repeat elements in promoter regions. In fact, recently Zemojtel re-inspected the original data describing the p53 binding sites and found that, unknowingly, the authors had used sites that were contained in Alu repeats. A related study of Zemojtel's shows the general influence of the CpG-mutation-deaminiation process in creating new transcription factor binding sites.

Gene networks

(Ewa Szczurek, Navodit Misra, Mahsa Ghanbari, Julia Lasserre)

With many of the efforts of the research group being directed at the detection of interactions between transcription factors and target genes, or between histone modifications and target gene expression, it is only natural to ask for the possibility to put all this information together into large gene networks. While we go about this task with a lot of respect for the intrinsic difficulties, we are still pursuing a few avenues which we believe lead towards larger networks. On the theoretical side, we are exploring Graphical Gaussian Networks which rely on estimating partial correlations. We are applying this to study the relationships among histones (Julia Lasserre). For network inference from gene expression data, we are exploring ways to upgrade the very tentative models with additional biological information (Mahsa Ghanbari). Bayesian Networks are also being used to model the co-occurrence of mutations in tumours (Navodit Misra). In another cancer-related study, Ewa Szczurek identified changes in regulatory networks which occur in tumours. Currently, she attempts to identify pairs of genes, which are rarely found to mutate jointly in a tumour. Such pairs might be indicators of "weak spots" in a tumour because the viability of the tumour appears to be decreased when the genes are both mutated.

Selected publications

Göke J, Schulz MH, Lasserre J, Vingron M (2012). *Estimation of pairwise sequence similarity of mammalian enhancers with word neighbourhood counts*. Bioinformatics 28(5):656-63

Myšicková A, Vingron M (2012). *Detection of interacting transcription factors in human tissues using predicted DNA binding affinity*. BMC Genomics 13 Suppl 1:S2

Göke J, Jung M, Behrens S, Chavez L, O'Keeffe S, Timmermann B, Lehrach H, Adjaye J, Vingron M (2011). *Combinatorial binding in human and mouse embyonic stem cells identifies conserved enhancers active in early embryonic development*. PLoS Comput Biol 7(12):e1002304

Thomas-Chollier M, Hufton A, Heinig M, O'Keeffe S, Masri NE, Roider HG, Manke T, Vingron M (2011). *Transcription factor binding predictions using TRAP for the analysis of ChIP-seq data and regulatory SNPs.* Nat Protoc 6(12):1860-9

Chung HR, Dunkel I, Heise F, Linke C, Krobitsch S, Ehrenhofer-Murray AE, Sperling SR, Vingron M (2010). *The effect of micrococcal nuclease digestion on nucleosome positioning data*. PLoS One 5(12):e15754.

Heinig M, Petretto E, Wallace C, Bottolo L, Rotival M, Lu H, Li Y, Sarwar R, Langley SR, Bauerfeind A, Hummel O, Lee YA, Paskas S, Rintisch C, Saar K, Cooper J, Buchan R, Gray EE, Cyster JG; Cardiogenics Consortium, Erdmann J, Hengstenberg C, Maouche S, Ouwehand WH, Rice CM, Samani NJ, Schunkert H, Goodall AH, Schulz H, Roider HG, Vingron M, Blankenberg S, Münzel T, Zeller T, Szymczak S, Ziegler A, Tiret L, Smyth DJ, Pravenec M, Aitman TJ, Cambien F, Clayton D, Todd JA, Hubner N, Cook SA (2010). *A trans-acting locus regulates an anti-viral expression network and type 1 diabetes risk.* Nature 467(7314):460-4

Karlic R, Chung HR, Lasserre J, Vlahovicek K, Vingron M (2010). *Histone modification levels are predictive for gene expression.* Proc Natl Acad Sci U S A 107(7):2926-2931

Gene Structure & Array Design Group

(Established: 01/2001)

Head
Dr. Stefan Haas (since 01/01)
Phone: +49 (30) 8413-1164
Fax: +49 (30) 8413-1152
Email: haas@molgen.mpg.de

Scientists
Ruping Sun (since 10/09)
Yasmin Aristei (10/10-09/11)
Hugues Richard* (09/06-09/10)

PhD students
Michael Love (since 05/10, IMPRS)
Anne-Katrin Emde
(since 10/08, IMPRS)
Shen Lin (01/09-01/10)
Marcel Schulz (08/06-05/10, IMPRS)

Scientific programmers
Sean O'Keeffe (11/06-12/10)
Ramu Chenna (12/06 – 10/09)

Scientific overview

The main focus of the group is the development of algorithms and tools for the analysis of next-generation sequencing (NGS) data. The recent advances in high-throughput sequencing technologies led to an enormous increase in the amount of data generated. Furthermore, steadily improving and newly emerging technologies require a continuous adaptation of existing software. In many cases dedicated software has to be developed to efficiently handle and analyze the vast amount of sequencing data. While a few processing steps like quality control and mapping are quite independent of the sequencing application (e.g. ChIP-seq, re-sequencing), for each application specific software is needed to address the questions of interest.
Therefore, the group put strong emphasis on setting up an efficient processing infrastructure that allows to cope with sequencing data even for large cohorts of samples in a short time period. As a basic requirement for mutation screening and transcriptome analysis we developed and published comprehensive tool sets for variant detection and transcript expression analysis, respectively. Our processing and data management pipeline is an essential prerequisite to successfully address questions in e.g. cancer genomics or diagnostics where large sample numbers have to be analyzed. All algorithms developed in the group were optimized and validated experimentally in collaboration with the respective laboratories (Ropers, Yaspo). As a result of these tight interactions we published not only new algorithms but also their application to large-scale projects that otherwise could not have been tackled.

* externally funded

Transcriptome sequencing (RNA-seq)

(Marcel Schulz, Hugues Richard, Ruping Sun, Stefan Haas)

Given our early experience with RNA-seq data, the group was among the first to develop algorithms to evaluate transcript expression based on NGS data. Our tools CASI and DASI allow the detection of alternative transcript expression within one cell-type or of differential expression of transcripts across different samples, respectively. In addition, we implemented an EM-based method, POEM, to even quantify expression of alternative transcripts. These predictions were subsequently validated experimentally on samples of HEK and B-cells in collaboration with the group of Marie-Laure Yaspo.

Besides the aspect of studying transcript expression, a main challenge is to assemble transcripts from short read sequences reliably. This task is complicated by the fact that the abundance of reads originating from different transcripts may vary in several orders of magnitude caused by different expression levels. We therefore developed a *de novo* assembly tool (Oases) that efficiently reconstructs transcripts taken estimated expression levels into account.

In a collaborative project with the Max Planck Institute for Neurological Research we recently started to apply state-of-the-art mapping tools as well as *de novo* assembly as a basis for the detection of fusion transcripts in samples of small-cell lung carcinomas. Such artificial transcripts may be prime candidate mutations driving tumour progression as shown for other tumour types.

Detection of disease-causing mutations (Re-sequencing)

(Anne-Katrin Emde, Michael Love, Ruping Sun, Hugues Richard, Sean O'Keeffe, Stefan Haas)

A major application of NGS is the sequencing of entire genomes or genomic regions of interest to determine the specific genotype of an individual. This information can be either used to unravel evolutionary relationships or e.g. to uncover mutations associated with a certain phenotype. In contrast to traditional methods NGS-based re-sequencing allows a comprehensive but also less biased screening for sequence variations (SV) at even lower costs.

In tight collaboration with the group of H.-Hilger Ropers we set up a computational processing pipeline to enable the large-scale analysis of re-sequencing data with the aim to detect potential disease-causing mutations from samples of patients suffering from intellectual disability (ID). Frequently, causal mutations disrupt gene function by changing the protein sequence, or by deletion/duplication of the gene or parts of it. Therefore, the computational pipeline needs to evaluate all the different types of sequence variation. However, depending on the size of the SVs, different computational approaches have to be applied to recover SVs comprehensively. In a first step, we evaluate the basic read mapping alignment for consistent deviations from the reference genome sequence. This strategy allows to determine base exchange variations and short (<=5 bp) insertions/deletions. In order to reduce the number of false-positive variant calls we correct for potential PCR amplification artefacts, and apply a robust quality-based read clipping.

In a second step, we apply our spliced mapping tool, SplazerS, to recover reads that cross boundaries of potential insertion/deletion events by generating artificial paired-end reads. Deviations from the expected distance of such read pairs allows us to predict not only short insertions (<=30 bp) and medium-sized deletions (<50 kb) but also putative retrocopies or pseudogenes.

Finally, we detect large duplication/deletions by evaluating read depth distribution along the genomic region of interest. Significant increase or decrease in read depth indicates potential duplication/deletion events, respectively. In case of e.g. exome enrichment data read depth is usually non-uniform but is rather skewed towards the ends of the enriched region depending on the enrichment technology used. We addressed this issue with our software ExomeCopy, which computes a representative background distribution of read depth against which a sample is compared. This strategy enables the detection of duplication/deletion even across data derived from different enrichment technologies.
On top of the comprehensive set of tools for SV prediction we provide functional annotations for all SVs in order to prioritize SVs according to their potential functional impact. This includes filtering for already known variations that are expected not be associated with diseases, but also annotating known disease-associations extracted from HGMD. In addition, we add information about sequence conservation, impact on protein sequence or splicing in order to further prioritize candidate mutations. All SVs detected, including detailed functional annotations, are finally stored in a database allowing querying for distinct regional or functional subsets of SVs.
Besides the development of software infrastructure to recover and annotate sequence variations, the group also applied these tools successfully on patient cohorts for the detection of putative disease-causing mutations. In a first project we analyzed 136 patients with autosomal-recessive ID where a genomic linkage interval was already known from previous studies. Targeted sequencing of these regions revealed 50 genes now newly associated with autosomal-recessive ID.
In a parallel project we analyzed the entire exome of chromosome X of >400 male patients suffering of X-linked ID. After filtering out common variants, our mutation analysis on average yields 3-4 candidate mutations per individual that are subsequently validated experimentally and checked for co-segregation within the family. These results are input for genetic counselling of the parents of the affected children.

Selected publications

Emde AK, Schulz MH, Weese D, Sun R, Vingron M, Kalscheuer VM, Haas SA, Reinert K. (2012). *Detecting genomic indel variants with exact breakpoints in single- and paired-end sequencing data using SplazerS.* Bioinformatics 28(5):619-27

Love, MI, Mysickova A, Sun R, Kalscheuer V, Vingron M; Haas SA (2011). *Modeling Read Counts for CNV Detection in Exome Sequencing Data.* Statistical Applications in Genetics and Molecular Biology 10:1, Article 52

Najmabadi H, Hu H, Garshasbi M, Zemojtel T, Abedini SS, Chen W, Hosseini M, Behjati F, Haas S, Jamali P, Zecha A, Mohseni M, Püttmann L, Vahid LN, Jensen C, Moheb LA, Bienek M, Larti F, Mueller I, Weissmann R, Darvish H, Wrogemann K, Hadavi V, Lipkowitz B, Esmaeeli-Nieh S, Wieczorek D, Kariminejad R, Firouzabadi SG, Cohen M, Fattahi Z, Rost I, Mojahedi F, Hertzberg C, Dehghan A, Rajab A, Banavandi MJ, Hoffer J, Falah M, Musante L, Kalscheuer V, Ullmann R, Kuss AW, Tzschach A, Kahrizi K, Ropers HH. (2011). *Deep sequencing reveals 50 novel genes for recessive cognitive disorders*. Nature 478(7367):57-63

Mechanisms of Transcriptional Regulation Group

(Established: 09/2009)

Head
Dr. Sebastiaan H. Meijsing
(since 09/09)
Phone: +49 (0)30 8413-1176
Fax: +49 (0)30 8413-1152
Email: meijsing@molgen.mpg.de

Scientist
Sergey Prykhozhij (10/10-05/12)

PhD students
Stefanie Schöne (since 10/11)
Stephan Starick (since 09/10)
Jonas Telorac (since 06/10)

Technicians
Katja Borzym (since 08/10)
Edda Einfeldt (since 09/09)

Scientific overview

The long-term goal of our group is to understand how a single transcription factor can regulate vastly different sets of genes depending on the cell type and to identify and study processes that influence the expression level of individual genes. We study transcriptional regulation using the glucocorticoid receptor (GR), a member of the steroid hormone receptor family. Upon hormone stimulation, GR binds to specific DNA sequences to regulate the expression of target genes. Although GR is expressed throughout the body, the genes regulated and the genomic loci bound by GR show little overlap between cell-types. Current efforts are aimed to investigate the role of sequence motifs and chromatin in cell-type-specific genomic binding and transcriptional regulation. Further, we study signals involved in fine-tuning expression levels of individual target genes, specifically the role of DNA as a ligand that allosterically modulates the activity of GR.

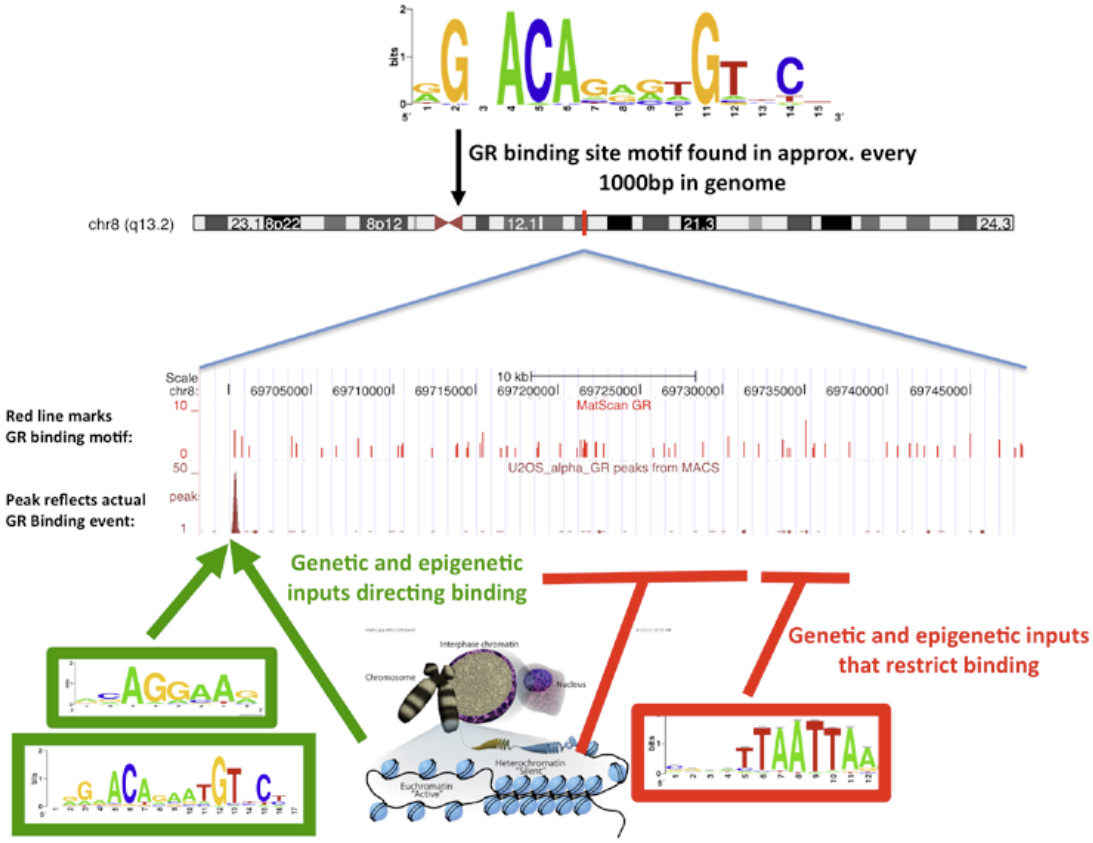


Figure 3: Integration of genetic and epigenetic signals influence where in the genome GR can bind.

Focus 1:
Role of genetic and epigenetic landscape in guiding GR to specific genomic loci

Binding-site motifs of eukaryotic transcription factors typically only have a handful of constrained nucleotide positions. Hence, these sequences are ubiquitously found in the genome whereas binding only occurs at a small subset of these putative genomic binding sites. Consequently, the binding site motif in itself provides insufficient information and a combination of inputs needs to be integrated to specify where transcription factors bind (Figure 3). Our major goal is to identify genetic (DNA sequence elements) and epigenetic (e.g. histone modifications) inputs that specify where GR binds genomically.

Scientific methods

We map the genomic loci of GR binding using ChIP-seq in cell types derived from different tissues including a cell line with a well-characterized epigenome (IMR90: data includes DNA methylation, DNAse-I sensitivity, >20 histone modifications, RNA-seq, www.roadmapepigenomics.org.). To identify genetic and epigenetic features that correlate with genomic binding we employ bioinformatical approaches including linear modeling. The relevance of the identified features is probed in cell lines and animal models including zebrafish. For example by assaying how perturbation of these signals (e.g. knockdown of genes responsible for depositing certain histone modifications) alters the binding profile of GR. Similarly, transcription reporter constructs allow us to test the role of identified DNA sequence motifs in determining where GR binds.

Findings

We have identified genetic and epigenetic features that correlate with GR binding. Interestingly, many of these sequence signals are cell type specific and functional studies using transcriptional reporters indicate that they play a critical role in directing cell-type specific transcriptional regulation by GR.
We also find sequence motifs and epigenetic features that are depleted at sites of GR binding making these features candidates to prevent the binding of GR. Studies using zinc finger nucleases to create isogenic cell lines with genomically integrated reporters and experiments in zebrafish showed that these depleted motifs interfere with GR binding and with GR-dependent activation of transcription (>90% reduction in activation).

Focus 2: DNA as an allosteric modulator of GR structure and activity

DNA guides TFs to defined genomic loci to regulate the expression of genes. The role of DNA binding sites was traditionally thought to be restricted to simply recruiting TFs. Recent studies however revealed that DNA sequences induce alternative conformations in the associated TFs. Furthermore, DNA binding site sequence variants show different coactivator requirements for transcriptional activation, indicating that different regulatory protein complexes are assembled depending on the sequence of the DNA bound (Figure 4). Our aim is to understand the role of DNA as an allosteric modulator of transcription factor activity.

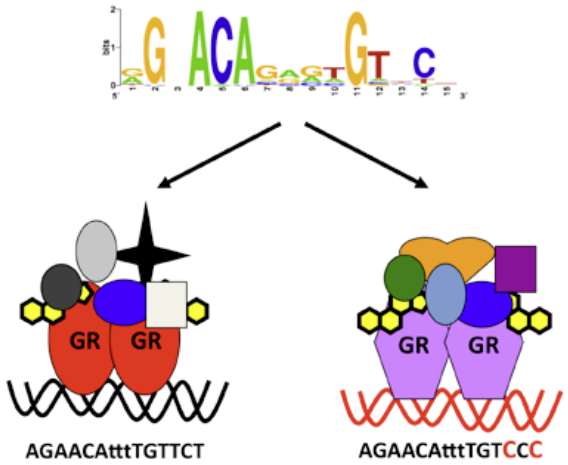


Figure 4: Binding site sequence variants nucleate the assembly of distinct transcriptional regulatory protein complexes.

Scientific methods

We use a combination of cell biology, molecular biology, biochemistry, genetics and structural biology to understand how DNA sequences influence GR signaling. One approach is to study how perturbation of domains that change conformation depending on the DNA sequence bound influences GR activity, communication between GR domains and interactions with coregulators. Other approaches are aimed to identify binding site specific coregulators of GR using DNA pull down assays combined with mass spectrometry (with David Meierhofer) to compare the regulatory complex composition of DNA-sequence variants. In parallel we study how individual DNA sequences direct GR activity in a genomic context by targeting transcriptional reporters to defined genomic loci using Zinc Finger Nucleases.

Findings

Recent studies have elucidated how structural changes at the DNA-protein interface may be transmitted to other parts of GR to elicit context-dependent activities. In collaboration with Ulrich Stelzl (OWL), we identified candidate DNA sequence-specific coregulators of GR. Moreover, we have generated isogenic panels of cell lines with integrated reporters that differ in the exact sequence of the GR binding site. These studies have shown that sequence variants indeed induce different transcriptional responses by GR in a genomic context and that higher activity does appear to require higher levels of GR binding, consistent with our working hypothesis that sequence variants direct the assembly of distinct transcriptional regulatory complexes by GR.

Selected publications

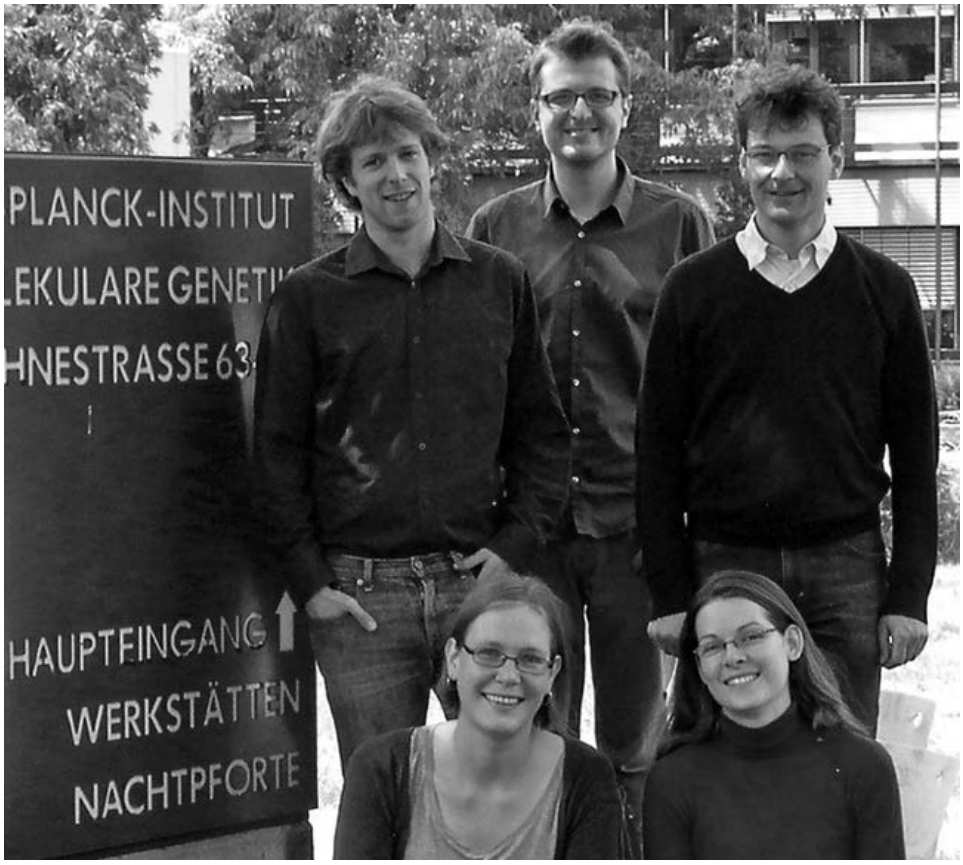
Ziv L, Muto A, Schoonheim PJ, Meijsing SH, Strasser D, Ingraham HA, Schaaf MJM, Yamamoto KR, Baier H. (2012) *An affective disorder in zebrafish with mutation of the glucocorticoid receptor*. Molecular Psychiatry, in press.

DeKelver RC, Choi VM, Moehle EA, Paschon DE, Hockemeyer D, Meijsing SH, Sancak Y, Cui X, Steine EJ, Miller JC, Tam P, Bartsevich VV, Meng X, Rupniewski I, Gopalan SM, Sun HC, Pitz KJ, Rock JM, Zhang L, Davis GD, Rebar EJ, Cheeseman IM, Yamamoto KR, Sabatini DM, Jaenisch R, Gregory PD, Urnov FD (2010). *Functional genomics, proteomics, and regulatory DNA analysis in isogenic settings using zinc finger nuclease-driven transgenesis into a safe harbor locus in the human genome.* Genome Research 20:1133-1142

Meijsing SH, Pufall MA, So AY, Bates DL, Chen L, Yamamoto KR (2009). *DNA binding site sequence directs glucocorticoid receptor structure and activity.* Science 17: 407-410

Research Group Evolutionary Genomics

(Established: 10/2003)



Head
Dr. Peter Arndt (since 10/03)
Phone: +49 (30) 8413-1162
Fax: +49 (30) 8413-1152
Email: arndt@molgen.mpg.de

Scientists
Irina Czogiel (since 10/11)
Brian Cusack (06/08–12/11)

PhD students
Barbara Wilhelm
(since 09/09, IMPRS)
Yves Clement (since 09/08, IMPRS)
Carina Mugal* (07/11–12/11)
Paz Polak (10/06–03/11, IMPRS)
Federico Squartini
(09/05–01/10, IMPRS)

Students
Florian Massip (since 04/12)

Scientific overview

Unravelling the evolutionary forces responsible for variation in genomes within one species or divergence between two species is a major scientific challenge. Today, genomes of many species and of individuals within species have been sequenced. This gives us the unprecedented opportunity for a quantitative analysis of these data with respect to evolutionary aspects. Due to advances in next generation sequencing technologies and the availability of public databases this analysis is possible with more power and precision than before.

* externally funded

We use data on variation within one genome and comparative genomics to learn more about the processes that shape the genome of humans and other species. We investigate processes on short length scales, e.g. nucleotide substitutions, insertions and deletions and long length scales, e.g. insertions of repetitive elements and duplications. Our analyses are complemented by studies of the mathematical underpinnings of models for nucleotide substitutions and phylogeny as well as experimental approaches to study selection *in vitro*.

Comparative analysis of nucleotide substitutions

(Yves Clement, Paz Polak, Peter Arndt)

Mammalian genomes show large-scale regional variations of GC-content (the isochors), but the substitution processes responsible for this structure are poorly understood. We have shown that meiotic recombination has a major impact on substitution patterns in humans and mice driving the evolution of GC-content. Furthermore, other cellular processes also have an influence on nucleotide substitutions on a more local scale.

A regional analysis of nucleotide substitution rates along human genes and their flanking regions allowed us to quantify the effect of mutational mechanisms associated with transcription in germ line cells. Our results revealed three distinct patterns of substitution rates. First, a sharp decline in the deamination rate of methylated CpG dinucleotides, which is observed in the vicinity of the 5' end of genes. Second, a strand asymmetry in complementary substitution rates, which extends from the 5' end to 1 kbp downstream from the 3' end, associated with transcription-coupled repair. Finally, a localized strand asymmetry, i.e. an excess of C->T over G->A substitution in the non-template strand confined to the first 1-2 kbp downstream of the 5' end of genes.

Mathematics of evolutionary models

(Barbara Wilhelm, Federico Squartini, Peter Arndt)

Markov models describing the evolution of the nucleotide substitution process are widely used in phylogeny reconstruction. They usually assume the stationarity and time reversibility of this process. Although corresponding models give meaningful results when applied to biological data, it is not clear if the two assumptions hold and, if not, how much sequence evolution processes deviate from them. To this end, we introduced two sets of indices to quantify violations of the above two assumptions using the Kolmogorov cycle conditions for time reversibility.

In the future we try to answer questions about the limitations of parameter estimations in comparative genomics. Especially we want to explore whether the addition of more species improves the estimation of nucleotide substitution rates along a given branch in a phylogeny.

Models of genome evolution

(Florian Massip, Peter Arndt)

In the recent past it has become clear that besides nucleotide substitutions also the insertion and deletion of short pieces of DNA as well as the insertion of repetitive elements have a substantial influence on the evolution of GC isochors in mammals. We found that in the case where insertions happen to be segmental duplications of adjacent sequences, this process is able to generate correlations of

the GC-content that fall off like a power law. We have shown that simple expansion randomization systems (ERS) are able to generate long-range correlation of the GC content, which is one of the hallmarks of isochors. A wide range of such ERSs fall within one universality class and the characteristic decay exponent of the correlation function can easily be calculated from the rates of the underlying processes. This result gives us also a simple method to simulate long-range correlated sequences and recently we were able to quantify the influence of such correlations on the alignment statistics of sequence, which turned out to be quite substantial.
Currently we are working on models that describe the evolution of segmental duplications of neutral genomic sequences. These duplications are thought to have no direct function and therefore dissolve into the genomic background by random mutations. However this process carries some fascinating statistical properties, which we are analyzing at the moment.

Spatial and temporal dynamics of the immune repertoire in mice

(Irina Czogiel, Peter Arndt)

In a joint project with immunologists from the Max Planck Institute for Infection Biology (Hedda Wardemann, Christian Busse) we try to accurately quantify the overall size, clonality, and histoanatomical distribution of the immune globulin gene repertoire in mice. Our wet-lab partners have developed a novel experimental approach for acquiring the necessary data that moves away from sequencing bulk isolated B cells. Instead, single B cells will be isolated from different histoanatomical locations (i.e. from all lymphoid tissues and several non-lymphoid tissues) so that the acquired dataset will contain information of previously inaccessible detail. For the first time, we will be able to quantify the diversity of the Ig gene repertoire on a monoclonal level. Moreover, we will develop a model for the underlying evolutionary phylodynamics of the B cell populations that will increase our understanding of the selection processes that constantly shape the antibody repertoire during B cell development and differentiation.

In vitro selection

(Barbara Wilhelm, Peter Arndt)

The advancements of next generation sequencing technologies give us a novel tool for the quantitative analysis of Systematic Evolution of Ligands by Exponential Enrichment (SELEX) experiments. Such experiments are conducted in close collaboration with the Glökler group (Dept. Lehrach). Starting from a highly diverse pool of DNA sequences, ligands to particular molecules, e.g. transcription factors or other molecules, relevant to cellular processes, are enriched through subsequent rounds of selection. In house sequencing capabilities give us the opportunity to sequence the DNA pools after each round of selection. This way we are going to study the dynamics of selection for strong binding ligands in lieu of a highly diverse background of unspecific ligands. Since very high diversities can be charted using Illumina sequencing we will also be able to study non-dominant secondary clones and follow the dynamics of their frequency in the population during rounds of selections. New approaches to cluster and analyze the clonal structure of synthetic sequence pools have to be developed.

Phenotypic mutations

(Brian Cusack, Peter Arndt)

Recent studies have hinted at the importance of 'phenotypic mutations' (errors made in transcription and translation) in molecular evolution. These are thought to facilitate positive selection for adaptations that require multiple-substitutions but the generality of this phenomenon has yet to be explored. Our research in this area focuses on the importance of phenotypic mutations to negative selection and to the maintenance of genomic robustness by selective constraint. We initially approached this in the context of Nonsense Mediated Decay (NMD)-based surveillance of human gene transcription. We have discovered a pattern of codon usage in human genes that compensates for the variable NMD efficiency by minimizing nonsense errors during transcription. Our future work will focus on whether phenotypic mutations due to other types of mis-transcription constitute a similar selective force.

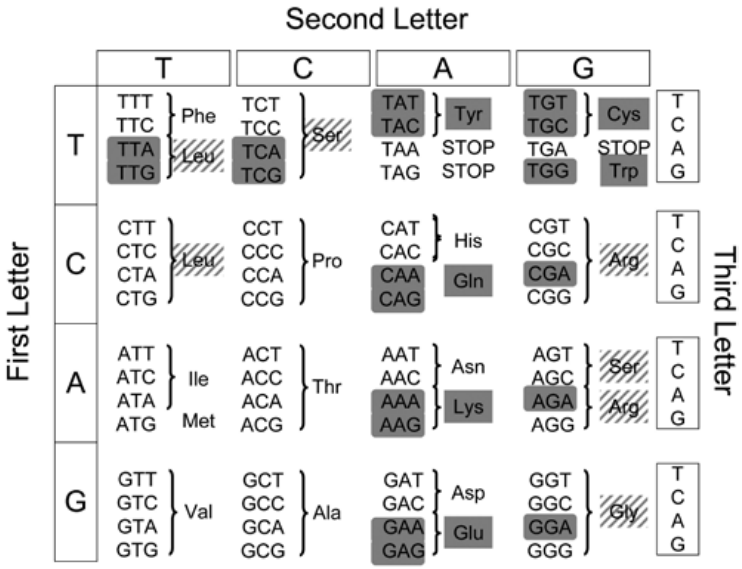


Figure 5: Sense codons differ in their propensity for conversion to STOP codons. The Standard Genetic Code contains 18 fragile codons (shaded) that can be changed into a STOP codon by a single point-mutation and whose mistranscription can therefore generate nonsense errors. The remaining 43 sense codons are "robust" to such errors. Six amino acids are encoded exclusively by fragile codons ("fragile amino acids", shaded), ten amino acids are encoded exclusively by robust codons ("robust amino acids", unshaded) and four amino acids can be encoded either by robust or fragile codons ("facultative amino acids", hatched shading).

