CRG
Centre for Genomic Regulation

Centre for Genomic Regulation
Annual Report 2009

Centres de recerca de Catalunya

© Copyright 2010

Produced by: Department of Communication & Public Relations
Centre for Genomic Regulation (CRG)
Dr. Aiguader, 88
08003 Barcelona, Spain
www.crg.es

Texts and graphics: CRG Researchers, Department of Communication and Public Relations

Graphic Design: Genoma ArtStudio SCP (www.genoma-artstudio.com)

Photography: Ivan Paulick, Sergi Espada

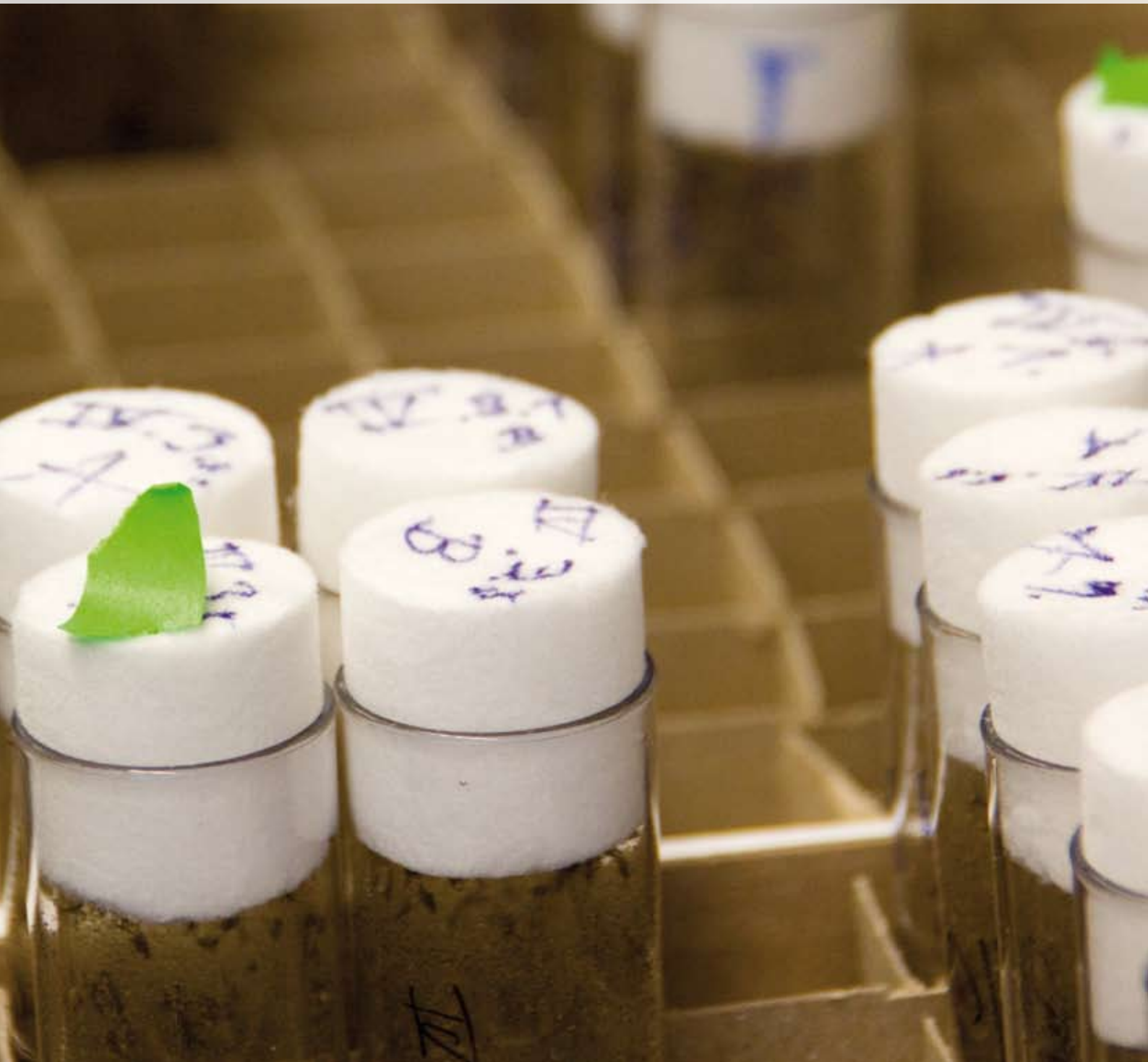
Printing: Novoprint, S.A.