General information

Complete list of publications (2009-2012)

2011

Clement Y, Arndt PF (2011). *Substitution Patterns Are Under Different Influences in Primates and Rodents.* Genome Biol Evol 3:236-45

Cusack BP, Arndt PF, Duret L, Roest Crollius H (2011). *Preventing dangerous nonsense: selection for robustness to transcriptional error in human genes*. PLoS Genet 7(10):e1002276

Schütze T, Wilhelm B, Greiner N, Braun H, Peter F, Mörl M, Erdmann, VA, Lehrach H, Konthur Z, Menger M, Arndt PF, Glokler J (2011). *Probing the SELEX process with next-generation sequencing*. PloS one 6(12):e29604

Zemojtel T, Kielbasa SM, Arndt PF, Behrens S, Bourque G, Vingron M (2011). *CpG deamination creates transcription factor binding sites with high efficiency*. Genome Biol Evol 2011, doi: 10.1093/gbe/evr107

2010

Polak P, Querfurth R Arndt PF (2010). *The evolution of transcription-associated biases of mutations across vertebrates*. BMC Evol Biol 10:187

Schütze T, Arndt PF, Menger M, Wochner A, Vingron M, Erdmann VA, Lehrach H, Kaps C, Glökler J (2010). *A calibrated diversity assay for nucleic acid libraries using DiStRO--a Diversity Standard of Random Oligonucleotides*. Nucleic Acids Res 38(4):e23

2009

Polak P, Arndt PF (2009). *Long-range bidirectional strand asymmetries originate at CpG islands in the human genome*. Genome Biol Evol 1:189–197

Singh ND, Arndt PF, Clark AG, Aquadro CF (2009). *Strong evidence for lineage and sequence specificity of substitution rates and patterns in Drosophila*. Mol Biol Evol 26:1591-1605

Zemojtel T, Kielbasa SZ, Arndt PF, Chung HR, Vingron M (2009). *Methylation and deamination of CpGs generate p53-binding sites on a genomic scale.* Trends in Genetics 25(2):63-66

Invited plenary lectures (Peter Arndt)

Breaking Sticks on Evolutionary Time Scales. 3rd International Conference on the Genomic Impact of Eukaryotic Transposable Elements, Asilomar, CA, Feb 25, 2012

Evolution von Genomen. 8. Treffpunkt Bioinformatik, Berlin, Sept 26, 2011

Mutagenic Processes and their Association with Transcription. Colloquium at the Dahlem Centre of Plant Sciences, Berlin, June 10, 2011

Mutagenic Processes and their Association with Transcription. EMBL Conference: Human Variation: Cause & Consequence, Heidelberg, June 6, 2010

Nucleotide Substitution Models - Mathematical Definitions and Genomic Applications. Molecular Evolution Meeting, Orange County, Coorg, India, Nov 30, 2009

Evolutionary Signatures of Mutagenic Processes Associated with Transcription. Fachtagung “Future of Computational Biology” Berlin/Potsdam, Sept 22, 2009

PhD theses

Paz Polak: *Discovering mutational patterns in mammals using comparative genomics.* Freie Universität Berlin, 12/2010

Federico Squartini: *Stationarity and reversibility in the nucleotide evolutionary process.* Freie Universität Berlin, 05/2010

Student thesis

Barbara Wilhelm, nee Keil: *Analyzing In Vitro Selection Experiments using Next Generation Sequencing Technologies*, Diploma Thesis,University of Greifswald, 08/2010

Teaching activities

Single lecture *Population Genetics and Evolutionary Game Theory*, Freie Universität Berlin, WS 2008/09

Lecture on *Population Genetics*, Universite Pierre et Marie Curie, Paris, France, October/November 2010

Lecture *Dynamical Models Describing Genomic Nucleotide Substitutions*, OIST Summer School on Quantitative Evolutionary and Comparative Genomics, Okinawa, Japan, May/June 2010

General information about the whole Department

Complete list of publications (2009-2012)

2012

Adams D, Altucci L, Antonarakis SE, Ballesteros J, Beck S, Bird A, Bock C, Boehm B, Campo E, Caricasole A, Dahl F, Dermitzakis ET, Enver T, Esteller M, Estivill X, Ferguson-Smith A, Fitzgibbon J, Flicek P, Giehl C, Graf T, Grosveld F, Guigo R, Gut I, Helin K, Jarvius J, Küppers R, Lehrach H, Lengauer T, Lernmark Å, Leslie D, Loeffler M, Macintyre E, Mai A, Martens JH, Minucci S, Ouwehand WH, Pelicci PG, Pendeville H, Porse B, Rakyan V, Reik W, Schrappe M, Schübeler D, Seifert M, Siebert R, Simmons D, Soranzo N, Spicuglia S, Stratton M, Stunnenberg HG, Tanay A, Torrents D, Valencia A, Vellenga E, Vingron M, Walter J, Willcocks S (2012). *BLUEPRINT to decode the epigenetic signature written in blood*. Nat Biotechnol 30(3):224-6. doi: 10.1038/nbt.2153

Banerjee A (2012). *Structural distance and evolutionary relationship of networks*. Biosystems 107(3):186-96

Emde AK, Schulz MH, Weese D, Sun R, Vingron M, Kalscheuer VM, Haas SA, Reinert K (2012). *Detecting genomic indel variants with exact breakpoints in single- and paired-end sequencing data using SplazerS*. Bioinformatics 28(5):619-627

Göke J, Schulz MH, Lasserre J, Vingron M (2012). *Estimation of pairwise sequence similarity of mammalian enhancers with word neighbourhood counts*. Bioinformatics 28(5):656-63

Heise F, Chung HR, Weber JM, Xu Z, Klein-Hitpass L, Steinmetz LM, Vingron M, Ehrenhofer-Murray AE (2012). *Genome-wide H4 K16 acetylation by SAS-I is deposited independently of transcription and histone exchange.* Nucl Acids Res 40:65-74

Huppke P, Brendel C, Kalscheuer V, Korenke GC, Marquardt I, Freisinger P, Christodoulou J, Pitelet G, Wilson C, Gruber-Sedlmayr U, Ullmann R, Haas S, Elpeleg O, Nürnberg U, Nürnberg P, Dad S, Moller LB, Kaler SG, Gärtner J (2012). *Mutations in SLC33A1 cause a lethal autosomal recessive disorder with congenital cataracts and hearing loss associated with low serum copper and ceruloplasmin*. Am J Hum Genet 90(1):61-68

Myšicková A, Vingron M (2012). *Detection of interacting transcription factors in human tissues using predicted DNA binding affinity*. BMC Genomics 13 Suppl 1:S2

Schulz MH, Zerbino DR, Vingron M, Birney E (2012). *Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels*. Bioinformatics 28(8):1086-92

Sun R, Love MI, Zemojtel T, Emde AK, Chung HR, Vingron M, Haas SA (2012). *Breakpointer: using local mapping artifacts to support sequence breakpoint discovery from single-end reads.* Bioinformatics 28(7):1024-5

Zemojtel T, Vingron M (2012). *P53 binding sites in transposons*. Front Genet 3:40

2011

Clement Y, Arndt PF (2011). *Substitution Patterns Are Under Different Influences in Primates and Rodents.* Genome Biol Evol 3:236-45

Cusack BP, Arndt PF, Duret L, Roest Crollius H (2011). *Preventing dangerous nonsense: selection for robustness to transcriptional error in human genes*. PLoS Genet 7(10):e1002276

Göke J, Jung M, Behrens S, Chavez L, O'Keeffe S, Timmermann B, Lehrach H, Adjaye J, Vingron M (2012). *Combinatorial binding in human and mouse embyonic stem cells identifies conserved enhancers active in early embryonic development*. PLoS Comput Biol 7(12):e1002304

Hafemeister C, Krause R , Schliep A (2011). *Selecting Oligonucleotide Probes for Whole-Genome Tiling Arrays with a Cross-Hybridization Potential*. IEEE/ACM Trans Comput Biol Bioinform 2011 Feb 24

Hallen L, Klein H, Stoschek C, Wehrmeyer S, Nonhoff U, Ralser M, Wilde J, Röhr C, Schweiger MR, Zatloukal K, Vingron M, Lehrach H, Konthur Z, Krobitsch S (2011). *The KRAB-containing zinc-finger transcriptional regulator ZBRK1 activates SCA2 gene transcription through direct interaction with its gene product, ataxin-2*. Hum Mol Genet 20(1):104-14

Homilius M, Wiedenhoeft J, Thieme S, Standfuß C, Kel I, Krause R (2011). *Cocos: Constructing multi-domain protein phylogenies.* PLoS Curr 3:RRN1240

Kielbasa SM, Wan R, Sato K, Horton P, Frith MC (2011). *Adaptive seeds tame genomic sequence comparison*. Genome Res 21(3):487-93

Kolanczyk M, Pech M, Zemojtel T, Yamamoto H, Mikula I, Calvaruso MA, van den Brand M, Richter R, Fischer B, Ritz A, Kossler N, Thurisch B, Spoerle R, Smeitink J, Kornak U, Chan D, Vingron M, Martasek P, Lightowlers RN, Nijtmans L, Schuelke M, Nierhaus KH, Mundlos S (2011). *NOA1 is an essential GTPase required for mitochondrial protein synthesis.* Mol Biol Cell 22(1):1-11

Lasserre J, Arnold S, Vingron M, Reinke P, Hinrichs C (2011). *Predicting the outcome of renal transplantation*. J Am Med Inform Assoc 2011 Aug 28

Lin S, Haas S, Zemojtel T, Xiao P, Vingron M, Li R (2011). *Genome-wide comparison of cyanobacterial transposable elements, potential genetic diversity indicators*. Gene 473(2):139-49

Love, MI, Mysickova A, Sun R, Kalscheuer V, Vingron M; Haas SA (2011). *Modeling Read Counts for CNV Detection in Exome Sequencing Data.* Statistical Applications in Genetics and Molecular Biology 10(1), Article 52

Najmabadi H, Hu H, Garshasbi M, Zemojtel T, Abedini SS, Chen W, Hosseini M, Behjati F, Haas S, Jamali P, Zecha A, Mohseni M, Püttmann L, Vahid LN, Jensen C, Moheb LA, Bienek M, Larti F, Mueller I, Weissmann R, Darvish H, Wrogemann K, Hadavi V, Lipkowitz B, Esmaeeli-Nieh S, Wieczorek D, Kariminejad R, Firouzabadi SG, Cohen M, Fattahi Z, Rost I, Mojahedi F, Hertzberg C, Dehghan A, Rajab A, Banavandi MJ, Hoffer J, Falah M, Musante L, Kalscheuer V, Ullmann R, Kuss AW, Tzschach A, Kahrizi K, Ropers HH (2011). *Deep sequencing reveals 50 novel genes for recessive cognitive disorders*. Nature 478(7367):57-63

Pfenninger CV, Steinhoff C, Hertwig F, Nuber UA (2011). *Prospectively*

isolated CD133/CD24-positive ependymal cells from the adult spinal cord and lateral ventricle wall differ in their long-term in vitro self-renewal and in vivo gene expression. Glia 59(1):68-81

Schraders M, Haas SA, Weegerink NJ, Oostrik J, Hu H, Hoefsloot LH, Kannan S, Huygen PL, Pennings RJ, Admiraal RJ, Kalscheuer VM, Kunst HP, Kremer H (2011). *Next-Generation Sequencing Identifies Mutations of SMPX, which Encodes the Small Muscle Protein, X-Linked, as a Cause of Progressive Hearing Impairment*. Am J Hum Genet 88(5):628-34

Schütze T, Wilhelm B, Greiner N, Braun H, Peter F, Mörl M, Erdmann, VA, Lehrach H, Konthur Z, Menger M, Arndt PF, Glokler J (2011). *Probing the SELEX process with next-generation sequencing*. PloS one 6(12):e29604

Serin A, Vingron M (2011). *DeBi: Discovering Differentially Expressed Biclusters using a Frequent Itemset Approach*. Algorithms Mol Biol 6(1):18

Szczurek E, Markowetz F, Gat-Viks I, Biecek P, Tiuryn J, Vingron M (2011). *Deregulation upon DNA damage revealed by joint analysis of context-specific perturbation data*. BMC Bioinformatics 12:249

Thomas-Chollier M, Defrance M, Medina-Rivera A, Sand O, Herrmann C, Thieffry D, van Helden J (2011). *RSAT 2011: regulatory sequence analysis tools*. Nucleic Acids Res 39 (Web Server issue):W86-91

Thomas-Chollier M, Herrmann C, Defrance M, Sand O, Thieffry D, van Helden J (2011). *RSAT peak-motifs: motif analysis in full-size ChIP-seq datasets*. Nucleic Acids Res 2011, doi: 10.1093/nar/gkr1104

Thomas-Chollier M, Hufton A, Heinig M, O'Keeffe S, Masri NE, Roider HG, Manke T, Vingron M (2011). *Transcription factor binding predictions using TRAP for the analysis of ChIP-seq data and regulatory SNPs.* Nat Protoc 6(12):1860-9

Wiedenhoeft J, Krause R, Eulenstein O (2011). *The plexus model for the inference of ancestral multidomain proteins*. IEEE/ACM Trans Comput Biol Bioinform 8(4):890-901

Zemojtel T, Kielbasa SM, Arndt PF, Behrens S, Bourque G, Vingron M (2011). *CpG deamination creates transcription factor binding sites with high efficiency*. Genome Biol Evol 2011, doi: 10.1093/gbe/evr107

2010

Behrens S, Vingron M (2010). *Studying the evolution of promoter sequences: a waiting time problem.* J Comput Biol 17(12):1591-606

Cheng X, Guerasimova A, Manke T, Rosenstiel P, Haas S, Warnatz HJ, Querfurth R, Nietfeld W, Vanhecke D, Lehrach H, Yaspo ML, Janitz M (2010). *Screening of human gene promoter activities using transfected-cell arrays*. Gene 450(1-2):48-54

Chung HR, Dunkel I, Heise F, Linke C, Krobitsch S, Ehrenhofer-Murray AE, Sperling SR, Vingron M (2010). *The effect of micrococcal nuclease digestion on nucleosome positioning data*. PLoS One 5(12):e15754

DeKelver RC, Choi VM, Moehle EA, Paschon DE, Hockemeyer D, Meijsing S, Sancak Y, Cui X, Steine EJ, Miller JC, Tam P, Bartsevich VV, Meng X, Rupniewski I, Gopalan SM, Sun HC, Pitz KJ, Rock JM, Zhang L, Davis GD, Rebar EJ, Cheeseman IM, Yamamoto KR, Sabatini DM, Jaenisch R, Gregory PD, Urnov FD (2010). *Functional genomics, pro-*

teomics, and regulatory DNA analysis in isogenic settings using zinc finger nuclease-driven transgenesis into a safe harbor locus in the human genome. Genome Res 20(8): 1133–1142

Emde AK, Grunert M, Weese D, Reinert K, Sperling SR (2010). *MicroRazerS: rapid alignment of small RNA reads.* Bioinformatics 26(1): 123-124

Heinig M, Petretto E, Wallace C, Bottolo L, Rotival M, Lu H, Li Y, Sarwar R, Langley SR, Bauerfeind A, Hummel O, Lee YA, Paskas S, Rintisch C, Saar K, Cooper J, Buchan R, Gray EE, Cyster JG; Cardiogenics Consortium, Erdmann J, Hengstenberg C, Maouche S, Ouwehand WH, Rice CM, Samani NJ, Schunkert H, Goodall AH, Schulz H, Roider HG, Vingron M, Blankenberg S, Münzel T, Zeller T, Szymczak S, Ziegler A, Tiret L, Smyth DJ, Pravenec M, Aitman TJ, Cambien F, Clayton D, Todd JA, Hubner N, Cook SA (2010). *A trans-acting locus regulates an anti-viral expression network and type 1 diabetes risk.* Nature 467(7314):460-4

Karlic R, Chung HR, Lasserre J, Vlahovicek K, Vingron,M (2010). *Histone modification levels are predictive for gene expression.* Proc Natl Acad Sci U S A. 107(7):2926-31

Kielbasa SM, Blüthgen N, Fähling M, Mrowka RN (2010). *Targetfinder.org: a resource for systematic discovery of transcription factor target genes.* Nucleic Acids Res 38 Suppl:W275-80

Kielbasa SM, Klein H, Roider HG, Vingron M, Blüthgen N (2010). *TransFind-predicting transcriptional regulators for gene sets.* Nucleic Acids Res 38 Suppl:W275-80

Kolanczyk M, Pech M, Zemojtel T, Yamamoto H, Mikula I, Calvaruso MA, van den Brand M, Richter R, Fischer B, Ritz A, Kossler N, Thurisch B, Spoerle R, Smeitink J, Kornak U, Chan D, Vingron M, Martasek P, Lightowlers RN, Nijtmans L, Schuelke M, Nierhaus KH, Mundlos S (2010). *NOA1 is an essential GTPase required for mitochondrial protein synthesis*. Mol Biol Cell 2010 Nov 30

Löhr U, Chung HR, Beller M, Jäckle H (2010). *Bicoid: Morphogen function revisited*. Fly (Austin) 4(3)236 - 240

Luksza M, Lässig M, Berg J (2010). *Significance analysis and statistical mechanics: an application to clustering*. Phys Rev Lett 105(22):220601

Manke T, Heinig M, Vingron M (2010). *Quantifying the effect of sequence variation on regulatory interactions*. Hum Mutat 2010 Feb 2

May P, Kreuchwig A, Steinke T, Koch I (2010). *PTGL: A database for secondary structure-based protein topologies*. Nucleic Acids Research 38 (suppl 1) D326 – 30

Meng G, Mosig A, Vingron M (2010). *A computational evaluation of overrepresentation of regulatory motifs in the promoter regions of differentially expressed genes.* BMC Bioinformatics 11:267

Polak P, Querfurth R Arndt PF (2010). *The evolution of transcription-associated biases of mutations across vertebrates.* BMC Evol Biol 10:187

Richard H, Schulz MH, Sultan M, Nürnberger A, Schrinner S, Balzereit D, Dagand E, Rasche A, Lehrach H, Vingron M, Haas SA, Yaspo ML (2010). *Prediction of alternative isoforms from exon expression levels in RNA-Seq experiments*. Nucleic Acids Res 38(10):e112

Rödelsperger C, Guo G, Kolanczyk M, Pletschacher A, Köhler S, Bauer

S, Schulz MH, Robinson PN (2010). *Integrative analysis of genomic, functional and protein interaction data predicts long-range enhancer-target gene interactions*. Nucleic Acids Res 2011; 39(7):2492-502

Schaefer AS, Richter GM, Nothnagel M, Manke T, Dommisch H, Jacobs G, Arlt A, Rosenstiel P, Noack B, Groessner-Schreiber B, Jepsen S, Loos BG, Schreiber S (2010). *A genome-wide association study identifies GLT6D1 as a susceptibility locus for periodontitis*. Hum Mol Genet 19(3):553-62

Schütze T, Arndt PF, Menger M, Wochner A, Vingron M, Erdmann VA, Lehrach H, Kaps C, Glökler J (2010). *A calibrated diversity assay for nucleic acid libraries using DiStRO--a Diversity Standard of Random Oligonucleotides*. Nucleic Acids Res 38(4):e23

Szczurek E, Biecek P, Tiuryn J, Vingron M (2010). *Introducing knowledge into differential expression analysis*. J Comput Biol 17(8):953-67

Warnatz HJ, Querfurth R, Gueramova A, Cheng X, Haas SA, Hufton AL, Manke T, Vanhecke D, Nietfeld W, Vingron M, Janitz M, Lehrach H, Yaspo ML (2010). *Functional analysis and identification of cis-regulatory elements of human chromosome 21 gene promoters*. Nucleic Acids Res 38(18):6112-23

Wiedenhoeft J, Krause R, Eulenstein O (2010). *Inferring Evolutionary Scenarios for Protein Domain Compositions*. Lecture Notes in Computer Science 6053/2010

Zackay A, Steinhoff C (2010). *MethVisual - visualization and exploratory statistical analysis of DNA methylation profiles from bisulfite sequencing.* BMC Res Notes 3:337

2009

Baek YS, Haas S, Hackstein H, Bein G, Santana MH, Lehrach H, Sauer S, Seitz H (2009). *Identification of novel transcriptional regulators involved in macrophage differentiation and activation in U937 cells*. BMC Immunology 10:18

Bozek K, Relogio A, Kielbasa SM, Heine M, Dame C, et al. (2009). *Regulation of Clock-Controlled Genes in Mammals*. PLoS ONE 4(3): e4882

Chavez L, Bais A, Vingron M, Lehrach H, Adjaye J, Herwig R (2009). *In silico identification of a core regulatory network of OCT4 in human embryonic stem cells using an integrated approach*. BMC Genomics 10:314

Chung HR, Vingron M (2009). *Comparison of sequence-dependent tiling array normalization approaches.* BMC Bioinformatics 10:204

Chung HR, Vingron M (2009). *Sequence-dependent Nucleosome Positioning.* J Mol Biol 386(5):1411-22

Diella F, Chabanis S, Luck K, Chica C, Chenna R , Nerlov C, Gibson TJ (2009). *KEPE - a motif frequently superimposed on sumoylation sites in metazoan chromatin proteins and transcription factors.* Bioinformatics 25:1-5

Gambin A, Szczurek E, Dutkowski J, Bakun M, Dadlez M. (2009). *Classification of peptide mass fingerprint data by novel no-regret boosting method.* Comput Biol Med 39(5):460-73

Gat-Viks I, Vingron M (2009). *Evidence for Gene-Specific Rather Than Transcription Rate–Dependent Histone H3 Exchange in Yeast Coding Regions.* PLoS Comput Biol 5(2): e1000282 (2009)

Hu H, Wrogemann K, Kalscheuer V, Tzschach A, Richard H, Haas SA, Menzel C, Bienek M, Froyen G, Raynaud M, Van Bokhoven H, Chelly J, Ropers H, Chen W (2009). *Erratum to: Mutation screening in 86 known X-linked mental retardation genes by droplet-based multiplex PCR and massive parallel sequencing.* Hugo J 3(1-4):83

Huang X, Vingron M (2009). *Maximum Similarity: A New Formulation of Phylogenetic Reconstruction.* J Comp Biol 16(7): 887-896

Hufton AL, Mathia S, Braun H, Georgi U, Lehrach H, Vingron M, Poustka AJ, Panopoulou G (2009). *Deeply conserved chordate noncoding sequences preserve genome synteny but do not drive gene duplicate retention.* Genome Res 19:2036-2051

Kanhere A, Vingron M (2009). *Horizontal Gene Transfers in prokaryotes show differential preferences for metabolic and translational genes.* BMC Evol Biol 9:9

Kielbassa J, Bortfeldt R, Schuster S, Koch I (2009). *Modeling of the U1 snRNP assembly pathway in alternative splicing in human cells using Petri nets*. Comput Biol Chem 33:46-61

Koch I (2009). *Petri nets and GRN Models In Computational Methodologies in Gene Regulatory Networks.* IGI Global PA (USA),chapt. 25:604-637

Koehler S, Schulz MH, Krawitz P, Bauer S, Doelken S, Ott CE, Mundlos C, Horn D, Mundlos S, Robinson PN (2009). *Clinical Diagnostics with Semantic Similarity Searches in Ontologies.* Am J Hum Genet 85(4): 457-464

Loehr U, Chung HR, Beller M, Jaeckle H (2009). *Antagonistic action of Bicoid and the repressor Capicua determines the spatial limits of Drosophila head gene expression domains.* Proc Natl Acad Sci U S A 106(51):21695-21700

Ott CE, Bauer S, Manke T, Ahrens S, Rödelsperger C, Grünhagen J, Kornak U, Duda G, Mundlos S, Robinson PN (2009). *Promiscuous and Depolarization-Induced Immediate-Early Response Genes are Induced by Mechanical Strain of Osteoblasts Original Study.* J Bone Miner Res 24(7):1247-1262

Pape UJ, Klein H, Vingron M (2009). *Statistical detection of cooperative transcription factors with similarity adjustment.* Bioinformatics 25(16):2103-2109

Polak P, Arndt PF (2009). *Long-range bidirectional strand asymmetries originate at CpG islands in the human genome*. Genome Biol Evol 1:189–197

Rausch T, Koren S, Denisov G, Weese D, Emde AK, Döring A, Reinert K (2009). *A consistency-based consensus algorithm for de novo and reference-guided sequence assembly of short reads.* Bioinformatics 25 (9):1118-1124

Rödelsperger C, Koehler S, Schulz MH, Manke T, Bauer S, Robinson PN (2009). *Short ultraconserved promoter regions delineate a class of preferentially expressed alternatively spliced transcripts.* Genomics 94(5):308-316

Roider HG, Lenhard B, Kanhere A, Haas SA, Vingron M (2009). *CpG-depleted promoters harbor tissue-specific transcription factor binding signals—implications for motif overrepresentation analyses.* Nucleic Acids Research 2009:1–11

Roider HG, Manke T, O'Keeffe S, Vingron M, Haas SA (2009). *PASTAA: identifying transcription factors associated with sets of co-regulated genes.* Bioinformatics 25(4):435-442

Roytberg MA, Gambin A, Noe L, Lasota S, Furletova E, Szczurek E, Kucherov G (2009). *On Subset Seeds for Protein Alignment.* IEEE/ACM Trans., Comp Biol Bioinform 6(3):483-494

Schulz MH Koehler S, Bauer S, Vingron M, Robinson PN (2009). *Exact Score Distribution Computation for Similarity Searches in Ontologies.* Proceedings WABI 2009, Springer LNCS, Vol. 5724

Steinhoff C, Paulsen M, Kielbasa S, Walter J, Vingron M (2009). *Expression profile and transcription factor binding site exploration of imprinted genes in human and mouse*. BMC Genomics 10:144

Stritt C, Stern S, Harting K, Manke T, Sinske D, Schwarz H, Vingron M, Nordheim A, Knöll B (2009). *Paracrine control of oligodendrocyte differentiation by SRF-directed neuronal gene expression.* Nat Neurosci 12(4):418-27

Szczurek E, Gat-Viks I, Tiuryn J, Vingron M (2009). *Elucidating regulatory mechanisms downstream of a signaling pathway using informative experiments.* Mol Systems Biol 5:287

Vingron M, Brazma A, Coulson R, van Helden J, Manke T, Palin K, Sand O, Ukkonen E (2009). *Integrating sequence, evolution and functional genomics in regulatory genomics.* Genome Biol 10:202

Weese D, Emde AK, Rausch T, Döring A, Reinert K (2009). *RazerS--fast read mapping with sensitivity control.* Genome Res 19:1646-1654

Ye K, Schulz MH, Long Q, Apweiler R, Ning Z (2009). *Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short read data.* Short-read SIG ISMB 2009, Bioinformatics 25(21):2865-2871

Zemojtel T, Kielbasa SZ, Arndt PF, Chung HR, Vingron M (2009). *Methylation and deamination of CpGs generate p53-binding sites on a genomic scale.* Trends in Genetics 25(2):63-66

Invited plenary lectures (Martin Vingron)

JOBIM (Journées Ouvertes en Biologie, Informatique et Mathématiques) 2012. Rennes, France, 07/2012

Biozentrum, University Basel, Switzerland, 11/2011

10th CRG Symposium, 11/2011

BioRegSIG at ISMB 2011, Vienna, Austria, 07/2011

22nd CPM 2011 Symposium, Palermo, Italy, 06/2011

University of Göttingen, 02/2011

5th annual BBSRC Systems Biology Workshop, London, UK, 01/2011

1. Epigenetics Meeting Freiburg, 12/2010

First RECOMB Satellite Conference on Bioinformatics Education, University of California, San Diego, La Jolla, CA, 03/2009

The Seventh Asia Pacific Bioinformatics Conference (APBC 2009), Beijing, China, 01/2009

Awards and scientific honours

Martin Vingron: *Elected Fellow of the International Society for Computational Biology (ISCB)*, 07/2012

Irina Czogiel: *Gustav-Adolf Lienert Prize*, Biometrical Society, German Region, 03/2012

Marcel Holger Schulz: *Otto Hahn Medal*, Max Planck Society, 06/2011

Rosa Karlic: L'Oreal Adria-UNESCO National Fellowship "For Women in Science", L'Oreal, UNESCO, 2010

PhD theses

Rosa Karlic: *Influence of Histone Modifications on mRNA Abundance and Structure.* Freie Universität Berlin, 12/2011

Akdes Serin: *Biclustering analysis for large scale data*. Freie Universität Berlin, 11/2011

Marta Luksza: *Cluster statistics and gene expression analysis.* Freie Universität Berlin, 08/2011

Ewa Szczurek: *Modeling signal transduction pathways and their transcriptional response.* Freie Universität Berlin 04/2011

Matthias Heinig: *Statistical methods for the analysis of the genetics of gene expression.* Freie Universität Berlin 12/2010

Marcel H. Schulz: *Data structures and algorithms for analysis of alternative splicing with RNA-seq data.* Freie Universität Berlin 08/2010

Holger Klein: *Co-occurrence of transcription factor binding sites.* Freie Universität Berlin, 05/2010

Stefan Bentink: *Transcriptional profiling of aggressive lymphoma*. Freie Universität Berlin, 12/2009

Benjamin Georgi: *Context-specific independence mixture models for cluster analysis of biological data.* Freie Universität Berlin, 06/2009

Student theses

Stephan Knorr: *Identification and Analysis of Repeat Associated Transcription Factor Binding Sites in the Human Genome.* Bachelor Thesis, 10/2011

Daniel Mehnert: *Quality assessment of protein-protein interaction networks.* Bachelor Thesis, 09/2011

John Wiedenhoeft: *Biclustering and Related Methods.* Master Thesis, 09/2011

Jan Patrick Pett: *Identification of putative regulatory conserved elements in coding exons of vertebrate Hox gene clusters.* Bachelor Thesis, 2011

Sandra Kiefer: *Transkriptionelle Regulation durch den Glucocorticoid-Rezeptor.* Bachelor Thesis, 08/2010

Sabrina Krakau: INSEGT Ein Programm für annotationsbasierte Expressionanalyse von RNA-Seq Date. Bachelor Thesis, 10/2009

Arie Zackay: *Visualization and Exploratory statistical Analysis of genome wide DNA Methylation Profiles.* Master Thesis, 08/2009

Rina Ahmed: *Statistical exploration and visualization of epigenetic genome-wide data.* Master Thesis, 03/2009

Teaching activites

The Department of Computational Molecular Biology contributes to teaching in the Bioinformatics curriculum of Free University both at the Bachelors and at the Masters level. Every other year, Vingron teaches Algorithmic Bioinformatics with the help of a staff member who is paid by the university (for a long time this was Roland Krause). The department also takes over one third of a class on *Algorithms in Systems Biology* every year.

In addition, members of the department are involved in teaching for master and bachelor students in *Functional Genomics* and *Epigenetics* and participate in several lecture series at FU and also at Humboldt University. In 2009 and 2011, Roland Krause organized an EMBO World Practical Course on *Computational Biology* in Shanghai, China, and another EMBO Practical Course on *Computational Biology* in Reykjavik, Iceland.

Guest scientists

Dmitri Petrov, Stanford University, 02.-31.07.2012

Alexander Bolshoy, University of Haifa, Israel, 01.09.2011–04.03.2012

Ivan Gesteira Costa Filho, University of Pernambuco, Brazil, 03.–28.01.2011

Matteo Pardo, Italian National Research Council (CNR), 01.06.2008-30.11.2010

Pierre Nicodeme, Laboratoire d'informatique (LIX), École polytechnique, Palaiseau Cedex (near Paris), France, 16.08.2010–03.09.2010

Lin Shen, Chinese Academy of Science, China, 11.01.2009-10.01.2010

Oliver Eulenstein, IOWA State University, Ames, Iowa, USA, 02.01–30.06.2009

Ina Koch, Beuth University of Applied Sciences, Berlin, 2006 – 02/2010

Organization of scientific events

Since 2009, Vingron acts as Chair of the Steering Committee of the *RECOMB conference series*, a renowned annual international conference on computational biology (see http://recomb.org).

Almost every year, the department organizes the *International Otto Warburg Summer School and Research Symposia*, which last about one and a half week and bring together several well-known researchers and PhD students from different backgrounds to discuss recent advances varying fields of computational molecular biology. Regulatory (Epi-) Genomics (2009), Evolutionary Genomics (2011), and Genes, Metabolism and Systems Modelling (2012) have been the themes of the last summer schools. For more details, please see http://ows.molgen.mpg.de/.

Every year, Martin Vingron, together with BioTOP Berlin-Brandenburg, organizes the *Treffpunkt Bioinformatik*, a local, German-speaking workshop with high-ranking speakers that allows students and local actors to meet each other and get an overview about the scientific activities in the Berlin-Brandenburg area on various aspects of computational molecular biology. The last Treffpunkt Bioinformatik have been focused on systems biology (2009), structural biology (2010), evolutionary biology (2011), and RNA technologies (2012).

In addition to these regularly events, members of the department also organized the following national and international workshops.

Workshop on Mathematical and Statistical Aspects of Molecular Biology (MASAMB), Berlin, 04/2012 (organizers: Martin Vingron, Julia Lasserre, Alena Mysickova)

Meeting of Section 2 (Information Sciences) of Leopoldina, Berlin, 02/2012 (organizer: Martin Vingron, together with Martin Grötschel, Zuse Institute Berlin, and Thomas Lengauer, Max Planck Institute for Informatics)

Bioinformatics Summerschool, EU-TRACC Consortium, Bergen, Norway, 07/2009 (co-organizer Christine Steinhoff)

Seminar and lectures given by external speakers in the department

2012

Xuegong Zhang, Tsinghua University, Beijing, China, 25.04.2012. *Studying the Expression of Alternative Splicing Genes with RNA-Seq Data*

Inbal Ipenberg, Technion, Haifa, Israel, 18.04.2012. *Heat Shock Protein 90 is required for the stability of KDM4B histone demethylase*

Thomas Conrad, MPI for Immunobiology and Epigenetics, Freiburg, 16.04.2012. *Dissecting MOF function and the mechanism of X chromosome dosage compensation in Drosophila*

Nina Henriette Uhlenhaut, Max-Delbrück-Centrum für Molekulare Medizin, Berlin-Buch, 06.03.2012. *Rethinking Repression: The Glucocorticoid Receptor at Inflammatory Crossroads*

Gary Stormo, Washington University Medical School, USA, 06.03.2012. *Computational and experimental determinations of Protein-DNA binding specificity*

Soumyadeep Nandi, University of Ottawa, USA, 29.02.2012. *Improved identification of cis-regulatory modules in proximalpromoters of human genes and exploiting the mutual positioning of factors*

Jan Tuckermann, Leibniz Institute for Age Research, 13.02.2012. *Molecular Mechanisms of beneficial and side Effects of Glucocorticoids in inflammatory bone diseases*

2011

Verena Heinrich, Charité, Berlin, 20.12.2011. *The allele distribution in next-generation sequencing data sets is accurately described as the result of a stochastic branching process*

Moritz Gerstung, ETH Zurich, Switherland, 07.12.2011. *Computational Cancer Genomics*

Florian Massip, AgroParisTech, Paris, France, 02.11.2011. *A New Model for Genomic Evolution - The Length Distribution of Segmental Duplications*

Rahul Siddharthan, Institute of Mathematical Sciences, India, 21.10.2011. *Evolution of centromeres in budding yeasts*

Jun Yin, University College Dublin, 10.10.2011. *High throughput transcriptomic analysis of zebrafish eye development*

Benjamin Georgi, University of Pennsylvania, 1.07.2011. *Genomic analysis of bipolar disorder in a genetic isolate*

Stefanie Schöne, Centre of Organismal Studies Heidelberg, 31.05.2011. *Determining the number of stem cells*

in the shoot apical meristem of Arabidopsis thaliana by lineage tracing

Andreas Kowarsch, Helmholtz Zentrum München, 30.05.2011. *Analysis of signaling networks:From miRNA-mediated regulation to temporally responses*

Sarah A. Teichmann, MRC Laboratory of Molecular Biology, Cambridge, UK, 26.04.2011. *A quantitative view of gene expression levels in T helper cells*

Sebastian Klie, MPIMP Golm, 22.04.2011. *Identifiaction of metabolites involved in sensing an signaling in E. coli*

Marc Johannes, DKFZ Heidelberg, 13.04.2011. *Integration of Pathway Knowledge into a Support Vector Framework using Reweighted Recursive Feature Elimination*

Marie Manceau, Harvard University, Boston, MA, USA,16.03.2011. *Formation and Evolution of Color Pattern in Natural Populations*

Jerome Gros, Harvard Medical School, Boston MA, USA,16.03.2011. *The Molecular and Cellular Events Shaping the Vertebrate Embryo*

2010

Ralf Jauch, Genome Institute of Singapore, 20.12.2010. *How proteins understand genomes – the structural biochemistry of transcription factors*

Remo Rohs, University of Southern California, Los Angeles, 07.12.2010. *The role of DNA shape in transcription factor-DNA recognition and nucleosome formation*

Navodit Misra, Carnegie Mellon University, Pittsburgh, USA, 17.11.2010. *Integer programming techniques for phylogeny reconstruction*

Ulf Andersson Örom,The Wistar Institute, Philadelphia, USA, 10.11.2010. *Long non-coding RNA and enhancers*

Julien Gagneur, EMBL Heidelberg, 01.11.2010. *On bidirectional promoters and antisense transcription in yeast*

Dirk J. Evers, Illumina Cambridge Ltd, Little Chesterford, UK, 14.06.2010. *A detailed look at SBS Sequencing Data and its Applications*

Ann Ehrenhofer-Murray, Universität Duisburg-Essen, Zentrum für Medizinische Biotechnologie, 26.02.2010. *Establishment of chromatin domains in the yeast genome*

Nina Stoletzki, University of Sussex, 10.02.2010. *Inferring selection from DNA sequence data - some statistical and biological considerations*

Sergey Prykhozhij, EMBL Heidelberg, 29.01.2010. *In the absence of Sonic Hedgehog, p53 induces apoptosis and inhibits retinal cell proliferation, cell-cycle exit and differentiation in zebrafish*

Peter Serocka, CAS-MPG Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 25.01.2010. *Software Tools for Visual and Statistical Analysis of Multivariate Image Data*

Karen Rusche, University of Leipzig, 12.01.2010. *The transcription factor NSCL-2 - neuronal control of bodyweight and adipose tissue structure*

2009

Laurent Duret, Université Lyon, Laboratoire de Biométrie et Biologie Évolutive, Lyon, France, 15.10.2009. *Accelerated evolution in the human genome: selection or biased gene conversion?*

Kai Ye, Leiden University Medical Center, 05.10.2009. *Detecting Breakpoints of Large Deletions and Medium Sized Insertions on the Low Coverage Samples of 1000Genomes Project and High Coverage Samples of the Cancer Genome Project from Pair-end Short Reads*

Christian M. Reidys, Center for Combinatorics, Nankai University, China, 17.09.2009. *RIP: RNA Interaction Prediction*

Jun Yan, CAS-MPG Partner Institute Shanghai, 16.09.2009. *Genomic Approaches to Circadian Rhythm, Sleep, and Hibernation*

Martin Weigt, Institute for Scientific Interchange, Torino, 15.09.2009. *Inference of protein-protein interactions from multi-species sequence data*

Ivan Gesteira Costa Filho, Federal University of Penambuco, 09.07.2009. *Robust Classification of Clinical Time Series*

Jörg Schulz, Universität Würzburg, Bio-Zentrum, 10.06.2009. *Positive Selection in Tick Saliva Proteins of the Salp15 Family*

Attila Gulyas-Kovacs, Rockefeller University, 09.06.2009. *Sequence Analysis*

Sebastiaan Meijsing, Cellular and Molecular Pharmacology, USCF, San Francisco, 27.01.2009. *DNA-binding site sequence directs glucocorticoid receptor structure and activity*

Roland Dosch, Department of Zoology, University of Geneva, 09.01.2009. *Molecular Control of Zebrafish Oogenesis*

Research Group Development & Disease

(Established: 05/2000)

Head
Prof. Dr. Stefan Mundlos
Phone: +49 (0)30 8413-1449
Fax: +49 (0)30 8413-1385
Email: mundlos@molgen.mpg.de

Scientists
Hardy Chan (since 07/12)
Dario Lupianez (since 05/12)
Malte Spielmann* (since 09/10)
Sandra Dölken* (since 07/09)
Peter Krawitz* (since 01/09)
Johannes Egerer* (since 09/07)
Claus Eric Ott* (since 03/05)
Mateusz Kolanczyk*(since 07/03)
Uwe Kornak* (since 03/03)
Sigmar Stricker (since 09/02)
Jochen Hecht* (since 10/01)
Peter Robinson* (since 10/00)
Pablo Villavicencio-Lorini* (02/05-10/12)
Eva Klopocki* (09/03-10/12)
Pia Kuss* (07/09-06/12)
Katrin Hoffmann* (01/04-12/09)
Petra Seemann* (01/06-02/09)

PhD students
Denise Emmerich* (since 08/12)
Sinje Geuer* (since 01/12)
Anja Will* (since 12/11)
Jürgen Stumm (since 10/11)
Magdalena Steiner* (since 04/11)
Pedro Vallecillo-Garcia (since 01/11)
Martin Franke (since 10/10)
Julia Grohmann* (since 06/10)
Daniel Ibrahim* (since 08/09)
Silke Lohan* (since 01/09)
Saniye Yumlu* (since 01/09)
Hendrikje Hein* (since 06/08)
Björn Fischer* (since 04/08)
Gao Guo* (02/08-06/12)
Hardy Chan (01/07-06/12)
Sebastian Bauer* (01/07-03/12)
Jirko Kühnisch* (09/05-10/11)
Nadine Kossler* (03/05-03/11)
Christian Rödelsperger* (09/07-01/11)
Katerina Dimopoulou* (11/07-07/10)
Anja Brehm* (10/05-12/09)
Wiebke Schwarzer* (05/06-12/09)
Julia Friedrich* (06/08-03/09)
Florian Witte* (01/05-02/09)

* externally funded

Diploma students
Jenny Viebig (since 04/12)
Krysztof Brzezinka (since 03/12)
Rieke Fischer (01/12-07/12)
Katerina Kraft (03/11-08/11)
Sarah Altmeyer (05/10-06/11)
Julia Dikow (12/10-03/11)
Dominik Jost (12/10-03/11)
Stephanie Wiegand (05/10-02/11)
Daniel Hirsch (04/10-12/10)
Denise Rockstroh (08/09-05/10)
Susanne Mathia (06/09-04/10)
Bianca Hennig (04/09-12/09)
Julia Meier (02/07-02/09)

Technicians
Nicole Rösener* (since 03/07)
Monika Osswald (since 05/05)
Norbert Brieske (since 05/00)
Asita Stiege (since 05/00)

Clinical Genetics
Denise Horn* (since 12/03)

Introduction

Structure of the research group

The research group *Development & Disease* focuses on fundamental questions regarding normal and abnormal development. A particular interest is in the basic mechanisms of skeletal development and growth. The research group is part of and works in close collaboration with the *Institute for Medical Genetics and Human Genetics (IMG)* at the Charité - Universitätsmedizin Berlin. The IMG provides clinical and diagnostic genetic service within Germany and the EU. Medical doctors in training get scientific education at the Development & Disease group and scientists from the MPIMG have the opportunity to specialize in Medical Genetics at the IMG. A shared infrastructure, exchange of technical achievements and expertise, as well as common research goals ensure a successful interdisciplinary approach to study the mechanisms of genetic disease. Thus, the research group Development & Disease and the IMG form a highly complementary unit that combines clinical expertise with a basic science approach to address genetic questions.
Development and regeneration are related and it is generally believed that developmental pathways get re-activated during healing processes. To synergistically use our expertise in the molecular control of cell differentiation and development with new advances in regenerative medicine we collaborate closely with the Berlin-Brandenburg Center for Regenerative Medicine (BCRT) which is funded by the BMBF. Two members of the lab, Jochen Hecht and Petra Seemann, were appointed as group leaders at the BCRT but remain affiliated with the department.

Cooperation within the institute

Cooperations over the past years have been with the Lehrach Department on an EU-funded large scale gene expression study using automated *in situ* hybridization technology, the Herrmann Department on novel technologies for 3D bone imaging and mouse transgenic technology, and the Ropers Department on array-CGH as well as on projects to identify genetic defects in conditions with mental retardation. Several cooperations exist with the Vingron Department on computational analysis of ChIP-Seq data and other bioinformatic projects. Intense collaborations exist with the mouse and the sequencing facilities.

* externally funded

Special facilities / equipment

The research group as well as the IMG is equipped with the standard facilities for research into genetics, developmental biology, cell biology, and molecular biology. Special equipment includes a histology unit for the MPIMG and a sequencing facility for the Charité.

Research concept

The mechanisms, by which DNA sequences influence human development, function, and aging has moved into the centre of medical research creating the basis for what is now called molecular medicine. By combining developmental biology, genetics, and clinical medicine we aim at generating in-depth knowledge of human disease, in particular those conditions that are related to abnormal development, growth, and aging of the musculoskeletal system. The MPIMG Development & Disease group focuses on analysis of normal and abnormal developmental processes in model systems. Through clinical and diagnostic services as well as collaborations, patient cohorts are generated that are analyzed for genetic defects. This involves patient recruitment, expert phenotyping, data management and analysis, as well as mutation detection, mapping, and disease gene identification. The latter has profited greatly from the recent technology developments in genome analysis such as next-generation sequencing. More recently, the focus of our work has shifted towards the mechanisms of gene regulation during development and the analysis of mutations that alter this process. At the Development & Disease group we are well equipped to test novel disease genes/mutations for their functional relevance in established *in vitro* and *in vivo* model systems. The major *in vivo* systems are genetically engineered mice, chicken embryos, and zebrafish. Thus, our approach synergistically combines basic science-oriented research at the MPIMG with the more clinically oriented work at the IMG for research into human genetic disorders

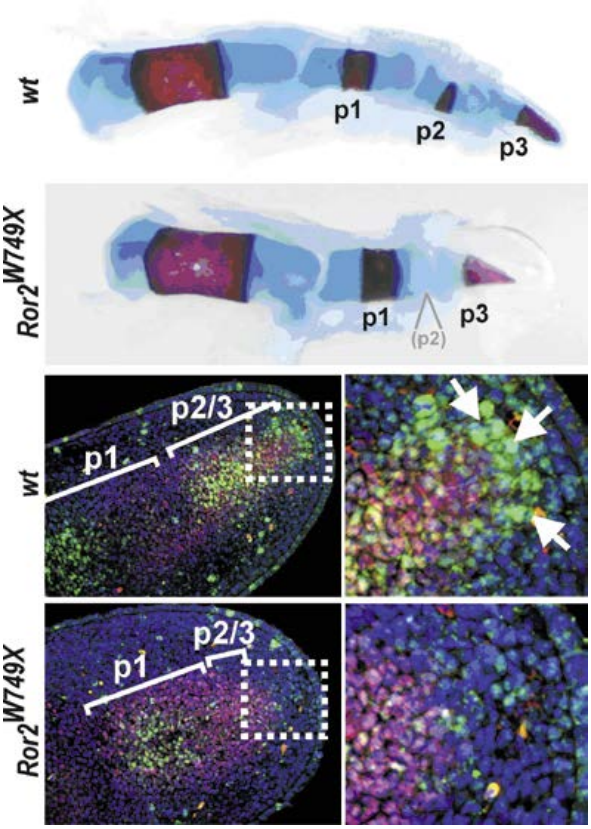
Scientific achievements / findings

Major achievements over the last years have been the identification and characterization of skeletal defects on a genetic, molecular and developmental level. The skeleton is a particularly informative model system for our phenotype driven approach, because of an almost unlimited number of distinct phenotypes. Over the past years our aim has been to learn about the origins of disease and how basic developmental mechanisms control phenotypic outcome. To achieve these goals the department is set up as an interactive unit with close cooperation between all members of the different research groups. Currently, our focus is on the following topics:

Mechanisms of limb development

We have been investigating basic mechanisms of limb development. One focus has been on normal and abnormal digit development mainly based on our clinical interest in brachydactylies. Using a disease and phenotype driven approach, we identified the BMP pathway as the major player in digit and joint formation and were able to correlate mutations and their mechanisms with disease phenotypes. The observation that the dysregulation of a particular pathway results in overlapping phenotypes, as exemplified in the brachydactylies, led to the con-

Figure 1: Phalanx forming region. Abnormal growth in the digits of a brachydactyly mouse (Ror-2W749X) with missing middle phalanx (p2). Phalanx forming region at the tip of the most distal digits (box) shows BMP-activity (green, arrows) in wt but not in mutant mice. Chondrocytes of the digit condensations are stained in red.



cept of molecular disease families. By studying several brachdactyly mutants, we were able to study their pathology and characterize a novel signaling center in mouse limb development, the phalanx forming region, which is essential for distal outgrowth mediated *via* mesenchymal bone morphogenetic protein (BMP) signaling (Figure 1).

Another major interest is in the function of homeobox genes in limb development. In humans mutations in HOXD13 results in synpolydactyly, a limb malformation characterized by an additional finger and a fusion between digits 3 and 4. The mutations that cause this condition are rather unusual as they comprise expansions of a polyalanine tract in the N-terminal region of the HOXD13 protein. The expansion results in protein aggregation and an interaction with other polyalanine containing proteins thereby inactivating them. The severity of pathology therefore correlates with the length of the expansion, an observation that was also confirmed *in vivo* by creating a mutant with a very long expansion (+ 21 alanines) (Figure 2). We showed that Hoxd13 controls bone formation in the limbs by directly activating Runx2, the transcription factor essential for osteoblast differentiation. Furthermore, we demonstrate that Hox genes determine the shape and identify of limb bones and that their inactivation causes a homeotic transformation of long bones (metacarpals) into round bones (carpals). The latter process involves the Wnt pathway and in particular Wnt5a thereby regulating cell polarity and, consequently, the shape of bones.

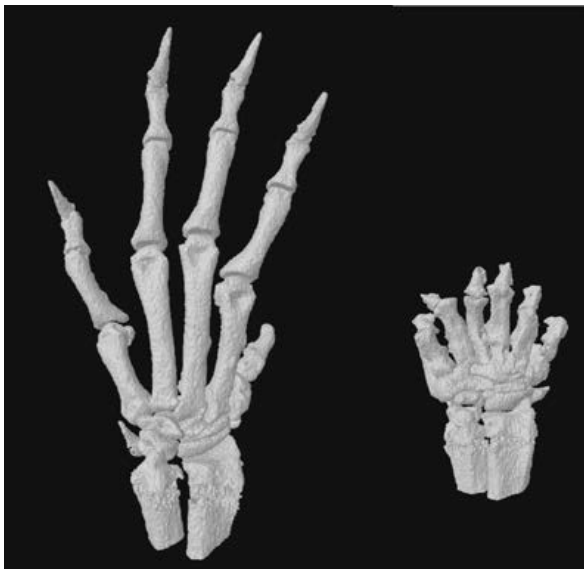


Figure 2: µCT image of digits in wt (left) and Hoxd13 mutant (right) mice. In the mutant the N-terminal repeat has been expanded by 21 alanines.

Together, our findings show that Hox genes are essential modifiers of shape and gestalt of the limbs by controlling stem cell differentiation into chondrocytes or osteoblasts. Furthermore, we have been able to unravel the basic mechanism how poly-alanine expansions in TFs result in disease.
Pia Kuss and Pablo Villavicencio-Lorini have been the driving force in the Hox project. Florian Witte and Sigmar Stricker have been studying general mechanisms of limb development.

Transcription factors in bone/limb development

Many developmental defects are due to mutations in transcription factors (TFs). The analysis of TF function is therefore essential to understand developmental abnormalities. We expanded our functional analysis of TFs by establishing methodologies to identify TF targets and binding sites within the genome. We adapted the technology of chromatin immunoprecipitation followed by deep sequencing (ChIP-Seq) to analyze TFs and their sequence variants in a standardized *in vitro* system. We use the chicken micromass system and tagged versions of the TFs for this purpose. Using this technology, we have been able to create a genomic binding profile for various TFs that play important roles in bone/limb development including Hox genes of the A and D cluster, Runx2, Pitx1, Twist and others. Furthermore, we tested mutations identified in our patient screen. One of these mutations (Q317K) was shown to convert HOXD13 into a TF with PITX1 binding properties. In another project we systematically analyze the genomic binding sites for 5' Hoxd and Hoxa genes with the aim to identify the mechanisms of Hox gene target identification.
Daniel Ibrahim, Hendrijke Hein, Ivana Jerkovic, and Jochen Hecht are in charge of this project.

Long range regulation

Most developmentally important genes have complex expression patterns that show distinct differences in temporal and spatial distribution. Cis-regulatory enhancer elements are believed to play an important role in this process. By screening large cohorts of patients with limb malformations *via* array-CGH, we have identified a series of duplications involving non-coding conserved elements (CNEs) that are located in the vicinity of developmentally important genes. This includes duplications involving BMP2 (brachydactyly type A2), SHH (mirror image polydactyly), SOX9 (Cooks syndrome), IHH (craniosynostosis with syndactyly), and MSX2 (cleidocranial dysplasia). Our findings identified duplications of CNEs as a novel mutational mechanism for human disease. In addition, they show that CNEs are important for fine tuning expression and that alterations in these regions can result in unexpected phenotypes.
To understand the mechanisms of disease and how these CNEs regulate gene expression, we are in the process of creating mouse models for several of the above mentioned loci. We use the sleeping beauty system of transposons linked with cre recombination sites. With this system we can create duplications but also deletions within the targeted region. Furthermore, the transposons can be mobilized to create further sites thereby staturating the region of interest. Other technologies to study the regulatory genome include ChIP to identify regulatory elements and chromatin conformation capture technologies.
Dario Lupianez, Martin Franke, Anja Will, Malte Spielmann, Claus Eric Ott, and Eva Klopocki are in charge of this project.

Osteoporosis and mechanisms of aging

In humans, aging is accompanied by loss of bone mass. Bone loss may result in an increased susceptibility to fractures and thus a significant disease burden. To elucidate the molecular processes that govern aging in bone, we studied a group of recessively inherited diseases. We have been able to identify disease causing mutations in three different genes (GORAB, ATP6V0A2, and PYCR1), two of which are involved in the Golgi network. We created a mouse model by inactivating Gorab and show that these mice recapitulate the human disease. Our findings provide new insights into the molecular mechanisms of skin aging and osteoporosis. Proper function of the Golgi apparatus appears to be important for the maintenance of healthy skin and bone. Increased susceptibility to apoptosis and/or senescence appears to be an important trigger for age-related changes in skin and bone.
Björn Fischer, Magdalena Steiner, Johannes Egerer, and Uwe Kornak are in charge of this project.

Muscle and connective tissue development

Muscles and the skeleton form a highly interdependent functional unit. The development of the skeleton is dependent on the correct function and interconnection of muscles to bones *via* the tendons, and skeletal homeostasis can only be reached through biomechanical interaction with the musculature. We are interested in the interaction of different mesenchymal lineages during musculoskeletal development and its alteration in disease. The main focus lies on the crosstalk of tissue progenitors, namely muscle cells, connective tissue cells (constituting loose connective tissue and tendons) as well as cartilage/bone. Connective tissue is an irreplaceable component of the musculoskeletal system; however it has attracted far less attention than other tissue types. We have identified transcription factors specifically expressed in limb connective tissue and are currently analyzing those factors *in vitro* (downstream targets *via* ChIP-Seq, microarrays) and in mouse and chicken models. Furthermore, we identified neurofibromin, the protein mutated in NF1, as a relevant regulator of muscle development.
Pedro Vallecillo-Garcia, Jürgen Stumm and Sigmar Stricker are in charge of this project.

Bioinformatics

One topic of the bioinformatics group is the application of ontologies to describe phenotypic features seen in hereditary and other forms of human disease. The Human Phenotype Ontology (HPO) was developed as a tool to study phenotypic features with bioinformatic means and other forms of computational analysis. The program has been adopted by international research groups for phenotyping, including the ClinVar project, Orphanet, ISCA, the DECIPHER group and the DDD project at the Sanger Center. Furthermore, the HPO has been used to develop a clinical diagnostics algorithm for human genetics that utilizes a novel statistical model of semantic similarities in ontologies to provide a ranking of the candidate differential diagnoses and have developed a novel graph algorithm that accelerates semantic searches in ontologies by many orders of magnitude. Other topics of the group include the development of next generation sequencing applications and support for ChIP-seq as well as microRNA analysis.
Peter Robinson leads the bioinformatics group.

Genome analysis for disease gene identification

Our focus is to identify genetic factors that cause or modify monogenic diseases. Learning about the cause of a disease helps to understand, or to start to study, the subsequent disease processes and aims to develop more effective diagnostics and eventually preventive or therapeutic strategies. Based on the recent technological advances in genome analysis we have been active in establishing and improving the technology, in particular the bioinformatics part. Currently we are focusing on the identification of regulatory mutations in non-transcribed sequences using a whole genome approach.

This project is interdisciplinary and involves several departments at the Charité and elsewhere, clinicians for sampling, diagnosing, and phenotyping as well as bioinformatitions for the analysis of phenotypic and sequence data, and sequencing technology for mutation identification. The continous supply of patient material provides us with a constant flow of novel genes and mutations. Novel gene mutations are being tested in the available model system. A recent example is the identification of duplications associated with ectrodactyly with tibial hemimelia. The gene included in the duplication, Bhlha9, was investigated in mice and zebrafish and shown to play role in the development of the zebrafish fin (Figure 3).

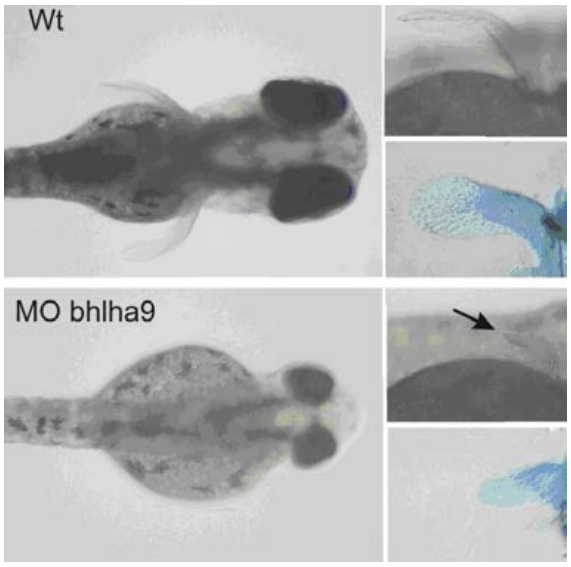


Figure 3: Duplications of BHLHA9 result in ectrodactyly with tibial hemimelia. Knock down of Bhlha9 in zebrafish results in small fins. Top shows native fins, bottom alcian blue staining of fin cartilage.

Denise Horn represents the clinical aspect of this project, Nick Robinson the bioinformatic part, and Jochen Hecht is in charge of the sequencing. Peter Krawitz has established analysis routines for the Illumina data and for using the exon-enriched NGS sequencing for clinical diagnostics.

Planned developments

We will stick to our basic research concept and plan to continue with the identification of novel disease genes in combination with functional studies with a focus on skeletal development. Novel technologies based on NGS will be used to identify the molecular basis of so far unknown conditions. However, due to the broad availability of NGS and intense activities in this field world wide, it is to be expected that the uncovering of novel disease genes will become increasingly difficult. The role of non-coding DNA or non-translated sequences in disease is, in contrast, largely unknown. Based on our previous findings, the expertise in the group and in the Institute, we will shift the focus of our studies towards the analysis of regulatory mutations. We will use array-CGH and whole genome sequencing in patients with yet unidentified mutations for this purpose. The focus will be on individuals with limb malformations. To reduce the genomic complexity, we plan to filter for those sequences that are relevant for limb development. To identify these sequences, we will use publicly available data and own ChIP experiments to create a limb “regulome”. Furthermore, we aim at systematically analyzing the regulatory landscape of selected key genes to get an in depth understanding of the gene’s developmental regulation. This includes the creation of deletions and/or duplications of regulatory elements within these regions in mouse models.

It is our aim to translate the new genomics into clinical practice. We are in the process of establishing genome analysis methodology for diagnostic purposes in rare Mendelian diseases. This is a collaborative project with several partners at the Charité and elsewhere. One important aspect is to standardize phenotypic assessment and documentation in order to correlate the phenotype with the genomic data. The development of the Human Phenotype Ontology has been a milestone in this process. Using this approach, we will be able to combine phenotypic and molecular data in a systematic and comprehensive manner allowing us to better understand and ultimately predict the consequences of mutations.

General information

Complete list of publications (2009-2012)

2012

Chen CK, Mungall CJ, Gkoutos GV, Doelken SC, Köhler S, Ruef BJ, Smith C, Westerfield M, Robinson PN, Lewis SE, Schofield PN, Smedley D (2012). *MouseFinder: Candidate disease genes from mouse phenotype data.* Hum Mutat 33(5):858-66.

Klopocki E, Kähler C, Foulds N, Shah H, Joseph B, Vogel H, Lüttgen S, Bald R, Besoke R, Held K, Mundlos S, Kurth I (2012). *Deletions in PITX1 cause a spectrum of lower-limb malformations including mirror-image polydactyly.* Eur J Hum Genet 20(6):705-8.

Klopocki E, Lohan S, Doelken SC, Stricker S, Ockeloen CW, Soares Thiele de Aguiar R, Lezirovitz K, Mingroni Netto RC, Jamsheer A, Shah H, Kurth I, Habenicht R, Warman M, Devriendt K, Kordass U, Hempel M, Rajab A, Mäkitie O, Naveed M, Radhakrishna U, Antonarakis SE, Horn D, Mundlos S (2012). *Duplications of BHLHA9 are associated with ectrodactyly and tibia hemimelia inherited in non-Mendelian fashion*. J Med Genet 49(2):119-25

Köhler S, Doelken SC, Rath A, Aymé S, Robinson PN (2012). Ontological phenotype standards for neurogenetics. Hum Mutat. 2012 May 9. doi: 10.1002/humu.22112

Krawitz PM, Murakami Y, Hecht J, Krüger U, Holder SE, Mortier GR, Delle Chiaie B, De Baere E, Thompson MD, Roscioli T, Kielbasa S, Kinoshita T, Mundlos S, Robinson PN, Horn D (2012). *Mutations in PIGO, a Member of the GPI-Anchor-Synthesis Pathway, Cause Hyperphosphatasia with Mental Retardation.* Am J Hum Genet 2012 Jun 6.

Murakami Y, Kanzawa N, Saito K, Krawitz PM, Mundlos S, Robinson PN, Karadimitris A, Maeda Y, Kinoshita T (2012). *Mechanism for release of alkaline phosphatase caused by glycosylphosphatidylinositol deficiency in patients with hyperphosphatasia mental retardation syndrome*. J Biol Chem 287(9):6318-25

Ott CE, Hein H, Lohan S, Hoogeboom J, Foulds N, Grünhagen J, Stricker S, Villavicencio-Lorini P, Klopocki E, Mundlos S (2012). *Microduplications upstream of MSX2 are associated with a phenocopy of cleidocranial dysplasia*. J Med Genet 2012 Jun 20.

Robinson PN (2012). *Deep phenotyping for precision medicine.* Hum Mutat 33(5):777-80.

Rosenfeld JA, Traylor RN, Schaefer GB, McPherson EW, Ballif BC, Klopocki E, Mundlos S, Shaffer LG, Aylsworth AS (2012). *Proximal microdeletions and microduplications of 1q21.1 contribute to variable abnormal phenotypes*. Eur J Hum Genet, doi: 10.1038/ejhg.2012.6.

Stricker S, Mathia S, Haupt J, Seemann P, Meier J, Mundlos S (2012). *Odd-skipped related genes regulate differentiation of embryonic limb mesenchyme and bone marrow mesenchymal stromal cells*. Stem Cells Dev 21(4):623-33

2011

Baasanjav S, Al-Gazali L, Hashiguchi T, Mizumoto S, Fischer B, Horn D, Seelow D, Ali BR, Aziz SA, Langer R, Saleh AA, Becker C, Nürnberg G, Cantagrel V, Gleeson JG, Gomez D, Michel JB, Stricker S, Lindner TH, Nürnberg P, Sugahara K, Mundlos S, Hoffmann K (2011). *Faulty initiation of proteoglycan synthesis causes cardiac and joint defects*. Am J Hum Genet 89(1):15-27

Blau O, Baldus CD, Hofmann WK, Thiel G, Nolte F, Burmeister T, Türkmen S, Benlasfer O, Schümann E, Sindram A, Molkentin M, Mundlos S, Keilholz U, Thiel E, Blau IW (2011). *Mesenchymal stromal cells of myelodysplastic syndrome and acute myeloid leukemia patients have distinct genetic abnormalities compared with leukemic blasts*. Blood 118(20):5583-92

Diez-Roux G, Banfi S, Sultan M, Geffers L, Anand S, Rozado D, Magen A, Canidio E, Pagani M, Peluso I, Lin-Marq N, Koch M, Bilio M, Cantiello I, Verde R, De Masi C, Bianchi SA, Cicchini J, Perroud E, Mehmeti S, Dagand E, Schrinner S, Nürnberger A, Schmidt K, Metz K, Zwingmann C, Brieske N, Springer C, Hernandez AM, Herzog S, Grabbe F, Sieverding C, Fischer B, Schrader K, Brockmeyer M, Dettmer S, Helbig C, Alunni V, Battaini MA, Mura C, Henrichsen CN, Garcia-Lopez R, Echevarria D, Puelles E, Garcia-Calero E, Kruse S, Uhr M, Kauck C, Feng G, Milyaev N, Ong CK, Kumar L, Lam M, Semple CA, Gyenesei A, Mundlos S, Radelof U, Lehrach H, Sarmientos P, Reymond A, Davidson DR, Dollé P, Antonarakis SE, Yaspo ML, Martinez S, Baldock RA, Eichele G, Ballabio A (2011). *A high-resolution anatomical atlas of the transcriptome in the mouse embryo*. PLoS Biol 9(1):e1000582

Guo G, Gehle P, Doelken S, Martin-Ventura JL, von Kodolitsch Y, Hetzer R, Robinson PN (2011). *Induction of macrophage chemotaxis by aortic extracts from patients with Marfan syndrome is related to elastin binding protein*. PLoS One 6(5):e20138.