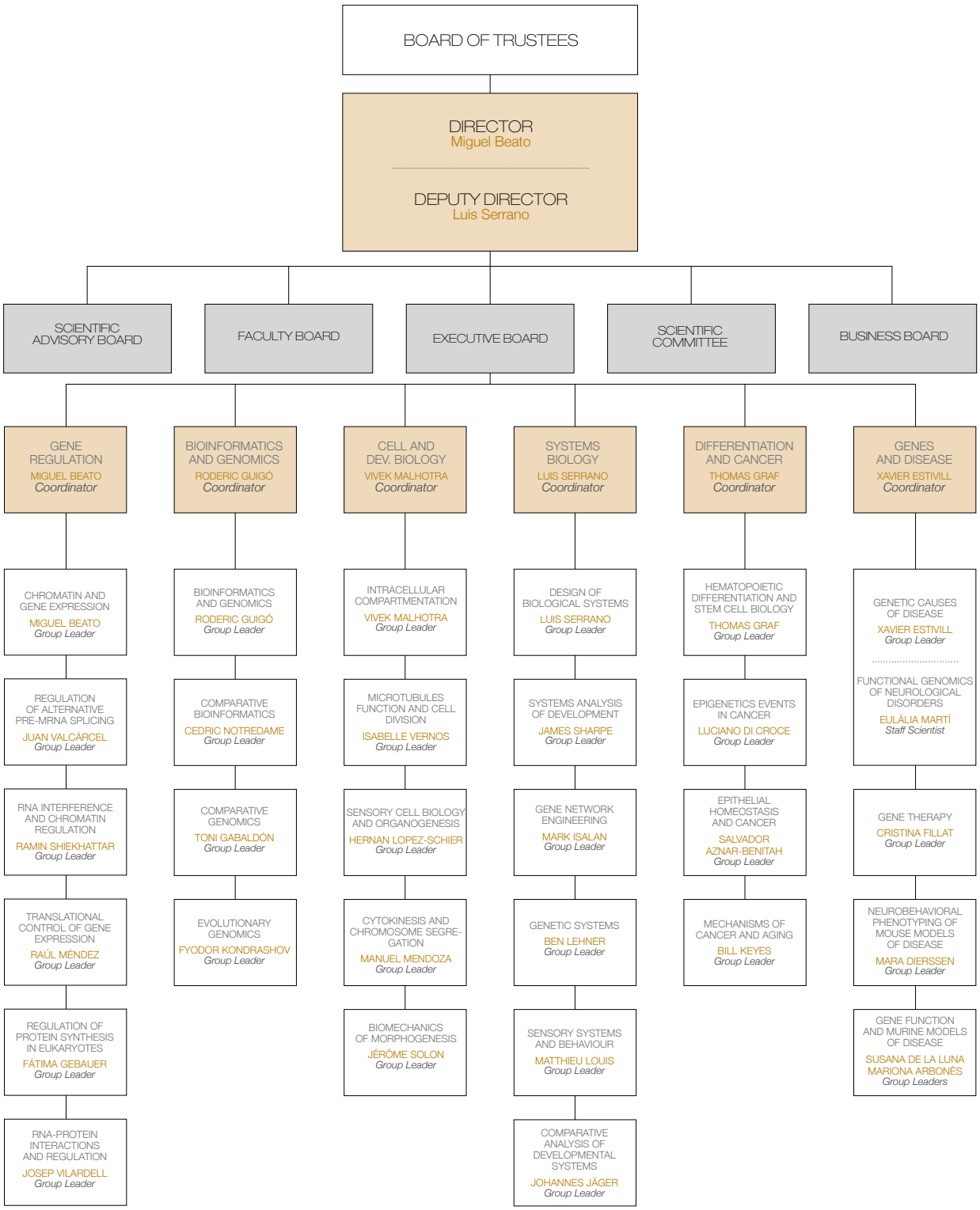
Legal deposit: B-31.812-2010


CRG