Heinrich V, Stange J, Dickhaus T, Imkeller P, Krüger U, Bauer S, Mundlos S, Robinson PN, Hecht J, Krawitz PM (2011). *The allele distribution in next-generation sequencing data sets is accurately described as the result of a stochastic branching process*. Nucleic Acids Res 40(6):2426-31. Epub 2011 Nov 29.

Horn D, Robinson PN (2011). *Progeroid facial features and lipodystrophy associated with a novel splice site mutation in the final intron of the FBN1 gene*. Am J Med Genet A 155A(4):721-4.

Jäger M, Ott CE, Grünhagen J, Hecht J, Schell H, Mundlos S, Duda GN, Robinson PN, Lienau J (2011). *Composite transcriptome assembly of RNA-seq data in a sheep model for delayed bone healing*. BMC Genomics 12:158

Joss S, Kini U, Fisher R, Mundlos S, Prescott K, Newbury-Ecob R, Tolmie J (2011). *The face of Ulnar Mam-*

mary syndrome? Eur J Med Genet 54(3):301-5

Kerschnitzki M, Wagermaier W, Roschger P, Seto J, Shahar R, Duda GN, Mundlos S, Fratzl P (2011). *The organization of the osteocyte network mirrors the extracellular matrix orientation in bone.* J Struct Biol 173(2):303-11

Klopocki E, Lohan S, Brancati F, Koll R, Brehm A, Seemann P, Dathe K, Stricker S, Hecht J, Bosse K, Betz RC, Garaci FG, Dallapiccola B, Jain M, Muenke M, Ng VC, Chan W, Chan D, Mundlos S (2011). *Copy-number variations involving the IHH locus are associated with syndactyly and craniosynostosis*. Am J Hum Genet 88(1):70-5

Klopocki E, Mundlos S (2011). *Copy-number variations, noncoding sequences, and human phenotypes*. Annu Rev Genomics Hum Genet 12:53-72

Köhler S, Bauer S, Mungall CJ, Carletti G, Smith CL, Schofield P, Gkoutos GV, Robinson PN (2011). *Improving ontologies by automatic reasoning and evaluation of logical definitions.* BMC Bioinformatics 12:418

Kolanczyk M, Mautner V, Kossler N, Nguyen R, Kühnisch J, Zemojtel T, Jamsheer A, Wegener E, Thurisch B, Tinschert S, Holtkamp N, Park SJ, Birch P, Kendler D, Harder A, Mundlos S, Kluwe L (2011). *MIA is a potential biomarker for tumour load in neurofibromatosis type 1*. BMC Med 9:82.

Kolanczyk M, Pech M, Zemojtel T, Yamamoto H, Mikula I, Calvaruso MA, van den Brand M, Richter R, Fischer B, Ritz A, Kossler N, Thurisch B, Spoerle R, Smeitink J, Kornak U, Chan D, Vingron M, Martasek P, Lightowlers RN, Nijtmans L, Schuelke M, Nierhaus KH, Mundlos S (2011). *NOA1 is an essential GTPase required for mitochondrial protein synthesis*. Mol Biol Cell 22(1):1-11

Kossler N, Stricker S, Rödelsperger C, Robinson PN, Kim J, Dietrich C, Osswald M, Kühnisch J, Stevenson DA, Braun T, Mundlos S, Kolanczyk M (2011). *Neurofibromin (Nf1) is required for skeletal muscle development*. Hum Mol Genet 20(14):2697-709

Lange C, Li C, Manjubala I, Wagermaier W, Kühnisch J, Kolanczyk M, Mundlos S, Knaus P, Fratzl P (2011). *Fetal and postnatal mouse bone tissue contains more calcium than is present in hydroxyapatite*. J Struct Biol 176(2):159-67

Lindblom A, Robinson PN. *Bioinformatics for human genetics: promises and challenges.* Hum Mutat. 2011 May;32(5):495-500.

Marchal JA, Ghani M, Schindler D, Gavvovidis I, Winkler T, Esquitino V, Sternberg N, Busche A, Krawitz P, Hecht J, Robinson P, Mundlos S, Graul-Neumann L, Sperling K, Trimborn M, Neitzel H (2011). *Misregulation of mitotic chromosome segregation in a new type of autosomal recessive primary microcephaly*. Cell Cycle 10(17):2967-77

Robinson PN, Krawitz P, Mundlos S (2011). *Strategies for exome and genome sequence data analysis in disease-gene discovery projects*. Clin Genet 80(2):127-32

Rödelsperger C, Krawitz P, Bauer S, Hecht J, Bigham AW, Bamshad M, de Condor BJ, Schweiger MR, Robinson PN (2011). *Identity-by-descent filtering of exome sequence data for disease-gene identification in autosomal recessive disorders*. Bioinformatics 27(6):829-36.

Rump P, Jongbloed JD, Sikkema-Raddatz B, Mundlos S, Klopocki E, van der Luijt RB (2011). *Madelung deformity in a girl with a novel and de novo mutation in the GNAS gene*. Am J Med Genet A 155A(10):2566-70

Schulz MH, Köhler S, Bauer S, Robinson PN (2011). *Exact score distribution computation for ontological similarity searches*. BMC Bioinformatics 12:441.

Spielmann M, Reichelt G, Hertzberg C, Trimborn M, Mundlos S, Horn D, Klopocki E (2011). *Homozygous deletion of chromosome 15q13.3 including CHRNA7 causes severe mental retardation, seizures, muscular hypotonia, and the loss of KLF13 and TRPM1 potentially cause macrocytosis and congenital retinal dysfunction in siblings.* Eur J Med Genet 54(4):e441-5

Stricker S, Mundlos S (2011). *FGF and ROR2 receptor tyrosine kinase signaling in human skeletal development.* Curr Top Dev Biol 97:179-206

Stricker S, Mundlos S (2011). *Mechanisms of digit formation: Human malformation syndromes tell the story*. Dev Dyn 240(5):990-1004

Warman ML, Cormier-Daire V, Hall C, Krakow D, Lachman R, LeMerrer M, Mortier G, Mundlos S, Nishimura G, Rimoin DL, Robertson S, Savarirayan R, Sillence D, Spranger J, Unger S, Zabel B, Superti-Furga A (2011). *Nosology and classification of genetic skeletal disorders: 2010 revision*. Am J Med Genet A 155A(5):943-68

2010

Baasanjav S, Jamsheer A, Kolanczyk M, Horn D, Latos T, Hoffmann K, Latos-Bielenska A, Mundlos S (2010). *Osteopoikilosis and multiple exostoses caused by novel mutations in LEMD3 and EXT1 genes respectively--coincidence within one family.* BMC Med Genet 11:110

Bauer S, Gagneur J, Robinson PN (2010). *GOing Bayesian: model-based gene set analysis of genome-scale data*. Nucleic Acids Res 38(11):3523-32.

Brancati F, Fortugno P, Bottillo I, Lopez M, Josselin E, Boudghene-Stambouli O, Agolini E, Bernardini L, Bellacchio E, Iannicelli M, Rossi A, Dib-Lachachi A, Stuppia L, Palka G, Mundlos S, Stricker S, Kornak U, Zambruno G, Dallapiccola B (2010). *Mutations in PVRL4, encoding cell adhesion molecule nectin-4, cause ectodermal dysplasia-syndactyly syndrome*. Am J Hum Genet 87(2):265-73

Cirstea IC, Kutsche K, Dvorsky R, Gremer L, Carta C, Horn D, Roberts AE, Lepri F, Merbitz-Zahradnik T, König R, Kratz CP, Pantaleoni F, Dentici ML, Joshi VA, Kucherlapati RS, Mazzanti L, Mundlos S, Patton MA, Silengo MC, Rossi C, Zampino G, Digilio C, Stuppia L, Seemanova E, Pennacchio LA, Gelb BD, Dallapiccola B, Wittinghofer A, Ahmadian MR, Tartaglia M, Zenker M (2010). *A restricted spectrum of NRAS mutations causes Noonan syndrome*. Nat Genet 42(1):27-9

Clayton P, Fischer B, Mann A, Mansour S, Rossier E, Veen M, Lang C, Baasanjav S, Kieslich M, Brossuleit K, Gravemann S, Schnipper N, Karbasyian M, Demuth I, Zwerger M, Vaya A, Utermann G, Mundlos S, Stricker S, Sperling K, Hoffmann K (2010). *Mutations causing Greenberg dysplasia but not Pelger anomaly uncouple enzymatic from structural functions of a nuclear membrane protein*. Nucleus 1(4):354-66

Harder A, Titze S, Herbst L, Harder T, Guse K, Tinschert S, Kaufmann

D, Rosenbaum T, Mautner VF, Windt E, Wahlländer-Danek U, Wimmer K, Mundlos S, Peters H (2010). *Monozygotic twins with neurofibromatosis type 1 (NF1) display differences in methylation of NF1 gene promoter elements, 5' untranslated region, exon and intron 1*. Twin Res Hum Genet 13(6):582-94

Horbelt D, Guo G, Robinson PN, Knaus P (2010). *Quantitative analysis of TGFBR2 mutations in Marfan-syndrome-related disorders suggests a correlation between phenotypic severity and Smad signaling activity*. J Cell Sci 123(Pt 24):4340-50.

Kantaputra PN, Klopocki E, Hennig BP, Praphanphoj V, Le Caignec C, Isidor B, Kwee ML, Shears DJ, Mundlos S (2010). *Mesomelic dysplasia Kantaputra type is associated with duplications of the HOXD locus on chromosome 2q*. Eur J Hum Genet 18(12):1310-4

Kantaputra PN, Mundlos S, Sripathomsawat W (2010). *A novel homozygous Arg222Trp missense mutation in WNT7A in two sisters with severe Al-Awadi/Raas-Rothschild/Schinzel phocomelia syndrome*. Am J Med Genet A 152A(11):2832-7

Klopocki E, Hennig BP, Dathe K, Koll R, de Ravel T, Baten E, Blom E, Gillerot Y, Weigel JF, Krüger G, Hiort O, Seemann P, Mundlos S (2010). *Deletion and point mutations of PTHLH cause brachydactyly type E*. Am J Hum Genet 86(3):434-9

Krawitz P, Rödelsperger C, Jäger M, Jostins L, Bauer S, Robinson PN (2010). *Microindel detection in short-read sequence data*. Bioinformatics 26(6):722-9

Krawitz PM, Schweiger MR, Rödelsperger C, Marcelis C, Kölsch U, Meisel C, Stephani F, Kinoshita T, Murakami Y, Bauer S, Isau M, Fischer A, Dahl A, Kerick M, Hecht J, Köhler S, Jäger M, Grünhagen J, de Condor BJ, Doelken S, Brunner HG, Meinecke P, Passarge E, Thompson MD, Cole DE, Horn D, Roscioli T, Mundlos S, Robinson PN (2010). *Identity-by-descent filtering of exome sequence data identifies PIGV mutations in hyperphosphatasia mental retardation syndrome*. Nat Genet 42(10):827-9

Lacombe D, Delrue MA, Rooryck C, Morice-Picard F, Arveiler B, Maugey-Laulom B, Mundlos S, Toutain A, Chateil JF (2010). *Brachydactyly type A1 with short humerus and associated skeletal features*. Am J Med Genet A 152A(12):3016-21

Liska F, Snajdr P, Stricker S, Gosele C, Krenová D, Mundlos S, Hubner N (2010). *Impairment of Sox9 expression in limb buds of rats homozygous for hypodactyly mutation*. Folia Biol (Praha) 56(2):58-65

Ott CE, Grünhagen J, Jäger M, Horbelt D, Schwill S, Kallenbach K, Guo G, Manke T, Knaus P, Mundlos S, Robinson PN (2010). *MicroRNAs differentially expressed in postnatal aortic development downregulate elastin via 3' UTR and coding-sequence binding sites*. PLoS One 6(1):e16250

Ott CE, Leschik G, Trotier F, Brueton L, Brunner HG, Brussel W, Guillen-Navarro E, Haase C, Kohlhase J, Kotzot D, Lane A, Lee-Kirsch MA, Morlot S, Simon ME, Steichen-Gersdorf E, Tegay DH, Peters H, Mundlos S, Klopocki E (2010). *Deletions of the RUNX2 gene are present in about 10% of individuals with cleidocranial dysplasia*. Hum Mutat 31(8):E1587-93

Ratzka A, Mundlos S, Vortkamp A (2010). *Expression patterns of sulfatase genes in the developing mouse embryo*. Dev Dyn 239(6):1779-88

Robinson PN (2010). *Whole-exome sequencing for finding de novo mutations in sporadic mental retardation*. Genome Biol 11(12):144.

Robinson PN, Mundlos S (2010). *The human phenotype ontology*. Clin Genet 77(6):525-34

Villavicencio-Lorini P, Kuss P, Friedrich J, Haupt J, Farooq M, Türkmen S, Duboule D, Hecht J, Mundlos S (2010). *Homeobox genes d11-d13 and a13 control mouse autopod cortical bone and joint formation*. J Clin Invest 120(6):1994-2004

Witte F, Bernatik O, Kirchner K, Masek J, Mahl A, Krejci P, Mundlos S, Schambony A, Bryja V, Stricker S (2010). *Negative regulation of Wnt signaling mediated by CK1-phosphorylated Dishevelled via Ror2*. FASEB J 24(7):2417-26

Witte F, Chan D, Economides AN, Mundlos S, Stricker S (2010). *Receptor tyrosine kinase-like orphan receptor 2 (ROR2) and Indian hedgehog regulate digit outgrowth mediated by the phalanx-forming region*. Proc Natl Acad Sci U S A 107(32):14211-6

2009

Bieler FH, Ott CE, Thompson MS, Seidel R, Ahrens S, Epari DR, Wilkening U, Schaser KD, Mundlos S, Duda GN (2009). *Biaxial cell stimulation: A mechanical validation*. J Biomech 42(11):1692-6

Dathe K, Kjaer KW, Brehm A, Meinecke P, Nürnberg P, Neto JC, Brunoni D, Tommerup N, Ott CE, Klopocki E, Seemann P, Mundlos S (2009). *Duplications involving a conserved regulatory element downstream of BMP2 are associated with brachydactyly type A2*. Am J Hum Genet 84(4):483-92

Dutrannoy V, Klopocki E, Wei R, Bommer C, Mundlos S, Graul-Neumann LM, Trimborn M (2009). *De novo 9 Mb deletion of 6q23.2q24.1 disrupting the gene EYA4 in a patient with sensorineural hearing loss, cardiac malformation, and mental retardation*. Eur J Med Genet 52(6):450-3

Elefteriou F, Kolanczyk M, Schindeler A, Viskochil DH, Hock JM, Schorry EK, Crawford AH, Friedman JM, Little D, Peltonen J, Carey JC, Feldman D, Yu X, Armstrong L, Birch P, Kendler DL, Mundlos S, Yang FC, Agiostratidou G, Hunter-Schaedle K, Stevenson DA (2009). *Skeletal abnormalities in neurofibromatosis type 1: approaches to therapeutic options*. Am J Med Genet A 149A(10):2327-38

Gao B, Hu J, Stricker S, Cheung M, Ma G, Law KF, Witte F, Briscoe J, Mundlos S, He L, Cheah KS, Chan D (2009). *A mutation in Ihh that causes digit abnormalities alters its signalling capacity and range*. Nature 458(7242):1196-200

Hucthagowder V, Morava E, Kornak U, Lefeber DJ, Fischer B, Dimopoulou A, Aldinger A, Choi J, Davis EC, Abuelo DN, Adamowicz M, Al-Aama J, Basel-Vanagaite L, Fernandez B, Greally MT, Gillessen-Kaesbach G, Kayserili H, Lemyre E, Tekin M, Türkmen S, Tuysuz B, Yüksel-Konuk B, Mundlos S, Van Maldergem L, Wevers RA, Urban Z (2009). *Loss-of-function mutations in ATP6V0A2 impair vesicular trafficking, tropoelastin secretion and cell survival*. Hum Mol Genet 18(12):2149-65

Kaplan FS, Xu M, Seemann P, Connor JM, Glaser DL, Carroll L, Delai P, Fastnacht-Urban E, Forman SJ, Gillessen-Kaesbach G, Hoover-Fong J, Köster B, Pauli RM, Reardon W, Zaidi SA, Zasloff M, Morhart R, Mundlos S, Groppe J, Shore EM (2009). *Classic and atypical fibrodysplasia ossificans progressiva (FOP) phenotypes are caused by mutations in the*

bone morphogenetic protein (BMP) type I receptor ACVR1. Hum Mutat 30(3):379-90

Kjaer KW, Tiner M, Cingoz S, Karatosun V, Tommerup N, Mundlos S, Gunal I (2009). *A novel subtype of distal symphalangism affecting only the 4th finger*. Am J Med Genet A 149A(7):1571-3

Köhler S, Schulz MH, Krawitz P, Bauer S, Dölken S, Ott CE, Mundlos C, Horn D, Mundlos S, Robinson PN (2009). *Clinical diagnostics in human genetics with semantic similarity searches in ontologies*. Am J Hum Genet 85(4):457-64

Kurth I, Klopocki E, Stricker S, van Oosterwijk J, Vanek S, Altmann J, Santos HG, van Harssel JJ, de Ravel T, Wilkie AO, Gal A, Mundlos S (2009). *Duplications of noncoding elements 5' of SOX9 are associated with brachydactyly-anonychia*. Nat Genet 41(8):862-3

Kuss P, Villavicencio-Lorini P, Witte F, Klose J, Albrecht AN, Seemann P, Hecht J, Mundlos S (2009). *Mutant Hoxd13 induces extra digits in a mouse model of synpolydactyly directly and by decreasing retinoic acid synthesis*. J Clin Invest 119(1):146-56

Mundlos S (2009). *The brachydactylies: a molecular disease family*. Clin Genet 76(2):123-36

Ott CE, Bauer S, Manke T, Ahrens S, Rödelsperger C, Grünhagen J, Kornak U, Duda G, Mundlos S, Robinson PN (2009). *Promiscuous and depolarization-induced immediate-early response genes are induced by mechanical strain of osteoblasts*. J Bone Miner Res 24(7):1247-62

Reversade B, Escande-Beillard N, Dimopoulou A, Fischer B, Chng SC, Li Y, Shboul M, Tham PY, Kayserili H, Al-Gazali L, Shahwan M, Brancati F, Lee H, O'Connor BD, Schmidt-von Kegler M, Merriman B, Nelson SF, Masri A, Alkazaleh F, Guerra D, Ferrari P, Nanda A, Rajab A, Markie D, Gray M, Nelson J, Grix A, Sommer A, Savarirayan R, Janecke AR, Steichen E, Sillence D, Hausser I, Budde B, Nürnberg G, Nürnberg P, Seemann P, Kunkel D, Zambruno G, Dallapiccola B, Schuelke M, Robertson S, Hamamy H, Wollnik B, Van Maldergem L, Mundlos S, Kornak U (2009). *Mutations in PYCR1 cause cutis laxa with progeroid features*. Nat Genet 41(9):1016-21

Rödelsperger C, Köhler S, Schulz MH, Manke T, Bauer S, Robinson PN (2009). *Short ultraconserved promoter regions delineate a class of preferentially expressed alternatively spliced transcripts*. Genomics 94(5):308-16

Schwarzer W, Witte F, Rajab A, Mundlos S, Stricker S (2009). *A gradient of ROR2 protein stability and membrane localization confers brachydactyly type B or Robinow syndrome phenotypes*. Hum Mol Genet 18(21):4013-21

Seemann P, Brehm A, König J, Reissner C, Stricker S, Kuss P, Haupt J, Renninger S, Nickel J, Sebald W, Groppe JC, Plöger F, Pohl J, Schmidt-von Kegler M, Walther M, Gassner I, Rusu C, Janecke AR, Dathe K, Mundlos S (2009). *Mutations in GDF5 reveal a key residue mediating BMP inhibition by NOGGIN*. PLoS Genet 5(11):e1000747

Seifert W, Beninde J, Hoffmann K, Lindner TH, Bassir C, Aksu F, Hübner C, Verbeek NE, Mundlos S, Horn D (2009). *HPGD mutations cause cranioosteoarthropathy but not auto-*

somal dominant digital clubbing. Eur J Hum Genet 17(12):1570-6

Shen Q, Little SC, Xu M, Haupt J, Ast C, Katagiri T, Mundlos S, Seemann P, Kaplan FS, Mullins MC, Shore EM (2009). *The fibrodysplasia ossificans progressiva R206H ACVR1 mutation activates BMP-independent chondrogenesis and zebrafish embryo ventralization*. J Clin Invest 119(11):3462-72

Türkmen S, Guo G, Garshasbi M, Hoffmann K, Alshalah AJ, Mischung C, Kuss A, Humphrey N, Mundlos S, Robinson PN (2009). *CA8 mutations cause a novel syndrome characterized by ataxia and mild mental retardation with predisposition to quadrupedal gait*. PLoS Genet 5(5):e1000487

Tuysuz B, Mizumoto S, Sugahara K, Celebi A, Mundlos S, Turkmen S (2009). *Omani-type spondyloepiphyseal dysplasia with cardiac involvement caused by a missense mutation in CHST3*. Clin Genet 75(4):375-83

van Wijk NV, Witte F, Feike AC, Schambony A, Birchmeier W, Mundlos S, Stricker S (2009). *The LIM domain protein Wtip interacts with the receptor tyrosine kinase Ror2 and inhibits canonical Wnt signalling*. Biochem Biophys Res Commun 390(2):211-6

Witte F, Dokas J, Neuendorf F, Mundlos S, Stricker S (2009). *Comprehensive expression analysis of all Wnt genes and their major secreted antagonists during mouse limb development and cartilage differentiation*. Gene Expr Patterns 9(4):215-23

Invited plenary lectures (Stefan Mundlos)

Clinical relevance of copy number variation. British Human Genetics Conference, University of Warwick, September 17-19, 2012

Regulatory mutations – the next frontier in Human Genetics. European Society for Human Genetics, Erlangen, Germany, June 23-25, 2012

Regulating skeletal development – lessons to be learned from rare disease. Paris Descartes University Hôpital Necker, Paris, France, March 15, 2012

Regulatory CNVs. Genomic Disorders 2012: The Genomics of Rare Disease, Sanger Center, Hinxton, UK, March 21-24, 2012

HOX Genes Sculpture our Bones. Keynote lecture at the Day of Clinical Research of the Department Clinical Research at the University of Bern, Bern, Switzerland, Nov 2, 2011

Structural variations of the human genome and their role in congenital disease. Sanger Center, Hinxton, UK, Oct 24, 2011

The molecular basis of skeletal disease. Spanish Society for Genetics, Murcia, Spain, Sept 21-23, 2011

Far, far away – Long Range Regulation in Skeletal Development and Disease. Gordon Research Conference Bone & Teeth, Les Diablerets, Switzerland, June 19-24, 2011

The Role of Hox Genes in Limb Development and Bone Formation. 3rd joint Meeting of the European Society of Calcified Tissues & the International Bone and Mineral Society, Athens, Greece, May 7-11, 2011

Digit Development, a Model for Skeletal Morphogenesis. Extracellular Matrix in Health and Disease, Boston, USA, April 14-15, 2011

Phenotypes and the Regulome. Wilhelm Johansen Symposium: The Impact of Deep Sequencing on the Gene, Genotype and Phenotype Concepts, Copenhagen, Denmark, March 21-23, 2011

Defects of Long Range Regulation. Lausanne Genomic Days, Lausanne, Switzerland, Feb 17-18, 2011

Genetics of Limb Malformations. Italian Society for Human Genetics, Florence, Italy, Oct 14-16, 2010

Far reaching consequences - mechanisms and problems of long range control. European Human Genetics Conference 2010, Gothenburg, Sweden, June 12-15, 2010

Chondrogenic Development and Disease Models. Current Concepts in Regenerative Orthopaedics, Düsseldorf, Germany, June 10, 2010

Chondrogenesis and Patterning. International Bone and Mineral Society, IBMS Davos Workshop, Davos, Switzerland, March 14-19, 2010

Genetics of Limb Malformation. 8th World Symposium on Congenital Malformations of the Hand and Upper Limb, Hamburg, Germany, Sept 10-12, 2009

Syndromes with Segmental Progeria as Models for the Ageing Bone and Skin. 9th International Skeletal Dysplasia Society, Boston, USA, July 16-19, 2009

Bone development and dysplasias. 2nd Joint Meeting of the British Society for Matrix Biology and Bone Research Society, London, UK, June 14-16, 2009

A Network of Transcription Factors Regulate Bone Formation in the Limb. Gordon Research Conference Cartilage Biology & Pathology, Les Diablerets, Switzerland, June 7-12, 2009

Awards

Malte Spielmann: *ESHG Young Scientist Award*, European Society of Human Genetics, 2012

Florian Witte: *Tiburtius Award*, Berlin Universities for best PhD thesis, 2011

Uwe Kornak: *Ulmer Dermatologiepreis*, University of Ulm, 2011

Eva Klopocki: *Finalist Trainee Award*, American Society of Human Genetics, 2009

Eva Klopocki: *ESHG Young Scientist Award*, European Society of Human Genetics, 2009

Eva Klopocki: *Vortragspreis,* Deutsche Gesellschaft für Humangenetik, 2009

Appointments of former members of the group

Katrin Hoffmann: *Professorship (W3) for Human Genetics*, Martin-Luther-Universität Halle, 2011

Uwe Kornak, Professorship (W2) for Functional Genetics, Charité – Universitätsmedizin Berlin, 2012

Petra Seemann, Professorship (W1) for Model Systems for Cell Differentiation, Berlin-Brandenburg Center for Regenerative Therapies, 2009

Eva Klopocki, Professorship (W2) for Human Genetics, University of Würzburg, 2012

Peter Robinson, Professorship (W2) for Medical Bioinformatics, Charité – Universitätsmedizin Berlin, 2012

Habilitationen/state doctorates

Katharina Dathe: *Molekulare Ursachen isolierter Handfehlbildungen am Beispiel des BMP-Signalwegs und von SHH*, 2010

Sigmar Stricker: *Molekulargenetik und funktionelle Analyse embryonaler Extremitätenfehlbildungen*, 2010

PhD theses

Sebastian Bauer (Dr. rer. nat.): *Algorithms for Knowledge Integration in Biomedical Sciences*. 2012

Wing Lee Chan (Dr. rer. nat.): *Molecular basis of Gerodermia Osteodysplastica, a premature ageing disorder.* 2012

Gao Guo (Dr. rer. nat.): *Fibrillin-1 and elastin fragmentation in the pathogenesis of thoracic aortic aneurism in Marfan syndrome*. 2011

Jirko Kühnisch (Dr. rer. nat.): *The ANK protein: pathologies, genetics and intracellular function*. 2011

Christian Rödelsperger (Dr. rer. nat.): *Computational Characterization of Genome-wide DNA binding Profiles*. 2011

Wibke Schwarzer (Dr. rer. nat.): *Phenotypic variability in monogenic disorders involving skelatal malformations*. 2010

Michael Töpfer (Dr. med.): *Der Transkriptionsfaktor Osr1 in der Extremitätenentwicklung*. 2010

Aikaterini Dimopoulou (Dr. med.). *Investigation of the genetical basis of autosomal recessive Cutis Laxa*. 2010

Kim Ryong (Dr. rer. medic.): *Assoziationsstudie zur klinischen Variabilität bei Patienten mit dem Nijmegen Breakage Syndrom*. 2010

Wenke Seifert (Dr. rer. nat.): *Pathology of Cohen syndrome: Expression analysis and functional characterization of COH1*. 2010

Florian Witte (Dr. rer. nat.): *Analyse der Ror2-Funktion in vivo und in vitro - Die Ror2 W749X-Maus als Modell für humane Brachydaktylie Typ B.* 2009

Uli Wilkening (Dr. rer. nat.): *Funktionelle Analyse von in der Skelettentwicklung differentiell regulierten Genen*. 2009

Chayarop Supanchart (Dr. med. dent.): *Characterization of an Osteopetrosis mouse model*. 2009

Friederike Kremer (Dr. med.): *Nonsense-mediated mRNA decay in collagen X*. 2009

Anja Brehm (Dr. rer. nat.): *Molekularbiologische Untersuchungen zum Pathogenesemechanismus der Skelettfehlbildungen SYN1 und BDA2 im BMP-Signalweg*. 2009

Pia Kuss (Dr. rer. nat.): *Molekulare Pathologie und Embryologie von Hoxd13-assoziierten Fehlbildungen der Extremitäten*. 2009

Charlotte Wilhelmina Ockeloen (Dr. med.): *Split hand/split foot malformation: determing the frequency of genomic aberrations with molecular-genetic methods.* 2009

Student theses

Sandra Appelt: *Identification and validation of variant calls in a gene panel screen*. Bachelor Thesis, 2012.

Sara Altmeyer: *Analyse des BMP Signalweges in Mausmodellen für humane Brachydaktylien*. Master Thesis, 2011

Stephanie Wiegand: *In vivo Analyse des Fingerphänotyps der Noggin Mausmutante*. Diploma Thesis, 2011

Dominik Jost: *Impairment of receptor-mediated endocytosis in ATP6V=A2 related cutis laxa*. Bachelor Thesis, 2011

Katerina Kraft: *Polarität von Chondrozyten in den Extremitäten der Hoxd13 Mausmutante synpolydactyly homolog (spdh)*. Biotechnol (Dipl.-Ing.), 2011

Denise Emmerich: *Charakterisierung der Interaktion des Golgins GORAB mit den kleinen GTPasen RAB6 und ARF5*. Diploma Thesis, 2010

Denise Rockstroh: *Untersuchung der Interaktion des BMP-Antagonisten Noggin mit der Rezeptor-Tyrosinkinase Ror2*. Diploma Thesis, 2010

Susanne Mathia: *Analyse des Knockdowns von Osr1 und Osr2 in Primärzellkulturen.* Biotechnol (MSc), 2010

Nina Günther: *Funktionelle Charakterisierung aktivierter RAS-Mutanten im Modellsystem Gallus gallus*. Biotechnol (MSc), 2010

Julia Meier: *Die Etablierung eines siRNA-Systems zur funktionellen Analyse der Odd-skipped-related-Gene Osr1 und Osr2 am Beispiel des Hühnerembryos.* Diploma Thesis, 2009

Annika Mahl: *Charakterisierung der Interaktion der Rezeptortyrosinkinase Ror2 mit dem Liganden Noggin*. Diploma Thesis, 2009

Nadine Gladow: *Molekulargenetische Untersuchungen des NSD1-Promotors bei Patienten mit Sotos Syndrom*. Diploma Thesis, 2009

Dajana Lichtenstein: *Deletionsanalyse im NSD1- und FOXL2-Gen bei Patienten mit Sotos- und BPES Syndrom*. Bachelor Thesis, 2009

Bianca Hennig: *Molekulare und funktionelle Charakterisierung von genomischen Aberrationen bei Oatienten mit Fehlbildungen der Extremitäten*. Diploma Thesis, 2009

Otto Schreyer: *Expression regulierter Zinkfingerproteine in der embryonalen Extremitätenentwicklung im Mausmodell für Synpolydaktylie*. Diploma Thesis, 2009

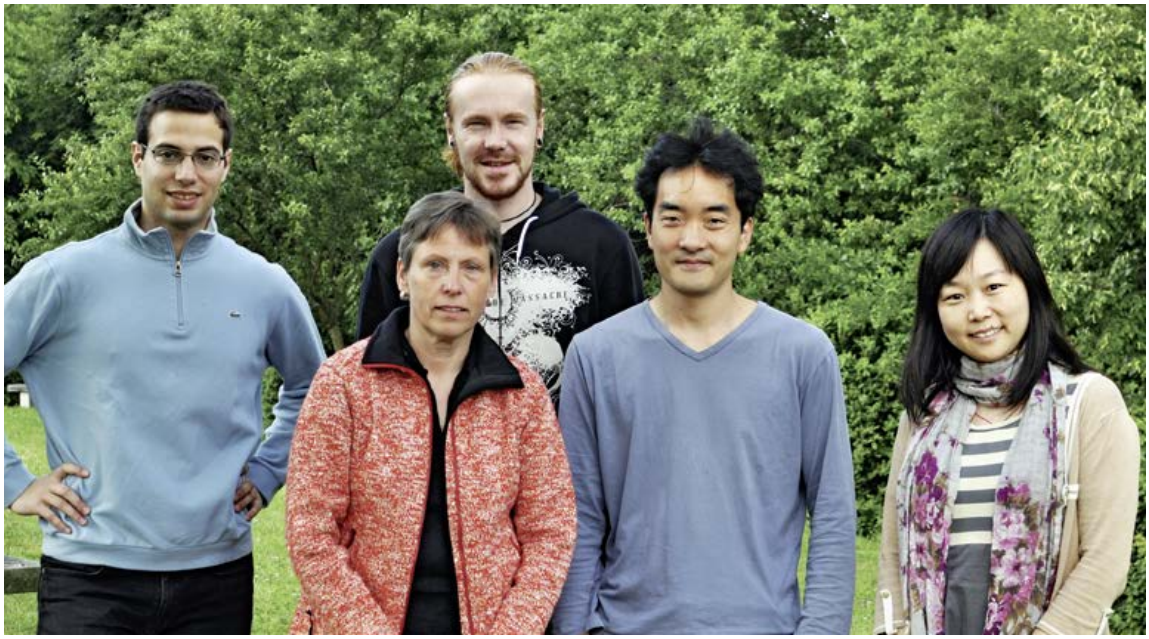
Teaching activities

The IMG together with the research group at the MPIMG runs the entire teaching for Human Genetics at the Charité. Furthermore, we are involved in teaching for students of the Master for Molecular Medicine (Module Human Genetics), as well as Genetics for Bioinformaticians at the Freie Universität.

Otto Warburg Laboratory

Max Planck Research Group Epigenomics

(Established: 12/2011)



Head
Dr. Ho-Ryun Chung (since 12/11)
Phone: +49 (0)30 8413-1122
Fax: +49 (0)30 8413-1960
Email:
ho-ryun.chung@molgen.mpg.de

Secretary of the OWL
Cordula Mancini
Phone: +49 (0)30 8413-1691
Fax: +49 (0)30 8413-1960
Email: mancini@molgen.mpg.de

Scientist
Na Li (since 04/12)

PhD students
Alessandro Mammana*
(since 09/11, IMPRS)
Johannes Helmuth
(since 04/12, IMPRS)

Technician
Ilona Dunkel (since 12/11)

Scientific overview

Introduction

Despite their constant genome sequence cells of multicellular organisms have different morphologies and functions due to the execution of distinct gene expression programs. In this context, transcriptional regulation is very important, as it controls the production rate of mRNAs, which together with the degradation rate determines the steady state level of mRNAs. Transcriptional control depends on the action of transcription factors, which bind to distinct DNA sequences in so-called *cis*-regulatory elements. These binding events in turn influence the recruitment and activity of RNA polymerases.

In eukaryotes, both the binding of transcription factors as well as transcription itself take place in the context of chromatin. The major repeating unit of chro-

* externally funded

matin is the nucleosome, which consists of two copies each of the four core histones H2A, H2B, H3 and H4 (histone octamer) and 147 base pairs of DNA, which is wrapped around the histones in a flat left-handed superhelix. Nucleosomes form every ~200 base pairs along the complete length of the chromosomal DNA. The mere presence of nucleosomes modulates the accessibility to specific DNA sequences, like promoters and other *cis*-regulatory sequences. Furthermore, histones are frequently modified by covalent addition of for example acetyl- or methyl-groups. These histone modifications can influence the stability of the DNA-histone complex and/or may serve as binding sites for protein complexes. Thus, histone modifications on the one hand may be read out during processes acting on chromatin, like transcription, or on the other hand may constitute a memory of past regulatory decisions. Hence, unraveling the *cis*- and *trans*-determinants of nucleosome positioning and the role of histone modifications in transcription are central questions of biology in the post-genomic era.
In my previous work, I have started to address these questions, namely (1) the impact of the DNA sequence on nucleosome positioning and (2) the role of histone modifications in transcription. In my newly founded group I would like to further investigate these questions theoretically and most importantly also experimentally.

Sequence-preferences of the histone octamer

The central assumption in this (still ongoing) project is that the DNA sequence of transcription factor bound genomic regions (referred to as *cis*-regulatory elements) disfavours nucleosome formation, such that once the sequence preferences of histones are known, *cis*-regulatory regions may be readily identifiable from the DNA sequence alone. In a first study, we made use of publicly available data that measured nucleosome positions in yeast by means of chromatin immunoprecipitation followed by sequencing, where mono-nucleosomal-sized DNA fragments were generated by digestion with Micrococcal nuclease (MNase). The analysis of this data revealed two signals that help to predict nucleosome positions determined *in vivo* by this method: (1) an overall enrichment of G or C bases in the nucleosomal DNA and (2) a periodic enrichment of A or T bases and an out of phase enrichment of C or G bases with a period of 10 bases. The analysis also showed that the DNA sequence directs nucleosome formation to a minor but significant degree.
Later we recognized that the preference for GC base pairs (signal 1) may be due to the experimental procedure to obtain nucleosomal DNA fragments, i.e. the digestion of chromatin with MNase. MNase is well known to cut DNA almost exclusively at AT base pairs and also known to cut nucleosomal DNA (albeit to a lesser degree than linker DNA). These properties together with the size-selection step may lead to an artificial increase in GC-rich DNA fragments, because those have a lower probability of internal cuts – or in other words, it is much more likely to recover a GC-rich 150 base pair fragment than an AT-rich one even in the absence of nucleosomes. To test this, we performed an experiment, where we digested naked yeast genomic DNA with MNase, selected ~150 base pair fragments and end-sequenced these using 2nd generation sequencing. In line with our hypothesis, we found that the recovered DNA fragments were indeed GC-rich. Moreover, the corresponding coverage profile was well correlated to the nucleosome occupancy profiles obtained both *in vitro* and *in vivo*, suggesting that these measurements are heavily biased by the sequence preferences of MNase.

Histone modifications and transcription

It is well established that the presence of certain histone modifications is correlated to transcriptional activity. To elucidate the relationship between histone modifications and gene expression in humans, we made use of a publicly available data set that measured the abundance of 38 histone modifications and one histone variant in human CD4+ T-cells. We derived very simple models that relate the levels of histone modifications present at a promoter proximal region to the expression level. The analysis of this data revealed that there is a stable relationship between histone modifications and gene expression, which allows to predict gene expression from the levels of histone modifications in one cell type using a model trained in another cell type, suggesting that we uncovered relationships that are general. Moreover, we showed that only a small number of histone modifications are necessary for accurate prediction. Starting from 39 modifications, we could model gene expression almost as accurately by using only three modifications. An over-representation analysis identified H3K27ac, H2BK5ac, H3K79me1 and H4K20me1 as most important. This result suggests that there is a lot of redundancy in the information contained in the histone modifications and that we possibly have identified modifications that are crucial during the transcriptional process. Finally, we found that the important histone modifications differ in two promoter types, namely CpG island promoters and non-CpG island promoters. While in CpG island promoters H3K27ac (and H2BK5ac) and H4K20me1 turned out to be most important, in non-CpG island promoters H3K4me3 and H3K79me1 were identified. This result suggests that these two promoter types are regulated by different mechanisms.

Outlook

Sequence preferences of the histone octamer

The current state of the art technique to map nucleosomes on a genome-wide scale is based on the assumption that the presence of nucleosomes protects the underlying DNA from the cutting activity of micrococcal nuclease. In this scenario, the number of reads should correspond to the proportion of cells that have a nucleosome covering the protected DNA fragment. However, we have shown that the number of reads is highly correlated in digestions of chromatin and naked DNA *via* the GC content of the underlying DNA fragments. This high correlation can be explained by two scenarios: (i) the histone octamer prefers GC-rich regions and micrococcal nuclease has just the opposite specificity and (ii) the histone octamer has no preference for GC-rich regions and what we observe is a bias due to the experimental technique.

The current method is not able to distinguish between these two scenarios. We have established that the enrichment of GC-rich DNA fragments is due to the size selection step. Thus, we are currently developing a method to map nucleosome positions without size selection. This is accomplished by isolating the cut sites and comparing the cutting frequencies of micrococcal nuclease in digestions of chromatin to naked DNA. Genomic regions bound by the histone octamer will display a reduced cutting frequency compared to the naked DNA sample, while there is no difference in linker regions. In effect this technique shifts the paradigm from indirect evidence of protection by the ability to recover a certain DNA fragment to direct evidence, i.e. a reduced cutting frequency. To critically challenge the resulting nucleosome map, we will use other nucleases with different sequence specificities and compare the resulting maps. However, this technique

requires a much higher coverage than the previous method. Thus, we will use yeast with its small genome as a model system.
Once the data has been generated, we will be able to (i) estimate the probability that a nucleosome is bound to a region, (ii) calculate the underlying potential energies from these probabilities using methods developed in statistical mechanics, and (iii) derive a model of the sequence preferences of the histone octamer independent of the effect of statistical positioning.

Towards a histone code of transcription

Our previous work has shown that histone modifications and transcription are very well correlated. In fact, one can use the levels of a few histone modifications measured at a promoter proximal region to infer the expression level of the corresponding genes. This implies that the histone modifications are tightly connected to the regulatory network that controls the activity of RNA polymerase II. However, we were not able to establish the precise relationships between histone modifications and RNA polymerase II, i.e. whether these histone modifications act up- and/or downstream of Pol II.
To get insight into the changes that take place during the transcription cycle, we plan to measure the levels of histone modifications and RNA polymerase II phosphorylation states after the induction of transcription in a time resolved manner by chromatin immunoprecipitation followed by sequencing. Here, we will make use of the model system *Drosophila melanogaster*. During embryogenesis of Drosophila there is a unique time point, the so-called maternal-to-zygotic transition at which ~1500 genes are induced. This transition occurs in the interphase of cell cycle 14, during which also cellularization takes place, such that the degree of cellularization can be used as a proxy for the time. The embryos will be collected in close collaboration with the group of Bodo Lange. We are currently developing methods to sort Drosophila embryos by morphological characteristics. Since the amount of starting material for chromatin immunoprecipitiation is the main limiting factor we plan to automate the sorting. This will be done by a microfluidic approach coupled to a microscope with a high performance digital camera. The images will be used to classify the embryos and to sort them accordingly. Furthermore, we are exploring means to substitute the conventional microscope by an optofluidic microscope, which can be built at low cost. The latter approach will allow for parallelizing the sorting to get even higher amounts of "pure" material. On the other end we are testing experimental procedures to lower the amount of starting material for chromatin immunoprecipitation. Drosophila (as a model system) allows for testing the effect of the removal of certain chromatin modifiers as well as mutations in the histones themselves on the transcriptional process, such that hypothesis formulated from the time course data can be tested directly.
In case we are not able to gather enough starting material for the chromatin immunoprecipitation experiment in Drosophila, we plan to use human tissue culture cells, which are synchronized in the cell cycle. During mitosis transcription stops and recommences after cell division. Thus, by isolating cells, which have just completed cell division, we could effectively synchronize transcription (at least for some time).
With this approach we will be able to uncover the dynamics of histone modifications during transcription in relationship with changes in the phosphorylation status of RNA polymerase II. Thus, we will be able to unravel the cause-effect relationship between histone modifications and transcription, which in the long run will establish a histone code of transcription.

German contribution to the International Human Epigenome Consortium

Our group is part of a BMBF (Federal Ministry for Education and Research) -funded consortium entitled "Deutsches Epigenom Programm – DEEP". Our part in this project is the generation of 34 histone modification maps for cells involved in inflammatory diseases and the downstream analysis of the data as well as the integration with other data types such as DNase I hypersensitivity, DNA methylation etc., together with the group of Martin Vingron at the institute.
For the data generation, we will implement quality controlled standard operation procedures for cell-type dependent chromatin isolation, chromatin immunoprecipitation and sequencing. Chromatin immunoprecipitation will be performed by a ChIP-robot to ensure maximal reproducibility independent of the operator. The same robot will also prepare the libraries for sequencing. Sequencing will be performed in collaboration with the inhouse sequencing unit headed by Bernd Timmermann.
The whole process of data generation will be monitored by an expert bioinformatician. The main emphasis will lie on the critical evaluation of data quality and reproducibility. Furthermore, we will implement (together with our project partners) methods to transfer the data to the data center in Heidelberg at the DKFZ. Finally, we will participate in the downstream analysis and data integration (i) to identify epigenetic markers for certain disease states and (ii) to unravel epigenetic mechanisms underlying the emergence of the disease state.

Cooperation within the institute

Within the institute, the Epigenomics group closely cooperates with the following people and their groups: Martin Vingron, Sebastiaan Meijsing, Bernd Timmermann, Bodo Lange, Peter Arndt.

Cooperation outside the institute

Outside the MPIMG, we cooperate with the following labs:

- Ann Ehrenhofer-Murray, Universität Duisburg-Essen, Germany
- Herbert Jäckle, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany
- Thomas Manke & Thomas Jenuwein, Max Planck Institute of Immunobiology and Epigenetics, Freiburg, Germany
- Karolin Luger, Colorado State University, USA
- Alexander Bolshoy, Haifa University, Israel

General information

Complete list of publications (2009-2012)

2012

Heise F, Chung HR, Weber JM, Xu Z, Klein-Hitpass L, Steinmetz LM, Vingron M, Ehrenhofer-Murray AE (2012). *Genome-wide H4 K16 acetylation by SAS-I is deposited independently of transcription and histone exchange*. Nucleic Acids Res 40:65–74

Sun R, Love MI, Zemojtel T, Emde AK, Chung HR, Vingron M, Haas SA (2012). *Breakpointer: using local mapping artifacts to support sequence breakpoint discovery from single-end reads*. Bioinformatics 28:1024–1025

2010

Chung HR, Dunkel I, Heise F, Linke C, Krobitsch S, Ehrenhofer-Murray AE, Sperling SR, Vingron M (2010). *The effect of micrococcal nuclease digestion on nucleosome positioning data*. PLoS ONE 5:e15754

Karlić R, Chung HR, Lasserre J, Vlahoviček K, Vingron M (2010). *Histone modification levels are predictive for gene expression*. Proc Natl Acad Sci USA 107:2926–2931

Löhr U, Chung HR, Beller M, Jäckle H (2010). *Bicoid - morphogen function revisited*. Fly (Austin) 4

2009

Chung HR, Vingron M (2009). *Comparison of sequence-dependent tiling array normalization approaches*. BMC Bioinformatics 10:204

Chung HR, Vingron M (2009). *Sequence-dependent nucleosome positioning*. J Mol Biol 386:1411–1422

Löhr U, Chung HR, Beller M, Jäckle H (2009). *Antagonistic action of Bicoid and the repressor Capicua determines the spatial limits of Drosophila head gene expression domains*. Proc Natl Acad Sci USA 106:21695–21700

Zemojtel T, Kielbasa SM, Arndt PF, Chung HR, Vingron M (2009). *Methylation and deamination of CpGs generate p53-binding sites on a genomic scale*. Trends Genet 25:63–66

Invited plenary lectures

Nucleosome positioning and histone octamer sequence preferences, Nucleosome positioning, chromatin structure and evolution, Haifa, Israel, Mai 2012

Student theses

Johannes Helmuth: *Statistical Sequence Analysis and Epigenetic Characterization of Human Promoters*, Diploma thesis, University of Jena, 2012

Teaching activities

Participation in lectures: Methoden der Genetik und Molekularbiologie, Freie Universität Berlin, each semester since 2009

Lecture: Functional Genomics, Freie Universität Berlin, winter term 2011/2012

Otto Warburg Laboratory

Minerva Group Neurodegenerative disorders

(Established: 09/2008)

Head
Dr. Sylvia Krobitsch
Phone: +49 (0)30-8413-1351
Fax: +49 (0)30-8413-1960
Email: krobitsc@molgen.mpg.de

Secretary of the OWL
Cordula Mancini
Phone: +49 (0)30 8413-1691
Fax: +49 (0)30 8413-1960
Email: mancini@molgen.mpg.de

Scientists
Franziska Welzel*
(11/11-09/12, part time)
Linda Hallen* (09/10-03/11)

PhD students
Gunnar Seidel* (since 02/11)
Christian Kähler* (since 05/09)
Christian Linke* (since 04/08)
Anja Nowka* (02/08-12/11)
Franziska Welzel (07/06-10/11)
Linda Hallen (03/06-08/10)

Master students
Anika Günther (since 05/12, guest)
Cathleen Drescher (since 04/12, guest)

Technician
Silke Wehrmeyer* (11/04-04/12)

Scientific overview

Introduction

Human life expectancy is steadily rising in industrialized western countries and as a result fatal late-onset neurodegenerative disorders, such as Alzheimer's disease, Parkinson's disease, or the polyglutamine related diseases, are among the leading causes of disability and death representing one of the major challenges of today's modern medicine. Millions of people worldwide suffer from these devastating disorders or are at risk, and a marked rise in the economic and social burden caused by these disorders will be noticed over the upcoming decades. Even though these diseases are quite common, the mechanisms responsible for their pathologies are in most cases still poorly understood and effective preventative therapies are currently not at hand. For the heritable forms of these neurodegenerative disorders, linkage studies have led to the discovery of the causative genes. Current knowledge of the underlying molecular mechanisms accountable for the observed neurodegenerative processes was gained mainly from studying

* externally funded

inherited disease variants, and these resulted in the identification of genetic and metabolic factors modulating disease onset and progression. Of note, similarities in the clinical and neuropathologic features have suggested that neurodegenerative diseases may share similar mechanisms of pathogenesis related to abnormal protein folding, aggregation, cellular dysfunction and cell death. In consequence, a comprehensive characterization of the molecular mechanisms implicated in the clinical heterogeneity of specific neurodegenerative disorders should help in defining the complete picture of potential pathomechanisms.
The main research interest of my group is to elucidate molecular mechanisms contributing to neurodegenerative processes in the polyglutamine disorder spinocerebellar ataxia type 2 (SCA2) and whether and how these pathways can be correlated to other polyglutamine or neurodegenerative disorders, such as spinocerebellar ataxia type 1 (SCA1) and amyotrophic lateral sclerosis (ALS), on different cellular levels by combining yeast genetics, humanized yeast models, and functional genomic approaches. Moreover, we are interested in studying the biology of stress granules and P-bodies, central self-assembling structures regulating mRNA metabolism, and their relevance in age-related human disorders including cancer.

Functional characterization of ataxin-2 proteins

To gain insight into the cellular function of ataxin-2 (ATXN2), the disease-causing protein in SCA2, we have performed comprehensive protein-protein interaction studies using the yeast-2-hybrid system. These studies revealed that ATXN2 is found in association with a number of proteins implicated in the cellular mRNA metabolism. Further analyses showed that ATXN2 and a number of its protein interaction partners are components of stress granules, cellular sites assembling in mammalian cells as response to specific cellular stresses that are central for regulating and controlling mRNA degradation and translation. In particular, we were interested in investigating the interaction between ATXN2 and the splicing factor FOX-2. The rationale for this was based on the finding that FOX-2 is part of a main protein interaction hub in a network related to human inherited ataxias. In addition, we discovered that the *SCA2* gene bears two putative FOX-binding sites ~ 30-100 nucleotides downstream of exon 18 in the ATXN2 transcript, suggesting that FOX-2 could potentially be involved in ATXN2 pre-mRNA splicing. RNAi experiments revealed that this splicing event indeed depends on FOX-2 activity, since reduction of FOX-2 levels led to an increased skipping of exon 18 in ATXN2 transcripts. To relate this finding to the pathogenesis of SCA2, we will investigate in the future perspective, whether mutant ATXN2 has an impact on FOX-2 splicing activity in general and particularly on ATXN2 transcripts *per se,* and whether and how alterations in ATXN2 transcripts and their cellular consequences affect SCA2 pathogenesis. In this line, we also discovered that the localization and splicing activity of FOX-2 is affected in the presence of nuclear ataxin-1 inclusions, a pathological hallmark in SCA1. Most striking, we observed that splicing of ATXN2 transcripts is affected in the presence of these nuclear ataxin-1 inclusions as well. Since ATXN2 has been shown to modulate SCA1 pathogenesis, it is quite tempting to speculate that alterations in ATXN2 transcripts and their cellular consequences could affect SCA1 pathogenesis. Therefore, further insight into the cellular function of different ATXN2 splice variants and their regulation, and whether and how this relates to mechanisms underlying SCA1, will be an interesting aspect in the future.

Since valuable informations about molecular mechanisms contributing to disease pathogenesis in neurodegenerative disorders have been gained through studying the cellular function of paralogs of neurodegenerative proteins, we also included the ATNX2 paralog, termed ataxin-2-like, in our studies. First, we explored whether an overlap between our generated ATXN2 protein network and ataxin-2-like exists. These comparative yeast-2-hybrid analyses revealed that some interactions are common between both proteins (unpublished data). In this perspective, we were interested in further analyzing a potential functional overlap between ATXN2 and ataxin-2-like in regard to cellular RNA metabolism. Interestingly, we discovered that alterations in the intracellular concentration of ataxin-2-like affect the formation of stress granules and P-bodies, as reported earlier for ATXN2 (unpublished data). Current research addresses whether and how posttranslational modifications of ataxin-2-like are implicated in these processes.
In addition to our functional analyses of ATXN2 concerning the cellular mRNA metabolism, emphasis was laid on the identified Yeast-2-Hybrid (Y2H) interaction between ATXN2 and the KRAB-containing zinc-finger transcriptional regulator, ZBRK1. Interestingly, aberrant interactions between polyglutamine proteins and transcriptional regulators have been found in respective cell culture, animal models and in the brains of patients indicating that perturbation of transcription frequently results in neuronal dysfunction in polyglutamine disorders. These functional studies revealed that ZBRK1 overexpression increased ATXN2 levels, whereas interference on transcriptional and protein levels of ZBRK1 yielded in reduced ATXN2 levels, suggesting that a complex comprising ZBRK1 and ATXN2 regulates *SCA2* gene transcription. Moreover, a bioinformatic analysis utilizing the known ZBRK1 consensus DNA binding motif revealed ZBRK1 binding sites in the *SCA2* promoter, and these predicted sites were experimentally validated demonstrating that *SCA2* gene transcription is controlled by a ZBRK1/ATXN2 complex. Moreover, we discovered that *SCA2* gene transcription is significantly reduced in colon tumours possessing low *ZBRK1* transcripts. Thus, our results provided first evidence that ATXN2 acts as a co-regulator of ZBRK1 activating its own transcription, thereby representing the first identified ZBRK1 co-activator.

Analysis of TDP-43 toxicity and aggregation properties

The TAR DNA-binding protein (TDP-43), which is implicated in transcription, splicing and mRNA stability, has been described as component of inclusions of patients with a variety of neurodegenerative diseases, such as frontotemporal lobar degeneration, ALS, Alzheimer´s disease and Lewy-body disorders. Most striking, ATXN2 intermediate-length polyglutamine expansions were recently associated with increased risk for ALS. Moreover, abnormal ATXN2 localization was detected in ALS patients, whereas TDP-43 mislocalization was observed in SCA2 patients. In addition, exploiting a humanized yeast model for ALS,

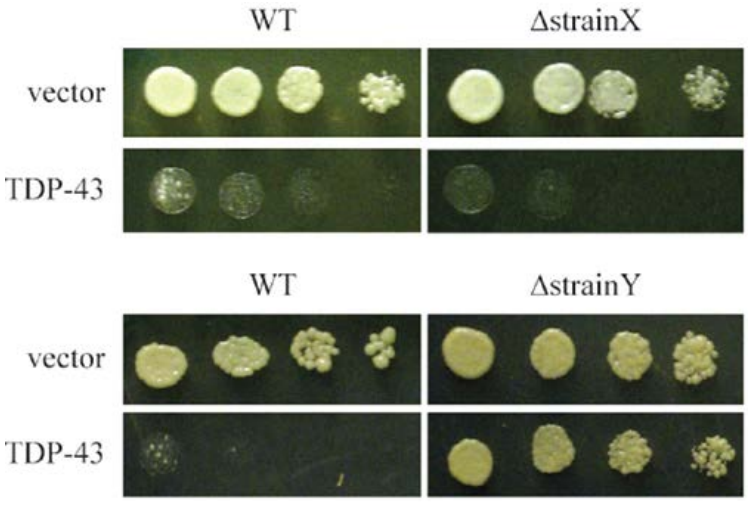


Figure 1: Analysis of TDP-43 toxicity in yeast. Expression of TDP-43 reduces growth of wild type (WT) cells. This TDP-43-induced toxicity is enhanced in yeast deletion strain X whereas it is abrogated in strain Y.

the yeast ATXN2 homolog Pbp1, a stress granule component, was shown to modify TDP-43 toxicity. Together these findings make research on TDP-43 function, its pathological abnormalities and the TDP-43/ATXN2 interaction a very promising task. In this light, we utilized a humanized yeast model to investigate whether other components of stress granules or P-bodies, which are central sites for mRNA storage or degradation, influence TDP-43 toxicity and its aggregation properties. Indeed, first analyses led to the identification of two yeast deletions strains, in which TDP-43 toxicity and aggregation are altered compared to wild type cells. Currently, further experiments are performed in yeast to corroborate these findings. Moreover, experiments are performed to analyze whether these results can be assigned to the human system. In addition, we are also investigating the cellular consequences of the TDP-43/ATXN2 interaction in mammalian cell lines.

Identification of genes causing early onset Alzheimer´s disease

In this project, we aimed in collaboration with Dr. Lars Bertram to identify novel early-onset Alzheimer´s disease (AD) genes and functionally characterize their respective gene products. The comprehensive mutational screening of AD patient DNA material was performed in Dr. Bertram's group. This approach resulted in the identification of a promising candidate gene, for which first functional studies were performed in my group. In addition, comparative yeast-2-hybrid studies revealed that proteins of the ataxin-2 network, which are implicated in transcriptional processes also interact with the amyloid precursor protein (APP) suggesting a potential function in AD as well.

Cell cycle regulation and neurodegenerative processes

The dysregulation of genes that are implicated in the control of cell cycle progression and DNA repair contribute to the degeneration of post-mitotic neurons under certain conditions. Strong evidence has been provided that loss of cell cycle control is associated with neurodegeneration and that re-entry into the mitotic cell cycle occurs before substantial brain pathology can be observed. We were addressing this issue by performing a systems biology approach in which we were studying particular cell cycle aspects in the yeast *S. cerevisiae.* The outcome is currently correlated to neurodegenerative disorders by investigating whether expression levels/localization/protein-protein interactions of particular cell cycle proteins are affected by the expression of mutant huntingtin, the disease-causing protein in Huntington´s disease. To this point, we have identified two cell-cycle specific transcription factors that seem to affect the aggregation properties of mutant huntingtin in yeast. Moreover, we observed that the expression of mutant huntingtin modulates protein-protein interactions controlling cell cycle. Currently, we are investigating whether these findings can be assigned to the human system.

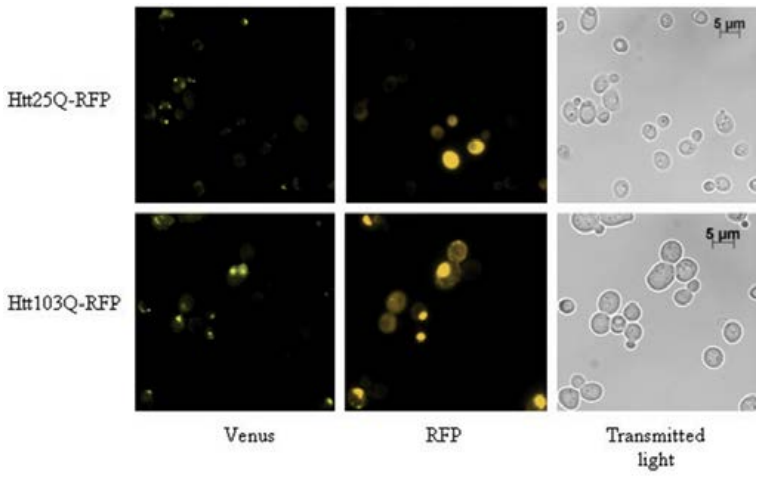


Figure 2: Expression of Htt103Q influences protein-protein interactions controlling cell cycle. Wild type yeast cells expressing cell cycle regulators fused to the N- and C- terminal part of the yellow fluorescent protein (Venus) and Htt25Q-RFP or Htt103Q-RFP were analyzed for bimolecular fluorescence complementation (BiFC).

Relevance of stress granules and P-bodies in cancer

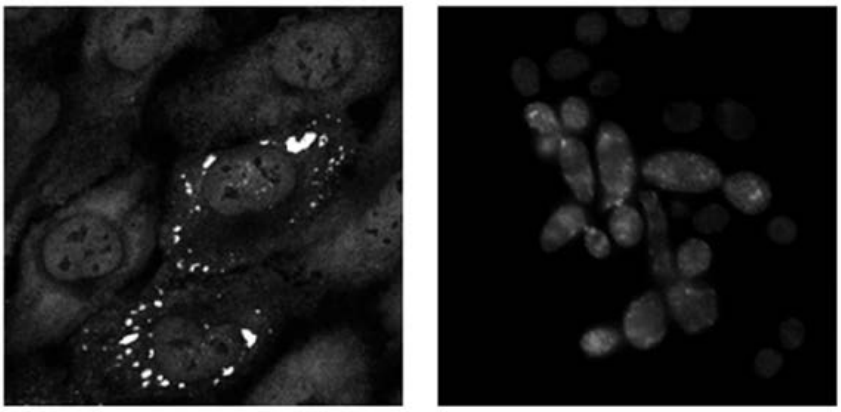


Figure 3: Treatment of mammalian cells (left panel) or yeast cells (right panel) with respective chemotherapeutic agents induces the formation of stress granules, which were visualized by the stress granule marker proteins TIA-R (left panel) and Pab-GFP (right panel).

In cancer therapy clinical applications of chemotherapeutics are often limited due to drug resistance for which the underlying mechanisms are not completely understood. Recently, stress granules have been linked to apoptotic processes and to cellular pathways contributing to chemotherapy resistance in cancer treatment. To dissect the underlying mechanisms in more detail, we are currently exploiting yeast genetic approaches and functional yeast studies. In particular, we have performed global yeast screens utilizing the yeast deletion strain library to identify strains sensitive for particular chemotherapeutics. Of note, these unbiased yeast screens identified candidate genes implicated in ribosomal function, tRNA modification, transport, and processing of mRNAs, amongst others. Moreover, some gene products are either components of stress granules or P-bodies *per se* or are involved in mediating interplay between stress-activated pathways and apoptosis. Remarkably, we observed that some chemotherapeutics cause formation of stress granules and P-bodies in yeast as well as in mammalian cells. Furthermore, stress granule formation was increased in mammalian cells exposed to stress, accompanied by changes in stress granule morphology. Therefore, it is likely that formation of stress granules might impair apoptotic pathways, thus counteracting cytotoxic drug effects opening up novel perspectives for cancer treatments.

Future perspectives

We will continue our efforts in understanding disease protein functions and mechanisms contributing to neurodegenerative processes with central focus on the cellular stress response. For this, we will further analyze protein-protein interactions and investigate their significance in health and disease. Studying cellular factors and mechanisms important for stress granule assembly/function, and how and whether modulation of stress granule composition/function is subject to pathomechanisms underlying age-related human disorders, is of central interest in the future as well.

Cooperation within the institute

Within the institute, the Neurodegenerative Disorders group closely cooperates with the following people and their groups: Zoltán Konthur, Michal Schweiger, Lars Bertram, Holger Klein/Martin Vingron, and Jörg Isensee/Tim Hucho.