Centre
for Genomic
Regulation

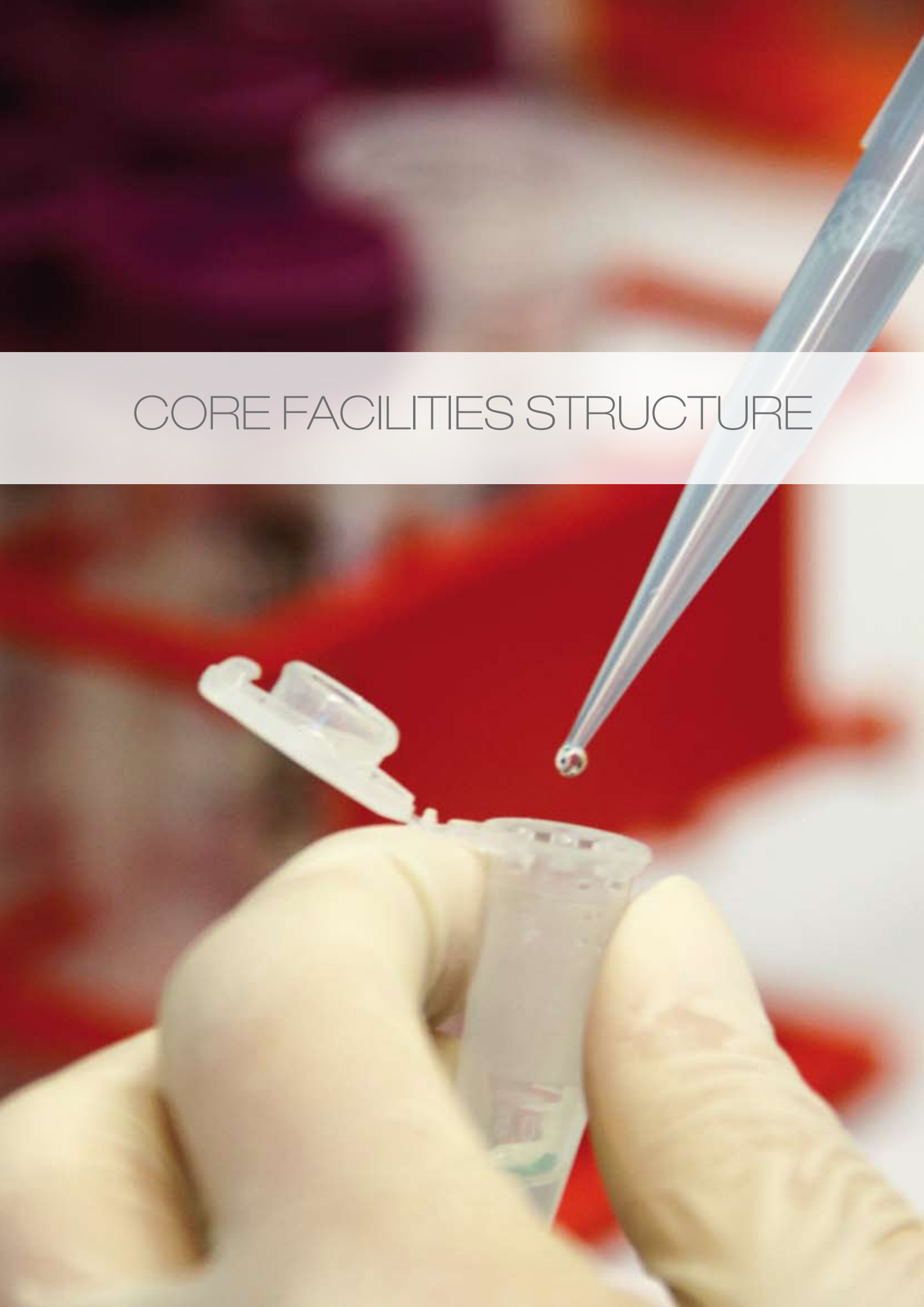