Cooperation outside the institute

Outside the MPIMG, we cooperate with the following labs:

- PD Dr. Stefan Kindler, University of Hamburg, Germany
- Dr. Karl Skriner, Charité, Department of Rheumatology and Clinical Immunology, Berlin, Germany
- Dr. Matteo Barberis/Prof. Dr. Edda Klipp, Humboldt-University, Berlin, Germany
- Dr. Sarah Stricker, Charité, Berlin, Germany
- Dr. Thomas Meinel, Structural Bioinformatics Group, Institute for Physiology, Charité, Berlin, Germany
- Prof. Dr. Georg Auburger, Dept. of Neurology, Goethe University Medical School, Frankfurt, Germany

General information

Complete list of publications (2009-2012)

2012

Barberis M, Linke C, Adrover MA, Lehrach H, Krobitsch S, Posas F, Klipp E (2012). *Sic1 potentially regulates timing and oscillatory behaviour of Clb cyclins.* Biotechnol Adv 30:108-130

Welzel F, Kaehler C, Isau M, Hallen L, Lehrach H, Krobitsch S (2012). *Fox-2 dependent splicing of ataxin-2 transcript is affected by ataxin-1 overexpression*. PLoS One 7(5):e37985

2011

Gostner JM, Fong D, Wrulich OA, Lehne F, Zitt M, Hermann M, Krobitsch S, Gastl G, Spizzo G (2011). *EpCAM Overexpression Is Associated with Downregulation of Wnt Inhibitors and Activation of Wnt Signalling in Human Breast Cancer Cell Lines.* BMC Cancer 11(1):45

Hallen L, Klein H, Stoschek C, Wehrmeyer S, Nonhoff U, Ralser M, Wilde J, Röhr C, Schweiger MR, Zatloukal K, Vingron M, Lehrach H, Konthur Z, Krobitsch S (2011). *The KRAB-containing zinc-finger transcriptional regulator ZBRK1 activates SCA2 gene transcription through direct interaction with its gene product ataxin-2*. Hum Mol Genet 20 (1):104-114

Kerick M, Isau M, Timmermann B, Sueltmann H, Herwig R, Krobitsch S, Schaefer G, Verdorfer I, Bartsch G, Klocker H, Lehrach H, Schweiger MR (2011). *Targeted High Throughput Sequencing in Clinical Cancer Settings: Formaldehyde fixed-paraffin embedded (FFPE) tumor tissues, input amount and tumor heterogeneity.* BMC Medical Genomics 4:68

Krobitsch S, Kazantsev AG (2011). *Huntington's Disease: from Molecular Basis to Therapeutic Advances*. Int J Biochem Cell Biol 43(1): 20-24

Meinel T, Schweiger MR, Ludewig HA, Chenna R, Krobitsch S, Herwig R (2011). *Ortho2ExpressMatrix--a web server that interprets cross-species gene expression data by gene family information.* BMC Genomics 12:483

2010

Bourbellion J, Orchard S, Benhar I, Borrebaeck C, de Daruvar A, Dübel S, Frank R, Gibson F, Gloriam D, Haslam N, Humphrey-Smith I, Hust M, Junker D, Koegl M, Konthur Z,

Korn B, Krobitsch S, Muyldermans S, Nygren PA, Palcy S, Polic B, Rodriguez H, Sawyer A, Schlapshy M, Snyder M, Stoevesandt O, Taussig M, Templin M, Uhlen M, van der Maarel S, Wingren C, Hermjakob H, Sherman D (2010). *Minimum information about a protein affinity reagent (MIAPAR)*. Nature Biotechnol 28 (7): 650-653

Chung HR, Dunkel I, Heise F, Linke C, Krobitsch S, Ehrenhofer-Murray AE, Sperling S, Vingron M (2010). *The effect of MNase on nucleosome positioning data*. PLoS ONE 5 (12):e15754

2009

Parkhomchuk D, Borodina T, Amstislavskiy V, Banaru M, Hallen L, Krobitsch S, Lehrach H, Soldatov A (2009). *Transcriptome analysis by strand-specific sequencing of complementary DNA*. NAR 37(18):e123

Habilitation / State doctorate

Sylvia Krobitsch: *Identifizierung von zellulären Mechanismen bei der Huntington-Krankheit und der Spinozerebellaren Ataxie Typ 2*. University of Hamburg, 2010

PhD theses

Christian Linke: *Identification of novel mechanisms controlling cell cycle progression in S.cerevisiae.* Freie Universität Berlin, 2012

Anja Nowka: *Identifikation von potentiellen Resistenzmechanismen gegenüber Camptothecin*. Freie Universität Berlin, 2012

Franziska Welzel: *Untersuchungen zur Rolle von Fox-2 in den Spinozerebellären Ataxien Typ 2 und Typ 1*. Freie Universität Berlin, 2011

Linda Hallen: *Spinozerebelläre Ataxie Typ 2: Untersuchungen zur Rolle von Ataxin-2 in der transkriptionellen Regulation*. Freie Universität Berlin, 2010

Student theses

Artemis Fritsche: *Untersuchungen zur Rolle der Protein-Kinase Ataxia telangiectasia mutated (ATM) in der Spinozerebellären Ataxie Typ 2*. Master Thesis, Freie Universität Berlin, 2012

Judith Hey: *Analysen zur Rolle zellzyklusspezifischer Transkriptionsfaktoren sowie des Sir2-Proteins bei Chorea Huntington*. Bachelor Thesis, Beuth Hochschule für Technik, 2012

Marcel Schulze: *Untersuchungen zur Interaktion von proteolytischen APP-Produkten und dem Transkriptionsrepressor ZBRK1*. Diploma Thesis, Freie Universität Berlin, 2010

Markus Terrey: *Untersuchung zur Rolle des Proteins Ataxin-2-like bei der zellulären Stressantwort*. Bachelor Thesis, Fachhochschule Gelsenkirchen/Recklinghausen, 2010

Fadel Arnaout: *Untersuchungen zur Rolle des RNA-Splicing Faktors NS1BP in neurodegenerativen Erkrankungen*. Diploma Thesis, Freie Universität Berlin, 2010

Tonio Schütze: *Identifizierung und Validierung funktioneller Intrabodies zur Inhibition von Protein-Protein-Wechselwirkungen bei SCA2*. Diploma Thesis, Freie Universität Berlin, 2010

Clara Schäfer: *Funktionelle Analyse zur Rolle des A2BP1-Proteins bei neurodegenerativen Erkrankungen*. Diploma Thesis, Freie Universität Berlin, 2010

Melanie Isau: *Untersuchungen zur zellulären Funktion des Proteins Ataxin-2*. Diploma Thesis, Freie Universität Berlin, 2009

Christian Kähler: *Funktionelle Charakterisierung von Ataxin-2-like: Untersuchungen zur Rolle des Ataxin-2-like Proteins im zellulären mRNA-Metabolismus*. Diploma Thesis, Freie Universität Berlin, 2009

Susanne Weber: *Identifizierung funktioneller Intrabodies für die Untersuchung von Protein-Protein-Wechselwirkungen in SCA2*. Bachelor Thesis, Technische Universität Braunschweig, 2009

Teaching activities

Practical course: Physikalische Übungen für Pharmazeuten, Universität Hamburg, (WS2008/2009; WS2009/2010)

Single lecture in the series „Gene und Genome: die Zukunft der Biologie“; Freie Universität Berlin (WS2008/2009; WS2009/2010; WS 2010/2011; WS 2011/2012)

In addition, several students have been supervised during their internships in the group.

Guest scientist

Prof. Dr. Susana Castro-Obregon, Instituto de Biotecnologia, Avenida Universidad, Mexico (09/09-08/10)

Otto Warburg Laboratory

Sofja Kovalevskaja Research Group
Long non-coding RNA

(Established: 01/2012)

Head
Dr. Ulf Andersson Ørom (since 01/12)
Phone: +49 (0)30 8413-1664
Fax: +49 (0)30 8413-1960
Email: oerom@molgen.mpg.de

Secretary of the OWL
Cordula Mancini
Phone: +49 (0)30 8413-1691
Fax: +49 (0)30 8413-1960
Email: mancini@molgen.mpg.de

PhD students
Anne Musahl (since 04/12)
Dubravka Vicicevic (since 04/12)

Scientific overview

Research concept

Advances in high-throughput sequencing, combined with genome-wide mapping of chromatin modification signatures, have resulted in the identification of a large number of experimentally supported transcriptionally active long non-coding RNAs (ncRNAs) in multiple experimental systems. Through these sequencing efforts thousands of long ncRNAs displaying tissue specific expression have been identified. Long ncRNAs have been described in processes of gene silencing such as X-inactivation, imprinting and dosage compensation. Recent large-scale studies have demonstrated that long-range transcriptional activation is another important function of long ncRNAs in mammals. The mechanisms of long ncRNA regulation are starting to emerge from pioneering work, showing the role of long ncRNAs in epigentic control, long transcriptional regulation and progression of disease. Additional, recent systems scale approaches have provided evidence of essential involvement of long ncRNAs in regulating complex networks of signaling pathways, and important roles in regulating the p53 pathway.
Long ncRNAs transcribed from enhancers are reported to be a widespread phenomenon in human, and have also been observed for thousands of cases in the mouse. Functional knock-down studies in human tissue culture experiments have shown that these enhancer-derived long ncRNAs are responsible for the activating function of several transcriptional enhancers previously thought to work exclusively at the DNA level. These observations reveal a novel aspect of enhancers, placing long ncRNAs in gene regulation in a new light. Enhancers working through a functional transcribed long ncRNA give the possibility of modulat-

ing their function through siRNA-based approaches. Such approaches for regulation of gene expression have great potential, as it allows for manipulation of the regulatory elements controlling a transcriptional network, rather than targeting single genes. The various novel aspects of long ncRNAs provide great potential for furthering the understanding of complex organisms, aspects that can be applied to further the understanding of cellular and molecular biology and are very likely to provide new strategies for therapeutic approaches. A deep and thorough understanding of long ncRNA biology, biogenesis and function is a goal that we should pursue with a massive effort to expand our knowledge of molecular biology and gene regulation to the fullest extent possible. Understanding how long ncRNAs influence the regulation of cellular pathways is one of the research areas expected to impact most of the understanding of gene regulation in the next years.

Scientific methods and achievements/findings

The main aim of the group is to elucidate details of the molecular mechanisms of long ncRNA function in transcriptional regulation in human. The group uses large-scale approaches as RNA-sequencing and Chromosome Conformation Capture sequencing as well as traditional molecular biology and specialized RNA techniques.

Chromosome conformation capture (3C) to identify direct long ncRNA targets

Long ncRNAs have been shown to physically connect the genomic regions of regulated genes with their own genomic locus, mediating activating effects on gene expression through the direct interaction with target genes. We address direct and indirect targeting by long ncRNAs using state of the art approaches for chromosomal conformation, such as 3C and the larger-scale derivative 4C, to identify the genomic loci directly interacting with the long ncRNAs correlating with regulation at the transcriptional level. Using this methodology, in principle, all regions physically associated with long ncRNA genomic regions can be identified.

Identifying protein complexes involved in long ncRNA function

Several protein complexes known to be involved in transcription have been identified to interact with long ncRNAs such as the PRC2 complex, hnRNP-K, WDR5 and the Mediator complex. The effects have been shown to be mutual dependent on both long ncRNAs and protein complexes. Given the largely unexplored nature of long ncRNAs and their regulatory functions, several of the more than 500 identified RNA binding proteins are likely to play important roles in mediating long ncRNA function. We explore the functional interaction of RNA binding proteins with long ncRNAs using an siRNA screening approach coupled with a functional ncRNA reporter system that has been established in the lab. With the successful integration of various models of long ncRNA action into reporter systems, different aspects and interaction partners of long ncRNAs can be addressed using screening for RNA binding proteins.

Delineating the extent of protein-ncRNA complexes.

Using RNA immunoprecipitation (RIP) and ChIP coupled with large-scale sequencing along with RNA-seq, we establish functional interactions between RNA binding proteins and long ncRNAs. RIP-sequencing determines the fraction of the noncoding transcriptome associated to identified RNA binding proteins with a functional role in mediating transcriptional regulation through the complex formation with long ncRNAs. Identifying genomic loci bound and potentially regulated by long ncRNAs in complexes with RNA binding proteins are assessed by ChIP-sequencing of the identified RNA binding proteins.

Perspectives

The group is working towards a global cellular understanding of long non-coding RNAs and how they are involved in regulation of gene expression. A particular goal is to establish the molecular mechanisms of how chromatin structure is regulated by long ncRNAs, and how this can be applied to the enhancer-like functions that have been observed for a class of long ncRNAs.

Cooperation within the Institute

Within the institute, the Long non-coding RNA group closely cooperates with the following people and their groups: Annalisa Marsico, Martin Vingron, Bernd Timmermann.

General information

Complete list of publications (2009-2012)

2012

Ørom UA, Lim MK, Savage JE, Jin L, Saleh AD, Lisanti MP, Simone NL (2012). *microRNA-203 regulates caveolin-1 in breast tissue during caloric restriction*. Cell Cycle 11(7)

2011

Ørom UA, Shiekhattar R (2011). *Noncoding RNAs and enhancers: complications of a long-distance relationship*. Trends Genet 2011:433-9 (review)

Ørom UA, Shiekhattar R (2011). *Long non-coding RNAs and enhancers*. Curr Opin Genet Dev 2011:194-8 (review)

2010

Ørom UA, Derrien T, Guigo R, Shiekhattar R (2010). *Long noncoding RNAs as enhancers of gene expression*. Cold Spring Harb Symp Quant Biol 2010:325-31 (review)

Ørom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, Bussotti G, Lai F, Zytnicki M, Notredame C, Huang Q, Guigo R, Shiekhattar R (2010). *Long noncoding RNAs with enhancer-like function in human cells. Cell* 2010:46-58

Ørom UA, Lund AH (2010). *Experimental identification of microRNA targets.* Gene 2010:1-5 (review)

Invited plenary lectures

GRC conference on chromatin structure and function, Lucca, Italy, 2012

BIMSB Ringvorlesung on RNA biology, Berlin, Germany, 2012

EMBO conference on nuclear organization and function, L'isle sur la Sorgue, France, 2011

Awards

Ulf Ørom is recepient of the 2012 Sofja Kovaleskaja Award from the Alexander von Humboldt Foundation.

Patents

Harel-Bellan A, Echwald SM, Naguibneva I, Lund AH, Ørom UA (2009). *Novel oligonucleotide compositions and probe sequences useful for detection and analysis of microRNAs and their target mRNAs*.

Otto Warburg Laboratory

BMBF-Group
Nutrigenomics and Gene Regulation

(Established: 01/2008)

Head
Dr. Sascha Sauer* (since 01/08)
Phone: +49 (0)30 8413-1661
Fax: +49 (0)30 8413-1960
Email: sauer@molgen.mpg.de

Secretary of the OWL
Cordula Mancini
Phone: +49 (0)30 8413-1691
Fax: +49 (0)30 8413-1960
Email: mancini@molgen.mpg.de

Scientists
Christopher Weidner* (since 07/11)
Chung-Ting Han* (since 06/08)
David Meierhofer* (04/09-02/12)
Vitam Kodelja* (05/08-05/11)

PhD students
Toni Luge* (since 02/12)
Cornelius Fischer* (since 11/11)
Annabell Witzke* (since 11/11)
Radmila Feldmann* (since 09/08)
Susanne Holzhauser* (since 09/08)
Christopher Weidner* (05/07-06/11)

Engineers
Anja Freiwald* (since 01/08)
Magdalena Kliem* (since 01/08)
Claudia Quedenau* (06/08-12/11)

Technician
Beata Lukaschewska-McGreal* (01/09-02/10)

Students
Sylvia Wowro* (07/08-09/12)
Toni Luge* (06/11-01/12)
Stefanie Becker* (01/11-12/11)
Cornelius Fischer* (03/11-10/11)
Anne Geikowski* (02/11-05/11)
Annabell Witzke* (05/11-10/11)
Ilka Limburg* (08/09-01/10)

Scientific overview

Research concept

Many physiological processes are controlled by complex molecular mechanisms. This includes daily environmental factors such as nutrition. In order to prevent health decline and prolong the quality of life we aim to identify causal connections between diet and disease, to increase the acceptance of nutritional intervention for the prevention of disease processes. Functional food and nutraceuticals, i.e. extracts or compounds of edible biomaterials with validated beneficial effects on human health are attracting more and more scientific and public interest. Fur-

* externally funded

thermore, highly potent natural products may be useful to develop pharmaceutical products.
As an externally founded research group in 2008 at the Max Planck Institute for Molecular Genetics, our research group has been exploring health implications of the interaction between nutrition and genomics or the so-called "nutrigenomics". The regulation of genes plays an important role in various molecular processes of metabolic disorders such as insulin resistance or atherosclerosis. One emphasis of our research lies in analysing genome-wide the modulation of gene expression in cellular processes, for example during adipocyte or macrophage cell differentiation. These processes can be significantly influenced through the interaction between genes and naturally occurring compounds.
Consequently, as the second emphasis of our research group, we study the capability and mechanisms of natural products to interact with genes and gene products. In order to identify active natural products, we screened and systematically characterized natural substances derived from small molecule libraries that featured large structural variability.

Scientific methods and achievements / findings

1. Scientific/research achievements

Given worldwide increases in the incidence of metabolic diseases such as obesity, type 2 diabetes or atherosclerosis, alternative approaches for preventing and treating these disorders are required. The nuclear receptors PPARγ (peroxisome proliferator-activated receptor gamma) and liver x receptor alpha (LXRα) play central roles in metabolism; however, current drugs or drug candidates targeting these receptors are characterised by undesirable side effects.

Amorfrutins and LXR ligands

We discovered a family of natural products that bind to and activate specifically PPARγ. These compounds, the amorfrutins, are derived from edible parts of two legumes, *Glycyrrhiza foetida* and *Amorpha fruticosa*. The natural amorfrutins are structurally new powerful anti-diabetics with unprecedented effects for a dietary molecule.

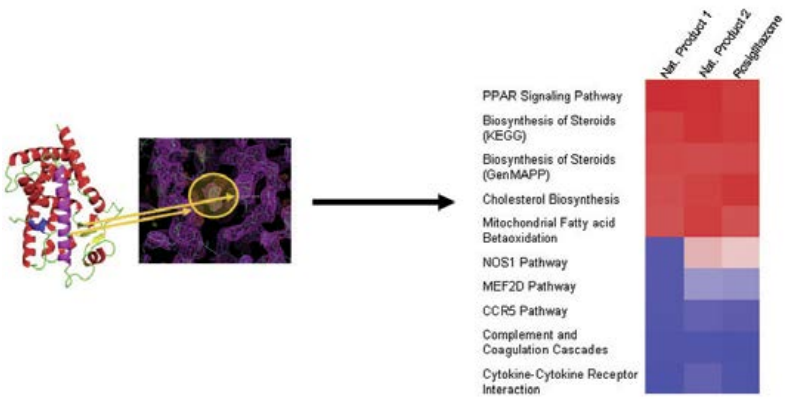


Figure 1: Structure-function relationship of novel selective nuclear receptor agonists

Moreover, we identified a new LXRα ligand, which is subtype-specific - in contrast to any other known LXR ligand. This specific LXRα ligand is in particular active in lipid-loaded foam cells that are involved atherosclerotic plaque formation. This LXR ligand will be explored as a chemical tool as well as for potential drugability.
Our results showed that selective nuclear receptor activation by (diet-derived) ligands constitutes a promising approach to combat metabolic disease.

Transcriptional networks of foam cells

Atherosclerosis is an important global health problem and a leading cause of cardiovascular disease. Adaptation of macrophages to physiological stimuli as lipid overload or elevated levels of cholesterol requires dynamic regulatory molecular networks. We deciphered the LXRα-dependent gene-regulatory architecture of atherosclerotic foam cells, as well as key networks triggered by LXRα-modulation for treating efficiently atherosclerosis, by using integrated genome-wide analysis of LXR-alpha. Functional analyses integrating genetic variation disease association data revealed cholesterol induced disease gene expression and suggest avenues for treating systematically foam cell development and atherosclerotic plaques, for example *via* specific LXRα-activation of the APOC-APOE gene cluster.

Gene regulation during mild stress response

The natural product resveratrol is a widely known molecule because of its reported health-beneficial and striking anti-aging effects. However, the mechanism of action of resveratrol remained essentially elusive. We showed that the beneficial cellular effects of resveratrol can be explained by its chemical degradation leading to reactive oxygen species, subsequent genome-wide remodelling of chromatin and gene regulation of interconnected pathways of cellular defence, thermogenesis, and modulated metabolic profiles. A concerted action of cellular defence mechanisms including genome remodelling may explain mechanistically reported aspects of resveratrol on the cellular and physiological level. Based on our data, we propose a hormesis model of the action of resveratrol. In particular this line of research led to intensified collaboration with the nutritional and cosmetics industry.

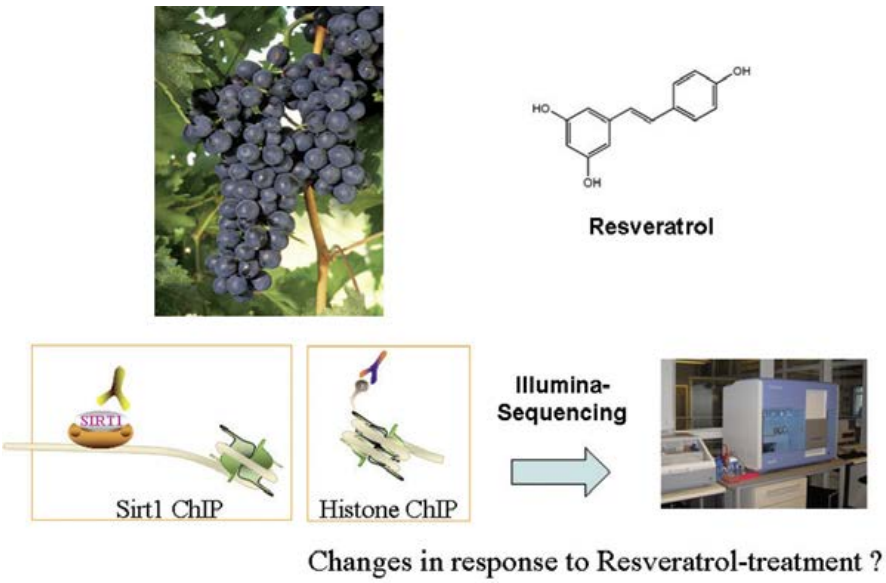


Figure 2: Mechanisms of Sirtuin 1 (SIRT1) activation: Chromatin dynamics triggered by natural products.

Proteomics

We further developed a bunch of proteomics and compound screening pipelines to decipher entire proteome-wide expression changes and post translational modifications, as well as discover new active natural products from diverse compound libraries.

Using SILAC (stable isotope labeling by amino acids in cell culture), or alternatively dimethyl labelling, we detected in treated diabetic mice more than 5000 proteins and including more than 9000 individual phosphorylation sites. Thereby we discovered molecular evidence for physiologically important (side) effects

such as heart failure. We furthermore apply mass spectrometry for protein characterisation to detect bacteria and bacteria-host interaction to understand important interactions as well as inflammatory processes that can be influenced by nutritional intervention. Moreover, we studied on the proteome-wide level protein networks of cell-cell communication, for example between fat cells and macrophages, to get mechanistic insights in the metabolic adaptation of these key cells, which are involved in the development of insulin resistance.
Furthermore, the group has set up the facilities and protocols required for studying systematically gene-regulation processes influenced by histone-modifying enzymes For example, we have developed a novel functional high-throughput mass spectrometry assay to screen and characterise natural products interacting with protein-modifying enzymes such as deacetylases such as sirtuin 1 (Sirt1) or acetyl transferases like p300.

2. Development of Research Infrastructures

a. Externally: The group coordinates the European Sequencing and Genotyping Infrastructure (ESGI, www.esgi-infrastructure.eu), which pools leading European genomics and bioinformatics facilities to provide the larger scientific community with access to new genomic technologies and the latest analytic tools. The aim of ESGI is to enable scientists across all disciplines to use emerging sequencing technologies to decipher the complex functions of genes, without breaking the bank. About 20 external collaborations are currently being coordinated by Sascha Sauer, with a particular focus in the areas of functional genomics of metabolic diseases or cell stress response, and on general mechanisms of gene regulation and chromatin biology.

b. Internally: The mass spectrometry based proteomics pipeline developed for our research questions and the various mass spectrometry methods described above are additionally being used for a number of internal collaborative projects to complement ongoing research in the institute. The postdoc responsible for the MS in our group has meanwhile been promoted by the Institute to head the MPIMG's MS service group.

Cooperation within the institute

Within the institute, the Nutrigenomics / Gene regulation group closely cooperates with Hans Lehrach and his department on a European Sequencing and Genotyping Infrastructure (ESGI), and with Martin Vingron and his department on bioinformatics analyses of 2nd generation sequencing data.

Special facilities/equipment of the group

The Nutrigenomics / Gene regulation group operates the following special equipment:

- Nano-HPLC LTQ Orbitrap XL (EDT) ESI Mass Spectrometer (Thermo) *
- Cap-LC HCT ultra mass spectrometer (Bruker)
- Genome Analyser IIx (Illumina) *

General information

Complete list of publications (2009-2012)

2012

Kliem M, Sauer S (2012). *The essence on mass spectrometry based microbial diagnostics*. Curr Opin Microbiol 15:1–6

Weidner C, de Groot JC, Prasad A, Freiwald A, Quedenau C, Kliem M, Witzke A, Kodelja V, Han CT, Giegold S, Baumann M, Klebl B, Siems K, Müller-Kuhrt L, Schürmann A, Schüler R, Pfeiffer AF, Schroeder FC, Büssow K, Sauer S (2012). *Amorfrutins are potent antidiabetic dietary natural products*. Proc Natl Acad Sci U S A 109(19):7257-62

2011

Freiwald A, Mao L, Kodelja V, Kliem M, Schuldt D, Schreiber S, Franke A, Sauer S (2011). *Differential analysis of Crohn's disease and ulcerative colitis by mass spectrometry*. Inflamm Bowel Dis 17(4):1051-2

Mertes F, Elsharawy A, Sauer S, van Helvoort JM, van der Zaag PJ, Franke A, Nilsson M, Lehrach H, Brookes AJ (2011). *Targeted enrichment of genomic DNA regions for next-generation sequencing*. Brief Funct Genomics 10(6):374-86

2010

Haseneyer G, Stracke S, Piepho HP, Sauer S, Geiger HH, Graner A (2010). *DNA polymorphisms and haplotype patterns of transcription factors involved in barley endosperm development are associated with key agronomic traits*. BMC Plant Biol 10:5

Konrad K, Dempfle A, Friedel S, Heiser P, Holtkamp K, Walitza S, Sauer S, Warnke A, Remschmidt H, Gilsbach S, Schäfer H, Hinney A, Hebebrand J, Herpertz-Dahlmann B (2010). *Familiality and molecular genetics of attention networks in ADHD*. Am J Med Genet B Neuropsychiatr Genet 153B(1):148-58

Sauer S, Kliem M (2010). *Mass spectrometry tools for the classification and identification of bacteria*. Nature Rev Microbiol 8(1):74-82

2009

Baek YS, Haas S, Hackstein H, Bein G, Hernandez-Santana M, Lehrach H, Sauer S, Seitz H. *Identification of novel transcriptional regulators involved in macrophage differentiation and activation in U937 cells*. BMC Immunol 10:18

Darii E, Lebeau D, Papin N, Rubina AY, Stomakhin A, Tost J, Sauer S, Savvateeva E, Dementieva E, Zasedatelev A, Makarov AA, Gut IG (2009). *Quantification of target proteins using hydrogel antibody arrays and MALDI time-of-flight mass spectrometry (A2M2S)*. N Biotechnol 25(6):404-16

Freiwald A, Sauer S (2009). *Phylogenetic classification and identification of bacteria by mass spectrometry*. Nat Protoc 4(5):732-42

Stracke S, Haseneyer G, Veyrieras JB, Geiger HH, Sauer S, Graner A, Piepho HP. *Association mapping reveals gene action and interactions in the determination of flowering time in barley*. Theor Appl Genet 118(2):259-73

Invited plenary lectures

Sequencing out effects of natural products, 3rd Annual Next Generation Sequencing Congress, London, UK, Nov 14-15, 2011

Make hypotheses!, The 4th Paris Workshop on Genomic Epidemiology, Paris, France, May 31, 2011

Mass spectrometry methods for microbiology and nutrition research, Rapid Methods Europe, Noordwijkerhout, The Netherlands, Jan 25, 2011

Modulation of gene regulation processes by small molecules, University Vienna, Austria, May 27, 2010

The mechanisms of action of Sirtuin 1, Human Genome Meeting Conference, Montpellier, France, May 20, 2010

Appointments of former members of the group

David Meierhofer: *Head of Service group Mass Spectrometry,* MPIMG, 03/2012

PhD theses

Christopher Weidner: *Prevention and therapy of type 2 diabetes by selective modulation of the human peroxisome proliferator-activated receptor gamma (PPARγ)*, Freie Universität Berlin, 06/2011

Student theses

Sylvia Wowro: *Wirkmechanismen der Resveratrol-Behandlung in Fibroblasten*, Master Thesis, Beuth Hochschule für Technik, Berlin, 04/2012

Cornelius Fischer, *Genome-wide liver x receptor alpha binding and chromatin accessibility in different macrophage models*, Master Thesis, Freie Universität Berlin, 11/2011

Toni Luge, *Identifikation von Phytoplasma-Transkripten und -Proteinen via RNA-Sequenzierung und Massenspektrometrie*, Master Thesis, Beuth Hochschule für Technik, Berlin, 11/2011

Annabell Witzke, *Cell Biological Studies on Adipocyte-Macrophage Communication*, University of Konstanz (10/2011)

Anne Geikowski, *Molekularbiologische Charakterisierung neuer potentiell anti-arteriosklerotisch wirkender LXRa-Liganden*, Bachelor Thesis, Beuth Hochschule für Technik, Berlin, 08/2011

Ilka Limburg, *Molekulare Charakterisierung der Interaktion von Naturstoffen mit dem nuklearen Rezeptor PPARα*, Master Thesis, Freie Universität Berlin, 02/2010

Teaching activities

- Members of the Nutrigenomics/ Gene regulation group have given the following lectures at the Freie Universität Berlin
- Lecture „Das Buch des Lebens. Historische und moderne Ansätze in der Genomforschung“, SS10
- Lecture „Nutrigenomik und Genregulation“, WS10/11
- Lecture „Nutrigenomik und Genregulation“, SS11
- Practical course „Molekularbiologische Methoden der Nutrigenomik und Chemischen Biologie“, WS11/12, FU Berlin
- Seminar „Molekularbiologische Methoden der Nutrigenomik und Chemischen Biologie“, WS11/12, FU Berlin
- Practical course „Molekularbiologische Methoden der Nutrigenomik und Chemischen Biologie“, SS12, FU Berlin
- Seminar „Molekularbiologische Methoden der Nutrigenomik und Chemischen Biologie“, SS12, FU Berlin

Guest scientists

Dr. Sheng Yu Huang, University Taipeh, China, 04/10 – 09/10

Dr. Lei Mao, Charité - Universitätsmedizin Berlin, Germany, 03/05 – 03/10

Organisation of scientific events

The 4th Paris Workshop on Genomic Epidemiology, Paris, France, May 31-June 2, 2011

Otto Warburg Laboratory

Max Planck Research Group Molecular Interaction Networks

(Established: 06/2007)

Head
Ulrich Stelzl (since 06/07)
Phone: +49 (0)30 8413-1264
Fax: +49 (0)30 8413-1960
Email: stelzl@molgen.mpg.de

Secretary of the OWL
Cordula Mancini
Phone: +49 (0)30 8413-1691
Fax: +49 (0)30 8413-1960
Email: mancini@molgen.mpg.de

Scientists
Jonathan Woodsmith (since 01/11)
Nouhad Benlasfer (since 05/10)
Anna Hegele* (06/07-04/12)
Petra Birth (06/08-12/11)
Reynaldo López-Mirabal* (07/08-06/10)

PhD students
Stefanie Jehle (since 05/12)
Thomas Corwin (since 09/10)
Luise Apelt (since 07/10)
Mareike Weimann (12/07-02/12)
Josphine Worseck* (09/07-11/11)
Atanas Kamburov* (01/10-12/11, part time)
Arndt Grossmann* (07/07-12/11)

Undergraduate students
Federico Apelt (since 11/10, part time)
Ziya Özkan (since 08/09, part time)
Franziska Wachsmuth (10/11-06/12)
Chrysovalantis Sourlis (10/09-06/10)
Sylvia Wowro (07/08-05/09, part time)

Scientific overview

Research concept

A major goal in current genome research is to predict the influence of human genetic variation on disease phenotypes. One idea is that large sequencing endeavors, e.g. whole genome sequencing of multiple individuals or large GWAS studies, will provide enough information to make better predictions for risk, cause, pathogenesis or medication of patients vs. control groups (Figure 1a). Molecular interaction networks, such as protein-protein interaction (PPI) networks, are very

* externally funded

useful for studying genotype to phenotype relationships. Integrative computational analyses that largely rely on molecular interaction data result in a more accurate interpretation of genomic variation but remains probabilistic (Figure 1b). However, we want to go beyond improving statistical predictions more towards specific information about an individual. This means we need to consider different sources of molecular, environmental and behavioral variation in addition to individual genomic sequence information. Thus we need to measure – e.g. at the level of cellular networks – to generate additional molecular information about the individual reflecting this variation. To do so, high quality molecular network information will be a necessity (Figure 1c). Differential network analysis will be informative on sets of decisive molecules such as drivers in cancer or transcriptional "master-regulator" proteins and their connections and then guide measurements to better understand and classify individual phenotypes in model systems and ultimately humans.

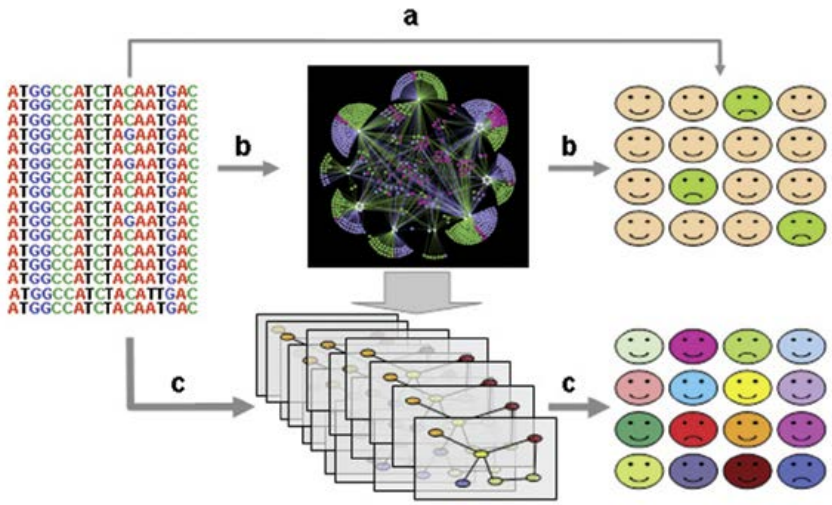


Figure 1: Molecular network information is required to predict genotype – phenotype relationships. To strengthen predictions about the phenotype from genomic information (a) cellular interaction networks will be useful (b) but remain probabilistic over groups of phenotypes. Differential network analyses will discover relevant sets of key molecules that reflect molecular, environmental and behavioral variation and can be measured (c) and will be necessary to predict phenotypes for individual cells /organisms from genomic information.

In the OWL group *Molecular Interaction Networks* we do not work specifically on the interpretation of genetic variation, but aim to provide the network information to better describe disease relevant cellular processes and thus contribute to genotype to phenotype predictions. We work on the generation and analysis of human interaction data and study interaction network dynamics. The later point is particular important as interactome networks are extensively re-wired during a cellular response e.g. during development or during the processing of internal or environmental cues. Differential interaction patterns imply mechanistic changes that are the result of these responses and will thus be most informative when studying genotype to phenotype relationships (Figure 1c).
Specifically, our work aims at i) improving data generation and analysis of human protein-protein interaction networks, ii) integrative approaches to analyze

protein interaction dynamics and iii) experimental approaches to directly investigate conditional protein-protein interactions, such as interactions that e.g. require triggered phosphorylation of one interaction partner mediating the response to changing conditions.

Scientific methods and achievements / findings

Systematic generation of high quality human protein-protein interaction networks

Several studies to systematically map protein-protein interaction (PPI) networks on a large scale have been successful and proven very useful in further studies. Nevertheless, for most species including human only a small fraction of all possible interactions has been mapped today. High quality PPI data collection and independent assessment of data quality remain important tasks.

To provide independent measures of interaction confidence, we have developed a cluster-based method for the assessment of protein-protein interaction confidence and implemented this as a web tool. The method (CAPPIC) exploits the modular network architecture independently of prior parameters or reference sets for confidence scoring of interactions.

In an international collaboration led by the Vidal Lab (Harvard/CCSB, Boston) we have assessed protein interaction data empirically demonstrating that systematic Y2H interaction data, including those generated with our setup are of high precision. The study also revealed that the coverage of the data is low due to relatively low sensitivity of the method. To overcome this limitation, we developed a Y2H-seq approach which enables very high PPI sampling through a second generation short read sequencing readout. Importantly, the method has significantly improved sensitivity and provides a quantitative readout that is indicative of the quality of the PPI information. It will accelerate large-scale interactome mapping efforts.

As Y2H-seq test case, we comprehensively screened proteins involved in methylation and demethylation, i.e. protein methyltransferases and demethylases such as AOF2/LSD1, for interacting partners. Protein methylation of non-histone proteins is a largely unexplored posttranslational modification. We report 523 interactions between 22 methyltransferases or demethylases and 324 interacting proteins. The methyltransferase network is experimentally validated, comprehensively annotated and defines novel cellular roles of non-histone protein methylation. It will thus serve as a major informational resource to the scientific community and is the basis to study methylation dependent protein interactions in the lab in more detail (see below).

Focusing on neurodegenerative diseases, we generated a PPI network connecting proteins implicated in Alzheimer's disease (AD) with the Aloy Lab (IRB, Barcelona). The study suggests novel roles for central proteins in the network that link between oxidative stress, inflammation, and mitochondrial dysfunction in AD. With the Beyer Lab (TU Dresden) a map of human protein interactions was inferred using combined random forest / Bayesian networks to distinguish functional from physical interactions. The map was in part experimentally validated and used to explore the relationships of candidate genes from GWAS of neurodegenerative diseases, such as AD.

Dynamic alterations in protein-protein interaction networks: integrative approaches

The goal here is to reveal differential network states that describe changing cellular processes *in vivo*. Successful analysis of network dynamics through data integration will focus on a specific biological process with interesting dynamic behavior and requires data of very high quality.

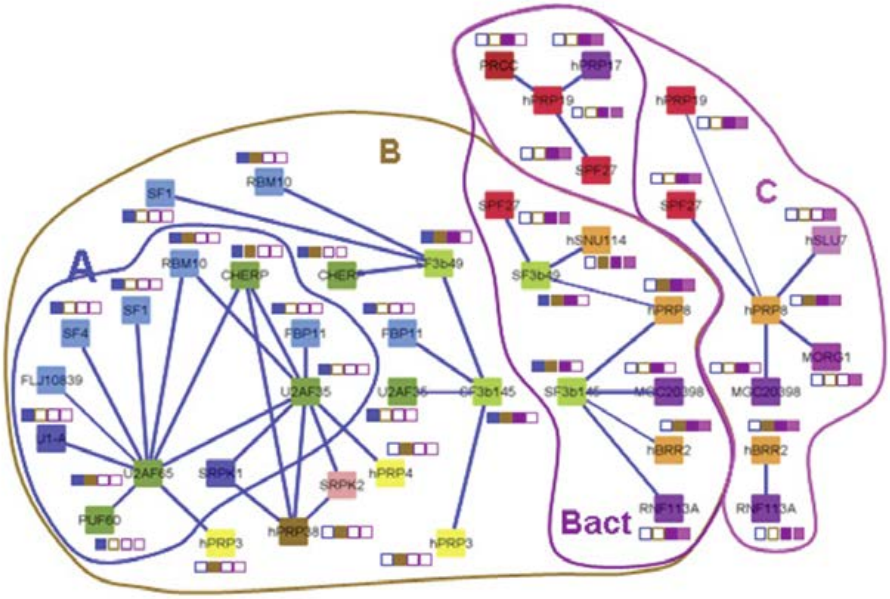


Figure 2: PPI dynamics involving SF3b-complex proteins and hPRP8 (Hegele, Mol Cell 2012). Selected interactions from the (U2AF35,U2AF65), the (SF3b145,SF3b49), the hPRP19 and the hPRP8 modules are shown. Distinct PPI patterns for proteins are suggested for different stages (i.e. A, B, Bact and C complexes) of the spliceosomal assembly cycle.

In a recent study, we focused on PPI dynamics of the splicing cycle. Pre-mRNA splicing is catalyzed by the spliceosome, a highly complex, dynamic and protein rich ribonucleoprotein complex that assembles *de novo* on each intron to be spliced. During spliceosome assembly, activation, catalysis and disassembly, defined large complexes are formed in an ordered, stepwise manner. A data set describing 632 interactions between 200 human spliceosomal proteins was generated including e.g. the first contact sites between the U5 proteins and specific U2-SF3b components at the heart of the spliceosome. We then integrated co-complex purification information from 76 purifications of active spliceosomal complexes with our data and performed PPI clustering. This approach revealed several interesting dynamic PPI patterns with relevance for a better understanding of the splicing cycle. For example, changing PPIs during B to C transition (Figure 2) with one of the most central proteins, hPRP8, are found. Together with interaction competition experiments, these data suggest that during step 1 of splicing, hPRP8 interactions with the SF3b49 protein is replaced by hSLU7, positioning this essential second step factor close to the active site and that the DEAH-box helicases hPRP2 and hPRP16 cooperate through ordered interactions with the G-patch protein GPKOW.

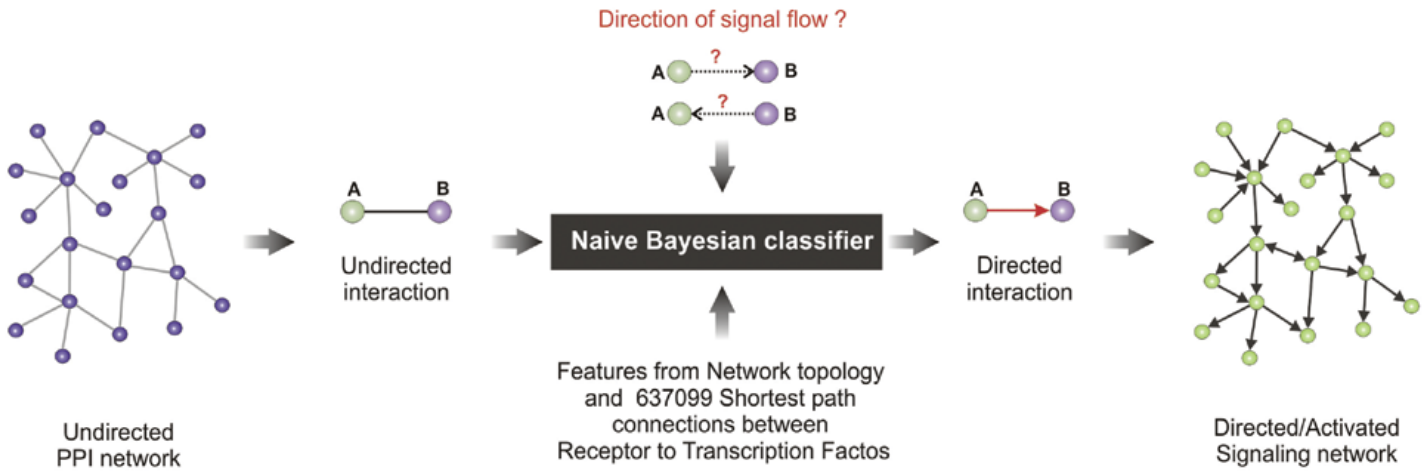


Figure 3: Inferring edge directions from PPI data (Vinayagam, Sci Signal 2011). For each interaction in the undirected PPI network, a naïve Bayesian classifier was used to predict the edge direction from topological network properties as well as shortest PPI paths connecting membrane receptors and transcription factors. An activated signaling network was assembled from all interactions that had a direction assigned.

In another study, performed together with the Wanker Lab (MDC, Berlin, Buch) we identified >2500 PPIs among human proteins broadly involved in cellular signaling. To provide information about how the signal is transmitted through the network, we developed a Bayesian learning strategy to assign direction to the interactions reflecting the potential signal flow among the proteins (Figure 3).
The resulting directed network is a unique resource for various modeling approaches. For example, we used the model to identify previously unknown modulators of the EGF/ERK pathway, of which 18 were validated with cell-based assays. It also enabled us to model EGF-induced protein phosphorylation dynamics. We could correlate *in vivo* phosphorylation dynamics with the output distance from the EGF/ERK pathway in our network resolving global protein phosphorylation events in a time-dependent manner.

In the two projects described, we exemplarily addressed PPI dynamics through combined experimental and computational approaches and successfully modeled how a signal spreads from an activated signaling pathway through a dense PPI network to the very distant proteins as well as identified crucial sites of changing PPI patterns that contribute to the exceptional compositional dynamics (and thus function) of the human spliceosome. In a next step, we want to take a direct experimental approach to analyze alterations of PPI patterns.

Planned developments

Conditional / modification-dependent protein interactions

In ongoing projects, we want to elucidate the role of posttranslational protein modifications (PTMs), such as phosphorylation (P) and methylation, for these dynamic processes and investigate how genetic variations, e.g. SNPs, may change protein-protein interaction patterns.

We have established a modified Y2H setup employing kinases to screen for P-dependent PPIs. We identified a novel P-dependent interaction between ADAP and Nck adaptor molecules that alter adhesion and migration of Jurkat T cells. We also contributed to the identification of a phosphorylation–triggered interaction between neuronal Fez1 and Munc18, mediating axonal transport of Syntaxin. The characterization of isoform-specific and P-dependent protein interactions between tumor suppressor protein NF2 (merlin) and AOF2, EMILIN1 and PIK3R3 suggests novel regulatory loops influencing NF2 conformation and function of the protein.

In a proteome wide approach, we have identified more than 300 novel pY-dependent PPIs that show high specificity with respect to human kinases and interacting proteins. P-dependent interactions are further analyzed in mammalian cell culture systems using e.g. co-immunoprecipitation, protein complementation and functional reporter readouts. The more detailed characterization of selected P-dependent GRB2 and PIK3R3 interactions exemplarily demonstrate how these PPIs are dynamically and spatially constrained to separate simultaneously triggered growth signals which are often altered in oncogenic conditions. Our screening approach is extended to other posttranslational protein modification such as methylation. Methyltransferase-substrate relationships discovered through Y2H-seq mapping of the methyltransferase interactome provide a reliable basis to exploit cellular functions of non-histone protein methylation. Finally, we integrate genetic variation data in our interaction studies and investigate how disease causing missense mutations change protein interaction profiles.

Cooperation within the institute

Within the institute, the Molecular Interaction Networks group cooperates with the following people and their groups: Ralf Herwig, Dept. of Vertebrate Genomics, on computational tools and network algorithms; Sebastiaan Meijsing, Dept. of Computational Molecular Biology, on splice variant specific protein interaction studies of GR; Hermann Bauer, Phillip Grote, and Heiner Schrewe, all from the Dept. of Developmental Genetics, on targeted protein interaction studies for proteins involved in mesoterm formation and of t-complex distorters; with Sascha Sauer, OWL, and David Meierhofer, Mass Spec group, on protein modification analyses, and with Bernd Timmermann (next generation sequencing).

General information

Complete list of publications (2009-2012)

2012

Chua JJ, Butkevich E, Worseck JM, Kittelmann M, Grønborg M, Behrmann E, Stelzl U, Pavlos NJ, Lalowski MM, Eimer S, Wanker EE, Klopfenstein DR, Jahn R (2012). *Phosphorylation-regulated axonal dependent transport of syntaxin 1 is mediated by a Kinesin-1 adapter.* Proc Natl Acad Sci U S A 109(15):5862-7

Hegele A*, Kamburov A*, Grossmann A, Sourlis C, Wowro S, Weimann M, Will CL, Pena V, Lührmann R, Stelzl U (2012). *Dynamic protein-protein interaction wiring of the human spliceosome*. Mol Cell 45(4):567-80

Kamburov A [#], Stelzl U, Herwig R (2012). *IntScore: a web tool for confidence scoring of biological interactions*. Nucleic Acids Res 2012 (40):W140-6

Worseck JM*, Grossmann A*, Weimann M*, Hegele A, Stelzl U (2012). *A stringent yeast two-hybrid matrix screening approach for protein-protein interaction discovery.* Methods Mol Biol 812:63-87

2011

Elefsinioti A, Saraç ÖS, Hegele A, Plake C, Hubner NC, Poser I, Sarov M, Hyman A, Mann M, Schroeder M, Stelzl U, Beyer A (2011). *Large-scale de novo prediction of physical protein-protein association*. Mol Cell Proteomics 10(11):M111.010629

Soler-López M*, Zanzoni A*, Lluís R, Stelzl U, Aloy P (2011). *Interactome mapping suggests new mechanistic details underlying Alzheimer's disease*. Genome Res 21(3):364-76

Vinayagam A*, Stelzl U*,#, Foulle R, Plassmann S, Zenkner M, Timm J, Assmus HE, Andrade-Navarro MA, Wanker EE[#](2011). *A directed protein interaction network for investigating intracellular signal transduction*. Sci Signal 4(189):rs8

2010

Sylvester M, Kliche S, Lange S, Geithner S, Klemm C, Schlosser A, Grossmann A, Stelzl U, Schraven B, Krause E, Freund C (2010). *Adhesion and degranulation promoting adapter protein (ADAP) is a central hub for phosphotyrosine-mediated interactions in T cells*. PLoS One 5(7):e11708

Vinayagam A, Stelzl U, Wanker EE (2010). *Repeated two-hybrid screening detects transient protein-protein interactions*. Theor Chem Acc 125(3-6):613-19

2009

Palidwor GA, Shcherbinin S, Huska MR, Rasko T, Stelzl U, Arumughan A, Foulle R, Porras P, Sanchez-Pulido L, Wanker EE, Andrade-Navarro MA (2009). *Detection of alpha-rod protein repeats using a neural network and application to huntingtin*. PLoS Comput Biol 5(3):e1000304

Venkatesan K*, Rual JF*, Vazquez A*, Stelzl U*, Lemmens I*, Hirozane-Kishikawa T, Hao T, Zenkner M, Xin X, Goh KI, Yildirim MA, Simonis N, Heinzmann K, Gebreab F, Sahalie JM, Cevik S, Simon C, de Smet AS, Dann E, Smolyar A, Vinayagam A, Yu H, Szeto D, Borick H, Dricot A, Klitgord N, Murray RR, Lin C, Lalowski M, Timm J, Rau K, Boone C, Braun P, Cusick ME, Roth FP, Hill DE, Tavernier J, Wanker EE, Barabási AL, Vidal M (2009). *An empirical framework for binary interactome mapping*. Nat Methods 6(1):83-90

Invited plenary lectures

A Y2H-seq approach to define the protein methyltransferase interactome, Integrative Network Biology 2012: Network Medicine, Helsingør, Denmark, May 11-13, 2012

Dynamic protein interaction networks in cellular signalling, Pharmacology and Molecular Sciences Wednesday Seminar Series, Johns Hopkins University School of Medicine, Baltimore, USA, Nov 30, 2011

Dynamic protein-protein interaction wiring of the human spliceosome, invited seminar at the MIT, Cambridge Boston, Nov 28, 2011

Dynamic protein interaction networks in cellular signalling, invited seminar at the Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Nov 14, 2011

Signaling dynamics in PPI Networks, British-German Frontiers of Science Symposium, May 12-15, 2011, Kavli International Centre, Chicheley Hall, UK (Flashtalk)

Towards the systematic analysis of protein-protein interaction dynamics, Lise-Meitner Kolloquium at Freie Universität Berlin, Germany, Dec 10, 2010

Systematic analysis of human protein-protein interactions, Symposium of the Biology and Medicine Section in the MPS, Harnack House, Berlin, November 22-23, 2010

Towards the analysis of protein-protein interaction dynamics, invited talk at the workshop: PPI Berlin: Current Trends in Network Biology, Max Delbrueck Communications Center Berlin-Buch, Germany, Oct 8-9, 2010

A protein interaction wiring of the human spliceosome, Joint Cold Spring Harbor Laboratory / Wellcome Trust Meeting on SYSTEMS BIOLOGY: NETWORKS, Hinxton, UK, August 11-15, 2010

Systematic analysis of protein-protein interaction networks, Informa conference Targets and Tools, Symposium X, Berlin, Germany, March 16-18, 2010

Yeast two-hybrid protein-protein interaction screening, Wellcome Trust 91st Advanced Course, Protein Interactions and Networks, Wellcome Trust Sanger Institute, Genome Campus Hinxton, Cambridge UK, Dec 13-18, 2009

Constructing directed protein interaction networks for activated EGF/Erk signalling, Green Seminar: Biotechnology Center, TU Dresden, Germany, Oct 16, 2009

Constructing directed protein interaction networks for cellular signalling, invited talk at the workshop: "What is a macromolecular Complex? Shades of Meaning Across Cellular, Systems, and Structural Biology". NKI, Amsterdam, NL, Oct 1-2, 2009

Constructing directed protein interaction networks for activated EGF/Erk signalling, EBI Seminars in Systems Biology: European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, UK, June 2, 2009

Constructing causal protein interaction networks for activated EGF/Erk signalling, CSHL Meeting Systems Biology: Networks, Cold Spring Harbor Laboratory, New York, USA, March 18-22, 2009

PhD theses

Mareike Weimann: *A proteome-wide screen utilising second generation sequencing for the identification of lysine and arginine methyltransferase-protein interactions*. 2012

Josephine Worseck: *Characterization of phosphorylation-dependent interactions involving neurofibromin 2 (NF2, merlin) isoforms and the Parkinson protein 7 (PARK7, DJ1)*. 2012

Atanas Kamburov: *More complete and more accurate interactomes for elucidating the mechanisms of complex diseases*. 2012

Student theses

Franziska Wachsmuth: *Characterization of GRB2 and PIK3R3 phospho-tyrosine dependent protein interactions using yeast two-hybrid and luciferase reporter analyses*. Diploma thesis, 2012

Ziya Özkan: *Darstellung von Hefe Knockout-Stämmen für die Analyse von Proteinwechselwirkung*. Bachelor thesis, 2010

Federico Apelt: *Analyse dynamischer Interaktionen von Proteinkinase A*. Bachelor thesis, 2010

Crysovalantis Sourlis: *Yeast two-hybrid protein interaction wiring of the human spliceosome: implications for the architecture of the U5 / U2 /Prp19 spliceosomal core*. Diploma thesis, 2010

Teaching activities

2nd module of the PhD program of the MPIMG

4th module of the PhD program of the MPIMG

Wellcome Trust 91st Advanced Course, Protein Interactions and Networks (guest speaker)

Auricher Wissenschaftstage (11.04.11-22.04.11), Schülerpraktikum

Guest scientists

Tonio Schütze, Freie Universität Berlin, Wahl Lab, Germany, since 10/2011

Scientific Services

Animal Facility

(Established: 2003)

Head
Dr. Ludger Hartmann (since 05/01)
Phone: +49 (0)30 8413-1189
Fax: +49 (0)30 8413-1197
Email: hartmann@molgen.mpg.de

Animal care takers
Nadine Lehmann (since 09/09)
Katharina Hansen (since 04/09)
Christin Franke (since 09/07)
Dijana Micic (since 09/07)
Eileen Jungnickel (since 09/06)
Sonja Banko (since 04/05)
Mirjam Peetz (since 01/05)
Carolin Willke (since 09/02)
Julia Wiesner (since 05/02)
Katja Reinsch (since 06/99)
Ulf Schroeder (master) (since 09/96)
Janina Hoppe (09/08-08/12)

Apprentices
Anna Damm (since 09/12)
Niclas Engemann (since 09/12)
Ceszendra Kaufmann (since 09/12)
Mareike Wegmann (since 09/11)
Laura Kühn (since 09/10)
Larissa Schmidtke (since 09/10)
David Brandenburg (09/09-08/12)
Sarah Hackforth (09/09-08/12)

Service
2.5 persons from a Service Company (cage washing etc.)
Edward Somera (08/09-12/10)

Overview

The *Animal Facility* of the MPIMG was completely brought into service in the year 2003. It provides an optimal research environment in the field of Laboratory Animal Science, which includes the basic animal breeding and maintenance service for approximately 300 genetically modified and 30 wildtype mouse strains and technical services with a highly motivated staff. The mouse strains are kept

under specified pathogen free (SPF) conditions in areas with restricted access. By using several physical barriers and standard working protocols, we have been strongly committed ourselves to keep our rodent colony free of rodent pathogens. All strains are housed in individually ventilated caging systems (approximately 6.000 cages) and are handled under sterile conditions (with changing hood).

The Animal Facility provides high standard services which includes:

- Animal husbandry
- Colony management
- Assistance in experimental design and techniques
- Experimental work
- Tissue biopsies
- Blood and organ collection
- Health monitoring
- Cryopreservation of mouse embryos and sperm freezing
- In vitro fertilisation (IVF)
- Sterile embryotransfer
- Training for researchers, caretakers and trainees
- Import & export of animals
- Quarantine
- Rederivation

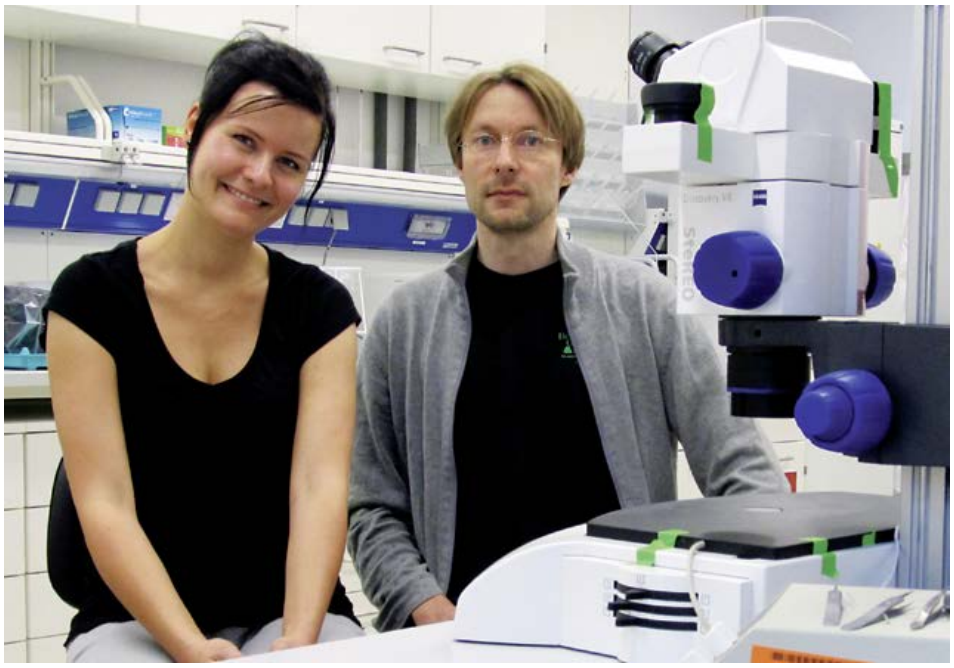
For the management of these mouse strains and the offered services, a mouse-colony management software program (PyRAT®) was established. This software enables scientists to see and modify all their research data online.

The *Zebrafish Facility* of our institute is set up to raise and keep up to 15,000 zebrafish (*Danio rerio*). The aquaria system is located in the animal house and consists of approximately 150 single tanks that are used for breeding and maintenance of zebrafish lines, as well as for providing eggs, embryos and larvae to the researchers of the institute. For zebrafish embryo manipulation, the facility offers a DNA/RNA microinjection setup.

Scientific Services

Transgenic Unit

(Established: 2004)



Head
Dr. Lars Wittler (since 08/10, Dept. of Developmental Genetics)
Phone: +49 (0)30 8413-1192/1453
Fax: +49 (0)30 8413-1197
Email: wittler@molgen.mpg.de

Technician
Judith Fiedler (since 07/11)
Larissa Mosch (05/05-12/09)

Overview

The *Transgenic Unit* of the MPIMG was established in 2004 and has been managed by Dr. Lars Wittler since 2010. It enables the successful and efficient generation of genetically modified mice for the Institute, providing a centralised resource and state-of-the-art technology for generating transgenic animals. It also gives expertise and support for the generation of murine ES cell lines and ES cell culture.

We mainly employ the diploid and tetraploid morula aggregation technologies, since they are an efficient method for generating chimaeric embryos with a high rate of ES cell contribution. By using the aggregation method, we established a robust routine platform, allowing us to utilise up to 6 ES-cell lines per week to generate mouse lines or embryos for direct phenotypic analyses.

For conservation and long-term storage of mouse lines, the transgenic unit cryopreserves murine sperm and embryos. Cryopreservation reduces maintenance costs and safeguards valuable mouse lines against loss through infection, disease, or breeding failure. Additionally, cryopreserved material allows the easy interchange of material between institutes, reducing the transport of live mice.

Scientific Services

Sequencing Facility

(Established: 2007)

Head
Bernd Timmermann (since 07/08)
Phone: +49 (0)30 8413-1121/1542
Fax: +49 (0)30 8413-1365
Email: timmermann@molgen.mpg.de

Bioinformaticians
Martin Kerick (since 05/12)
Martin Werber (since 05/11)
Heiner Kuhl (since 06/10)

Technicians
Miriam Mokros (since 02/12)
Norbert Mages (since 03/11)
Daniela Roth (since 03/11)
Sven Klages
(since 01/11, data analysis)
Ina Lehmann (since 08/10)
Ilona Hauenschild (since 09/09)
Sonia Paturej since 07/08)
Bettina Moser (01/08-10/11)
Sabrina Rau (03/11-09/11)
Isabelle Kühndahl (12/07-12/10)

Overview

New sequencing technologies are currently revolutionizing the field of genomic research. They enable researchers to investigate (epi)genomes and transcriptomes at an unexampled depth and level of detail. Founded in 2007, the Sequencing Facility at the MPIMG is a central unit which provides its service to all research groups of the institute. Since 2010, the expertise is also provided to scientists outside the MPIMG and the group acquired the status of a Sequencing Core Facility for the institutes of the biological-medical section (BMS) of the Max Planck Society.

The Sequencing Facility operates several next generation sequencers and maintains a fully equipped lab and staff able to perform a variety of sequencing applications - from sample preparation to data analysis. The unit was founded to help researchers process DNA and RNA samples in an efficient and economical manner. We established automation solutions for the sample preparation to

increase the throughput and minimize technical variation (Beckmann Spri-Te and 384 Beckman Multimek pipetting systems). By centralizing equipment and expertise, we have dramatically reduced the overall costs of sequencing, while increasing the efficiency and quality of the data generated.
Currently we are providing expertise for two different technical platforms: Roche/454 FLX+ and Illumina Hiseq 2000/GAIIx systems. At a read length of up to 1000 bases (modal read length around 800 bases), the 454 technology offers a great benefit especially for *de novo* genome sequencing, metagenome analysis, full length transcriptome analysis and amplicon sequencing. The high throughput of our Illumina systems in terms of Gigabases produced per run completes our sequencing service and offers a real advantage for many applications. Expression profiling (RNA-Seq), methylation analysis (MeDIP-Seq and Bisulphite-Seq), copy number analysis as well as the identification of protein binding sides (ChIP-Seq) and the analysis of whole exomes or genomes profite from the high output of this system to a great extent. The best practice protocols for both technologies established at our facility are frequently revised and carefully improved if necessary, to guarantee high standards of data quality and comparability.
Complementing the high throughput technologies, we additionally operate classical capillary systems (ABI 3730xl and 3130xl genetic analyzers) to provide Sanger sequencing to all departments of the institute.
Both platforms, the Roche/454 and the Illumina systems, have been extensively used for different projects. Currently the Sequencing Facility is involved as sequencing production side in several international genome projects. The 1000 Genomes Project (http://www.1000genomes.org/) aims to investigate genetic variations with an allele frequency >1% in multiple human populations. The OncoTrack Project (http://www.oncotrack.eu/) is directed at the identification of new biomarkers and their application for colon cancer. In this project, the Sequencing Facility is responsible for full length transcriptome analysis with the goal of identification of new fusion genes and splice variants.
As Sequencing Core Facility of the MPG we are currently involved in a couple of different projects, like the *de novo* sequencing of the canary genome (MPI for Ornithology, Manfred Gahr) or the analysis of different Nicotiana species genomes and transcriptomes (MPI for Chemical Ecology, Ian T. Baldwin).

Selected publications

Prigione A, Lichtner B, Kuhl H, Struys EA, Wamelink M, Lehrach H, Ralser M, Timmermann B, Adjaye J (2011). *Human induced pluripotent stem cells harbor homoplasmic and heteroplasmic mitochondrial DNA mutations while maintaining human embryonic stem cell-like metabolic reprogramming*. Stem Cells 29(9):1338-48

Timmermann B, Kerick M, Roehr C, Fischer A, Isau M, Boerno ST, Wunderlich A, Barmeyer C, Seemann P, Koenig J, Lappe M, Kuss AW, Garshasbi M, Bertram L, Trappe K, Werber M, Herrmann BG, Zatloukal K, Lehrach H, Schweiger MR (2010). *Somatic mutation profiles of MSI and MSS colorectal cancer identified by whole exome next generation sequencing and bioinformatics analysis*. PLoS One 5(12): e15661

1000 Genomes Project Consortium [MPI contributors: Sudbrak R, Albrecht MW, Amstislavskiy VS, Borodina TA, Dahl A, Davydov AN, Herwig R, Marquardt P, Mertes F, Nietfeld W, Parkhomchuk DV, Soldatov AV, Timmermann B, Tolzmann M, Lehrach H] (2010). *A map of human genome variation from population-scale sequencing*. Nature 467(7319): 1061-73

Scientific Services

Mass Spectrometry Facility

(Established: 03/2012)

Head
David Meierhofer, PhD (since 03/12)
Phone: +49 (0)30 8413-1567
Fax: +49 (0)30 8413-1380
Email: meierhof@molgen.mpg.de

Scientist
Klaus-Dieter Klöppel
(since 03/12, part time)

PhD student
Ina Gielisch (since 04/12)

Technician
Beata Lukaszewska-McGreal
(since 03/12)

Overview

The Mass Spectrometry Facility provides support for the entire institute in the field of proteomics and metabolomics and is funded by MPIMG. The following mass analyzers are available:
QTrap 5500 from AB/Sciex: This instrument is set up to detect known peptides or metabolites. For each peptide or metabolite, an individual method has to be established on the instrument. It is the most sensitive mass spectrometer in house, but it is essential to know the exact mass of the molecules of interest. The QTrap can be attached to an Eksigent nanoLC 2D Ultra liquid chromatography unit (routinely used for peptide identification), or alternatively to an Agilent 1290 Infinitiy ultra-high pressure liquid chromatography unit (routinely used for metabolite identification).
MALDI Ultraflex II from Bruker Daltonics: This instrument can be used for fast identification of simple protein/peptide mixtures.

*Orbitrap XL-ETD from Thermo Scientific:*We are using this mass analyzer in cooperation with Sascha Sauer. It is set up to analyze peptides proteome wide. In contrast to the QTrap, no informations about the petides and their post transcriptinal modifications need to be known in advance.
We developed a set of pipelines in order to decipher entire proteomes and post translational modifications, as well as single proteins, e.g. bands from SDS page gels, IPs, on bead digestions. We are able to detect more than 7000 proteins in a mammalian proteome and working on getting even more. As phosphorylations are of great interest, we can now detect more than 10.000 individual phosphorylation sites per proteome with our new enrichment method. For relative quantification, we are using SILAC (stable isotope labeling by amino acids in cell culture) and dimethyl labeling routinely.
For data analysis, we have a Mascot server (search engine that uses mass spectrometry peptide data to identify proteins from primary sequence databases). For more complex and larger mass spectrometry files, we have a powerful Dell server with MaxQuant tools installed (a quantitative proteomics software package, developed by Matthias Mann, MPI for Biochemistry, Martinsried).

Own research interests

Besides providing MS service for the institute, we pursue our own research topic: *Proteome and metabolome alterations in mitochondrial pathologies*.
Mitochondrial pathologies are a heterogeneous group of metabolic disorders that are frequently characterized by anomalies of oxidative phosphorylation, especially in the respiratory chain. These pathologies show a wide spectrum of clinical manifestations and variation in the mode of onset, course and progression with disease. The aim is to gain quantitative information on the regulatory network and interplay of proteins and metabolites in mitochondrial disorders. Cells, featuring known mitochondrial dysfunctions will be analyzed by high resolution mass spectrometers with a special focus on post translational modifications. The outcome of quantitative- proteomics and metabolomics will be combined and compared to achieve a complete picture of the investigated pathology.

Many cooperations have been established within the institute. In addition, we have external cooperations with:

- Peter Kaiser, Biological Chemistry, School of Medicine, University of California, Irvine
- Hans Mayr, Universitätsklinik für Kinder- und Jugendheilkunde, Salzburger Landeskliniken, Salzburg, Austria

Selected publications

Mayr JA, Zimmermann FA, Fauth C, Bergheim C, Meierhofer D, Radmayr D, Zschocke J, Koch J, Sperl W (2011). *Lipoic acid synthetase deficiency causes neonatal-onset epilepsy, defective mitochondrial energy metabolism, and glycine elevation*. Am J Hum Genet 89(6):792-7

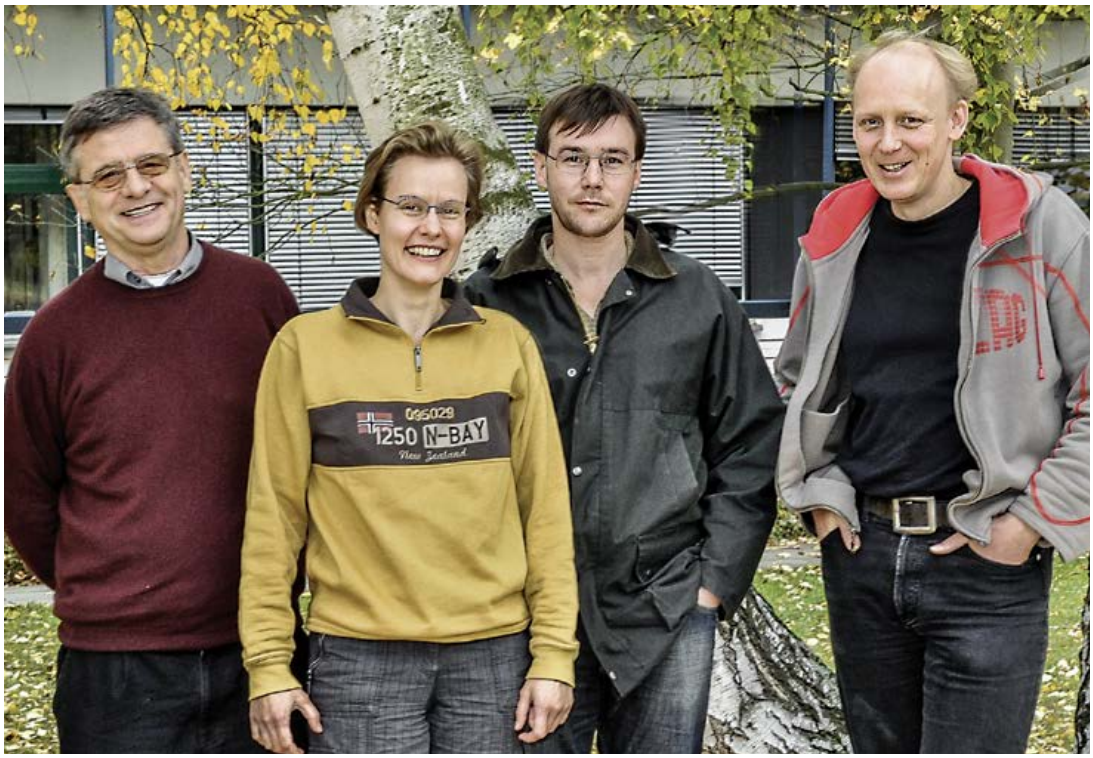
Meierhofer D, Wang X, Huang L, Kaiser P (2008). *Quantitative Analysis of global Ubiquitination in HeLa Cells by Mass Spectrometry*. J. Proteome Res 7(10):4566–76

Mayr JA*, Meierhofer D*, Zimmermann F, Feichtinger R, Kögler C, Ratschek M, Schmeller N, Sperl W, Kofler B (2008). *Loss of complex I due to mitochondrial DNA mutations in renal oncocytoma*. Clin Cancer Res 14(8):2270-5 * equal contributors

Scientific Services

Microscopy & Cryo Electron Microscopy Group

(Established: 1978/microscopy; 2004/cryo electron microscopy)



Heads

Dr. Rudi Lurz
(since 01/78, microscopy group)
Phone: +49 (0)30 8413-1644
Fax: +49 (0)30 8413-1385
Email: lurz@molgen.mpg.de

Dr. Thorsten Mielke*
(since 01/04, cryo-EM group
Phone: +49 (0)30 8413-1644
Fax: +49 (0)30 8413-1385
Email: mielke@molgen.mpg.de

Technicians

Beatrix Fauler (since 08/08)
Jörg Bürger* (since 08/06)
Matthias Brünner* (09/07-12/11)

Overview

From imaging service to structure determination of macromolecular complexes

For many years, the *microscopy group* headed by Rudi Lurz provides a broad range of imaging techniques for all departments of the institute, combining both, light and electron microscopy. The group operates two transmission electron microscopes, a 100 kV Philips CM100 and a 120 kV Tecnai Spirit (FEI), both equipped with CCD-cameras and optional cryo- or tomography holder. The lab has

* externally funded

established a wide range of cell-biological methods such as ultra-thin sectioning of plastic-embedded samples, immune-labelling of sections or isolated structures and visualization of nucleic acids and nucleic acid-protein complexes. Moreover, the group has a strong emphasis on fine-structure analysis of protein complexes and viruses using conventional EM as well as cryo-electron microscopy (cryo-EM).

In light-microscopy, the microscopy group is responsible for the service, maintenance and training of an increasing number of microscopes and users. Currently, we support four fluorescence/confocal microscopes (Zeiss LSM510, Zeiss LSM510meta, LSM 700 and AxioImager Z1). These instruments are operated as shared equipment and are accessible for all members of the institute. Furthermore, the group supports all users according to their specific biological questions and implements new techniques and applications. A Zeiss two-photon laser-scanning microscope LSM710-NLO was installed in May 2012.

Starting in 2004, the *cryo-EM group* headed by Thorsten Mielke established a state-of-the-art cryo-EM facility within the Berlin-Brandenburg research consortia "UltraStructure Network" (USN) and "Anwenderzentrum" (AWZ) for structure determination of macromolecular protein complexes using cryo-EM in combination with the single particle approach. Our facility provides a technology platform for sample screening, semi-automated sample vitrification, data acquisition and intense computing resources for image processing. Core instrument of the facility is a helium-cooled 300 kV Tecnai G2 Polara electron microscope (FEI) equipped with a 4k F416 CMOS camera (TVIPS), which is as well as all other equipment accessible for all groups at the MPIMG.

In order to account for the constantly increasing demands on imaging and visualizing biological structures within the institute and to guarantee the maximal professional and personnel continuity after the retirement of Rudi Lurz in December 2012, the microscopy group and cryo-EM group will be fused to one central scientific service group, which will then be headed by Thorsten Mielke. The joined group will continue to provide service and training in all aspects of light and electron microscopy and will continue the successful structural analysis of protein complexes using cryo-EM and single particle techniques.

Activities of the microscopy group

Besides the light microscopy support, most of our projects within the institute are on ultra-thin sectioning of tissues, cells and cell components. We hereby cooperate with all departments of the institute having a wet lab. The technical support is mainly performed by Beatrix Fauler and ranges from simple quality checks to long lasting cooperations.

Within the Berlin region we continued our EM studies on the aggregation of proteins responsible for degenerative brain diseases in cooperation with the groups of E. Wanker (MDC, Berlin) and G. Multhaup (FU Berlin). Outside Berlin, we cooperate with several international research groups on bacterial phages: SPP1 is a *B. subtilis* phage, which was introduced by the former director Thomas Trautner and which was studied in this institute for many decades. Together with Paulo Tavares (CNRS, Gif-sur-Yvette), we analysed numerous aspects of this phage using EM. Currently we are focusing on changes in the connector-tail region before and after infection. Image processing of SPP1 cryo data is done in cooperation with Elena Orlova (Birkbeck College, University of London).

Together with Pascale Boulanger (Université de Paris-Sud, ORSAY) and Cécile Breyton (CNRS, Grenoble), we analyse structural changes of the T5 phage

after interaction with the receptor protein. Individual proteins in the tail end structure are localized by immuno labelling. Other studies on phage morphology, repressor/operator interaction and IEM, respectively, were done in cooperation with the groups of M. Loessner (ETH Zürich), K. Geider (JKI Dossenheim) and R. Hertwig (BfR Berlin).

Furthermore, we use classical, but still powerful, mica-adsorption techniques to localize the position of proteins bound to DNA: Together with Christian Speck (Imperial College, London), we apply this method to visualize DNA-protein complexes involved in replication initiation in yeast. Bacterial initiation of replication is analysed with Rafał Donczew (Anna Zawilak-Pawlik, Institute of Immunology and Experimental Therapy, Wrocław) by binding of DnaA to the oriC region in *Helicobacter pylori*. Virtu Solano (Alicia Bravo, CIB, Madrid) maps the primary binding sites of the transcriptional regulator Mga*Spn* from *Streptococcus pneumoniae* on selected DNA fragments.

Current activities and future perspectives: Cryo-EM

Macromolecular protein complexes play a crucial role in all central biological processes such as replication, transcription, protein biosynthesis, metabolism as well as organization of the cell. Cryo-EM in combination with the single-particle approach has emerged as the key technology to gain structural information on protein complexes without the need to crystallize these complex and often dynamic systems. The molecules of interest are embedded in a thin layer of vitreous ice under near physiological conditions and imaged by transmission electron microscopy. State-of-art instruments such as our Polara microscope enabled us to solve various ribosomal protein complexes at sub-nanometer resolution in cooperation with the groups of Knud Nierhaus (MPIMG, now Charité) and Christian Spahn (Charité). Although commonly thought to act as highly organized molecular machines, it becomes more and more evident that protein complexes show a variable assembly, appear in various functional states and are subject to dynamical regulation. Since the single particle approach is an averaging technique, extrinsic or intrinsic sample flexibility is thus causing a serious heterogeneity problem.

To overcome this problem, we aim to implement new image

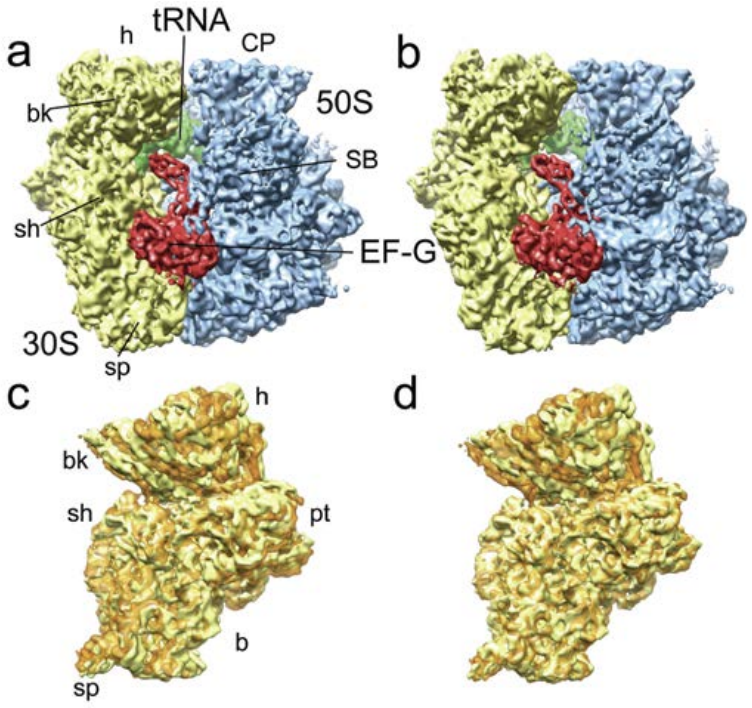


Figure 1: Cryo-EM reconstructions of sub-states (a) I (TIPRE) and (b) II (TIPOST) of the T. thermophilus 70S•EF-G•GDP•FA complex shown from the L7-stalk-site (yellow: 30S subunit; blue: 50S subunit; red: EF-G; green: tRNA). (c) The comparison of sub-states I and II of the 30S subunit with the maps aligned at the 50S subunit indicates the difference in terms of ratcheting. (d) The alignment to the body/platform domains of the 30S subunit highlights differences in head swiveling. The 30S of sub-state I (TIPRE) is rendered in transparent orange, while the 30S of sub-state II (TIPOST) is in solid yellow (from Ratje et al., Nature 2010).

processing tools for particle classification and sorting (collaborative research centre/SFB 740, project Z1). Applying our multiparticle refinement strategies, we could separate data sets of even biochemically well-defined bacterial 70S ribosomes, which were further stabilized by binding of the antibiotic fusidic acid, into more homogenous subsets (figure 1). The structures of these sub-states lead to the identification of a novel intra-subunit pe/E hybrid state showing a partly translocated tRNA. Multiparticle refinement of a fusidic acid-stalled 70S-tmRNA-SmpB-EF-G complex enabled us to identify a post-translocational intermediate (TI^{POST}), which presents the TLD-SmpB module in an intrasubunit ap/P hybrid site and a tRNA-fMet in an intrasubunit pe/E hybrid site that also shows a unique extra-large swivel movement of the 30S head. Similarly, we could observe distinct sub-states in the mammalian 80S ribosomal pre-translocation complex, which differ in large-scale conformational changes including intersubunit rotation of the ribosomal subunits as well as the binding mode of the tRNAs, whereby hybrid states are favoured within the mammalian complex.

Hence, analysis of sample heterogeneity is not only essential to improve the quality and resolution of a cryo-EM reconstruction, but also underlines the necessity to collect even larger data sets. We therefore implemented the Leginon system for automatic data collection. At the Spirit microscope, Leginon is now routinely used to collect data sets of about 1000-1500 digital images for initial 3D reconstructions including acquisition of tilt-pairs for random conical tilt analysis. After installing a new CMOS camera at our Polara microscope, we are currently evaluating the resolution of cryo-EM reconstructions obtained from Leginon data.

Within the framework of the collaborative research centre 740, we are also focusing on the implementation of molecular electron tomography as a tool for initial structure determination and structural analysis of protein modules with varying stability and/or only temporary associated subunits. Moreover, we offer cellular electron tomography to gain 3D information on e.g. thin-sectioned biological structures.

Selected publications

Ramrath DJF, Yamamoto H, Rother K, Wittek D, Pech M, Mielke T, Loerke J, Scheerer P, Ivanov P, Teraoka Y, Shpanchenko O, Nierhaus KH, Spahn CMT (2012). *The complex of tmRNA-SmpB and EF-G on translocating ribosomes*. Nature 485:526-9

Prigione A, Fauler B, Lurz R, Lehrach H, Adjaye J (2010). *The senescence-related mitochondrial/oxidative stress pathway is repressed in human induced pluripotent stem cells*. Stem Cells 28: 721-33

Ratje, AH, Loerke J, Mikolajka A, Brünner M, Hildebrand PW, Starosta A, Dönhöfer A, Connell SR, Fucini P, Mielke T, Whitford PC, Onuchic JN, Yu Y, Sanbonmatsu KY, Hartmann RK, Penczek PA, Wilson DN, Spahn CMT (2010). *Head swivel on the ribosome facilitates translocation by means of intra-subunit tRNA hybrid sites*. Nature 468:713-716

Scientific Services

IT Group

(Established: 1995)

Heads
Donald Buczek
(IT group since 12/96; head since 01/11)
Phone: +49 (0)30 8413-1433
Fax: +49 (0)30 8413-1432
Email: buczek@molgen.mpg.de

Peter Marquardt
(IT group since 08/95; head since 01/11)
Phone: +49 (0)30 8413-1430
Fax: +49 (0)30 8413-1432
Email: marquardt_p@molgen.mpg.de

IT staff
Matthias Rüster (since 09/12)
Tobias Dreyer (since 09/11)
Marius Tolzmann (since 08/07)
Sven Püstow (since 10/96)
Frank Rippel (since 01/95)
Alfred Beck (since 05/93)

Apprentices
Lena Zander (since 08/11)
Matthias Rüster (08/09-08/12)
Tobias Dreyer (09/08-08/11)

Student
Henriette Labsch (since 09/09)

With the move of Richard Reinhardt to Cologne in 2010, the responsibility for the IT group has been taken over by Donald Buczek and Peter Marquardt. The IT group is in charge of the operation and development of the IT-infrastructure of the institute. This includes workstation and server systems, storage, archives, wire based and wireless LAN, Internet access, Internet services and remote access.
The online storage capacity of the MPIMG on disk-based file servers exceeds 3.7 PB of data. The monthly backup volume sums up to about 85 TB, whereas the tape and disk-based offline archived data currently comprises about 1 PB. Presently the group serves about 450 Window based PCs and 300 Linux/Unix systems with a variety of hard- and software components and about 100 OSX systems. A variety of web servers are protected by a fire-wall installation, about 60 web servers are

active and maintained. The active hard- and software development of the group serves the scientific departments as well as the service and administration groups. Since 2008, the new requests of the NGS technology have dramatically increased the effort of the IT group to serve the computational and storage needs for data processing and analysis. We developed a new concept for short and long term data storage based on disk-arrays, installed a storage capacity of more than 3500 TB and increased the computational power by more than a factor of 100. Currently we are running about 3000 CPU cores with 13000 GB RAM spread over about 300 Linux systems ranging from single core systems with 256 MB RAM up to 64 multicore servers with 512 GB RAM.
Our internal network backbone is based on 10 GbE technology and is currently fed by approx. 50 interconnected network interfaces, from Isilon storage systems *via* multicore compute servers up to huge file- and archive servers. The inhouse LAN is segmented by about 180 manageable switches giving us the flexibility to control each segment and if necessary to configure each switch port individually. To supply a stable and reliable infrastructure for our IT equipment, we planned and implemented two physically separated server rooms. The storage and archive server room located in tower 4 is capable of supplying 180 kW cooling capacity and houses 20 server racks. The room has been reconstructed life without service interrupts from a laboratory equipment room with free flow air cooling to a closed cold aisle containment system. For the second server room, located in the new tower 3, we will establish a warm aisle containment system, which shall be capable of cooling down 450 kW. The 30 racks installed will be used for network, computing, storage and multicore clusters, and provide space for 1400 rack units. The IT group is very active in the training and education of young technicians, students, trainees and apprentices. Our IT apprentices participate at the Bundeswettbewerb Informatik regularly and entered the competitions round two out of three successfully. This year, Matthias Rüster was honored "top 5 best of" training school and also entered the last round of BWINF. In addition, the apprentices also presented an open source project developed at the MPIMG at the LinuxTag 2012 at the Berlin Fairgrounds.

Selected publications

Clarke L, […], The 1000 Genomes Project Consortium* (2012). *The 1000 Genomes Project: data management and community access*. Nat Methods 9:459-462 * MPI contributors: Lehrach H, Sudbrak R, Borodina T, Davydov A, Marquardt P, Mertes F, Nietfeld W, Soldatov A, Timmermann B, Tolzmann M, Albrecht M, Amstislavskiy V, Herwig R, Parkhomchuk D.

Sudmant PH, […],1000 Genomes Project*, Eichler EE (2010). *Diversity of human copy number variation and multicopy genes*. Science 330:641-6 *MPI contributors: Lehrach H, Sudbrak R, Borodina T, Davydov A, Marquardt P, Mertes F, Nietfeld W, Soldatov A, Timmermann B, Tolzmann M, Albrecht M, Amstislavskiy V, Herwig R, Parkhomchuk D

1000 Genomes Project Consortium* (2010). *A map of human genome variation from population-scale sequencing*. Nature 467(7319): 1061-73 *MPIMG contributors: Sudbrak R, Albrecht MW, Amstislavskiy VS, Borodina TA, Dahl A, Davydov AN, Herwig R, Marquardt P, Mertes F, Nietfeld W, Parkhomchuk DV, Soldatov AV, Timmermann B, Tolzmann M, Lehrach H

Research Support
Administration

Head
Dr. Manuela B. Urban, MBA
(06/00-09/12)
Phone: +49 (0)30 8413-1360
Fax: +49 (0)30 8413-1394
Email: urban@molgen.mpg.de

Secretaries
Tamara Safari (since 09/10, part-time)
Jeannine Dilßner
(since 02/96, part-time)
Sebastian Klein (06/05-09/10)
Sara Aziz (03/09-04/10, part-time)
Phone: +49 (0)30 8413-1399/1364
Fax: +49 (0)30 8413-1394
Email: safari@molgen.mpg.de
dilssner@molgen.mpg.de

Personnel department
Julia Zlotowitz (head; since 07/12)
Stephanie Cuber (since 11/11)
Gisela Rimpler (since 07/11,
parental leave substitution)
Kathleen Linnhoff
(since 11/07, on parental leave)
Jeanette Brylla (since 04/96)
Jeannette Bertone
(since 03/96, part-time)
Ruth Schäfer (head; 05/75-08/12)
Margrit Pomerenke (07/08-03/12)
Hilke Wegwerth (01/92-03/11)

Finance department
Angelika Brehmer
(head, since 10/97)
Ines Konietzko
(since 04/12, part time)
Malgorzata Klemm (since 04/01)
Petra Saporito (since 04/96)
Ursula Schulz
(04/98-06/12, part-time)

External project funding
Joachim Gerlach (since 06/97)
Anke Badrow (since 02/96)

Guest houses, apartments
Tamara Safari (since 09/10, part-time)
Marion Radloff (since 09/10)
Marianne Hartwig (since 10/07)
Sara Aziz (03/09-09/10, part-time)
Eleonora Volcik (08/09-02/10)

Purchasing department
Kerstin Tobis (head, since 06/12)
Karsten Krause
(since 02/05, part-time)
Kerstin Steudtner (since 06/03)
Rita Röfke-Bohnau (since 02/97)
Ute Müller (since 01/97)
Dirk Reichert (head, 05/11- 12/11)
Jutta Roll (head, 07/76 - 08/11)

Stock room
Karsten Krause
(head, since 05/07, part-time)
Dominik Buggenhagen (since 09/09)
Jürgen Joch (since 04/84)
Olaf Kischkat (01/04-08/11)

Reception, post office
Gabriele Hänsel
(since 04/12, part-time)
Monika Schweizer-Annecke
(since 05/01, part-time)
Sara Aziz (05/10-04/12, part-time)

Driver
Claus Langrock (since 07/91)

Overview

The administration of the MPIMG secures smooth operations and stable infrastructures for the institute. Besides the core administrative tasks like personnel and accounting, the administration takes care of purchasing and of all financial aspects of national and international grants. Researchers receive support in legal questions pertaining to technology transfer and patenting. This, like many other issues, is dealt with in close cooperation with the respective departments of Max Planck Headquarters in Munich.
The budget of the MPIMG comes to a large proportion from the Max Planck Society. In addition, researchers bring in substantial amounts of external funding from sources like BMBF, DFG, or the European Commission. Since 2010, funding has been gradually decreased due to the upcoming retirement of Hans Lehrach and H.-Hilger Ropers in 2014. Over the same period, there has been a considerable increase in prices and salaries, which has been covered only partially by adjustments in the budget. Like in many other research labs, the institute's energy expenses have grown to more than 1.5 million EUR p.a., which is an increase of more than 90 % since 2004. Furthermore, a growing share of cost contribution from the institute is being expected for formerly centrally funded independent research groups, Max Planck Research Schools, technology platforms and basic services such as e-journals. Likewise, the institute has had to complement the budget provided centrally for the upkeep of its buildings for the last two years.
Considering the legal framework, without the selective liberation measures of the last few years such as global budgets, the omission of staff appointment schemes, etc., the institute would not have been able to cope with fast-changing demands. Very positive experiences have been had with a clause of the public procurement law adopted in 2010, which allows procurement of scientific goods without open competitive bidding. Thanks to this clause, bureaucracy has been reduced considerably. Another positive example is the possibility to award additional bonuses for scientists and technical staff. Due to that regulation, the institute is able to successfully compete in hiring qualified staff. Unfortunately, the administrative and technical service units are still not eligible. This causes growing problems since salary levels outside the public sector are significantly higher for similar job specifications.

In contrast to those important facilitations, there is an increasing obligation for detailed reporting and capacious documentation of many facts and procedures. A prominent example is the Seventh Framework Programme of the European Commission with e.g. documented costs for single experiments or time sheets to report on „non-productive times“. In a similar manner, auditing and financial authorities construe legal provisions increasingly narrowly. The partial omission of the allowance to deduct input VAT for the Max Planck Society in 2008 is a striking example. The institute is looking forward to the draft “law on the freedom of science” (Wissenschaftsfreiheitsgesetz), which has recently been passed by the Federal Government and is intended to provide further flexibility and autonomy in the management of government-financed research institutions.

Concerning the personnel policy, one recent topic has been the performance-related payments as required by the collective labour agreement for the federal public service (Tarifvertrag für den öffentlichen Dienst, TVöD). In 2010, an agreement was reached for the Max Planck Society between the joint workers council and the management as a prerequisite for the distribution of the additional bonuses. Separate from performance reviews as a basis for additional bonuses, the management of the institute is engaged in establishing formal annual appraisal meetings. The main intent is to evaluate career opportunities for temporary employed staff as well as to improve working conditions for individuals.

The institute’s efforts in vocational training have resulted in eight graduations since 2009, thereof six animal keepers and two IT specialists for application development/software development. For the coming year, the administrative and technical services will engage again in vocational training.

Research Support

Technical Management & Workshops

Head
Dipl.-Ing. (FH) Ulf Bornemann
(since 04/01)
Phone: +49 (0)30 8413-1424
Fax: +49 (0)30 8413-1394
Email: bornemann@molgen.mpg.de

Building services engineering
Reinhardt Strüver (head)
(since 07/02)
Reinhard Kluge (since 12/10)
Bernd Roehl (since 08/97)
Frank Kalaß (since 01/88)
Thomas Oster (08/05-03/11)

Electrical engineering
Frank Michaelis (head, since 05/06)
Michael Pöschmann (since 07/12)
Lars Radloff (since 08/92)
Bernd Zabka (since 08/90)
Bernd Roßdeutscher (since 06/87)
Udo Abratis (05/74-09/12)
Frank Baumann (09/08-01/12)

Electromechanics
Karsten Beyer (since 07/07)
Florian Zill (since 08/95)
Carsten Arold (since 05/92)

Glass instruments construction
Peter Ostendorf
(since 05/72, part-time)

Technical supply service
Dirk Grönboldt-Santana (since 01/98)

Overview

The MPIMG was founded in 1964. In 1968, towers 1 and 2 and the administrative and workshop buildings at Ihnestrasse 73 in Berlin-Dahlem were built. After more than 40 years of use, a structural renovation is urgently required to meet current requirements for fire protection, occupational safety and energy efficiency. The plans foresee insulating the building shell, replacing the windows and doors and adjusting the floor layout to new laboratory use concepts. A fundamental reinstallation of technical facilities for electricity, cooling, water, gas and compressed air provision, ventilation equipment, and fire alarms is also urgently required. For this purpose, it is also necessary to adjust capacities to the sharply increased scientific requirements. The last few years have already been characterized by supply bottlenecks, technical failures, and burst pipes.
In 2003, the MPIMG received official approval from the management of the Max Planck Society to build tower 3 and renovate towers 1 and 2 from ground up one after the other directly afterwards. In 2007, the construction plan was submitted to the responsible "Joint Science Conference" (Gemeinsame Wisssenschaftskonferenz) of the Federal Government and Federal States for approval. At that time, all construction activities of the Max Planck Society were thoroughly reviewed by this same commission. The approval process for building tower 3 and renovating towers 1 and 2 was thus delayed by two years before approval was finally given in 2009. The construction of tower 3 started in April 2011. Completion and handover are planned for January 2013. While the plan has been to have tower 2 completed by December 2014 with the subsequent renovation of tower 1, this schedule is unfortunately now again in danger for reasons of budget problems.
Tower 3 is being built as an "office building" for scientists working theoretically. The building will provide a new main entrance for the whole institute and thus focussing attention on it. The ground floor will host an entrance hall and several seminar rooms for conferences and events. The upper floors shall be used by the Department of Computational Molecular Biology and other research groups working theoretically as well as the IT service group. In addition, a new server room will be created in tower 3. All construction works are carried out during full research operations, which in part lead to considerable restriction. It was e.g. necessary to empty rooms in tower 2 directly adjoining the construction work, accommodate groups in other parts of the building and relocate the cryo-electron microscope into the Fritz Haber Institute until the end of construction work.
The construction of tower 3 will solve the technical defects in the institute's cooling supply. New cables have also been laid and new transformers, switch panels and an emergency generator have been installed to ensure the institute has a secure continuous electricity supply. Nevertheless, due to upcoming supply problems in the infrastructure, the renovation of towers 1 and 2 is still of utmost importance.

Max Planck Institute for Molecular Genetics
Ihnestr. 63–73
14195 Berlin
Germany

Underground / U-Bahn

- U3/Oskar-Helene-Heim
- U3/Thielplatz

Urban rail system / S-Bahn

- S1 / Sundgauer Str.

Bus routes / Buslinien

- M11, 110 - Saargemünder Str./ Bitscher Str.
- X10, 115, 285 - Clayallee/Leichhardtstr. or Schützallee
- M48, 101 - Berliner Straße/Holländische Mühle