CONTENTS

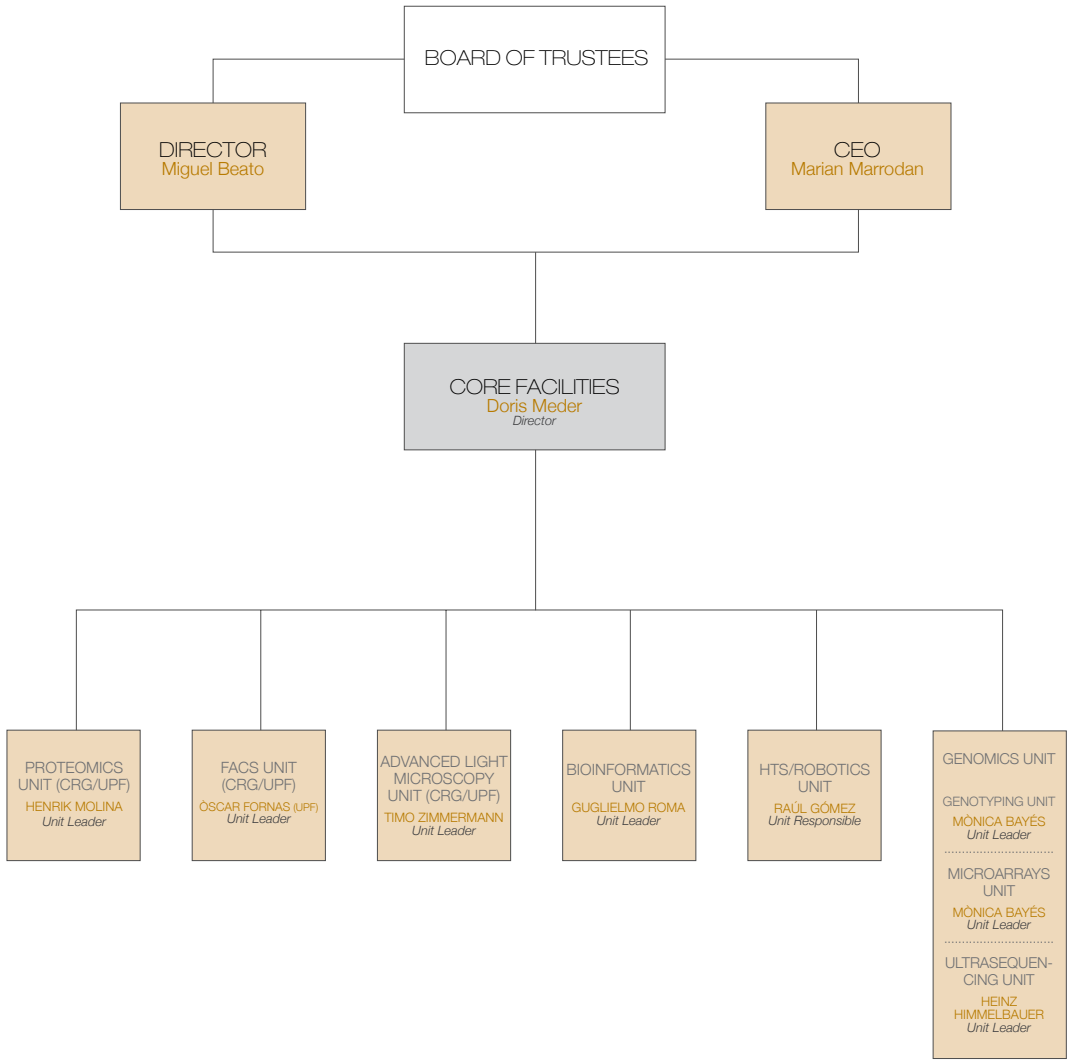
SCIENTIFIC STRUCTURE



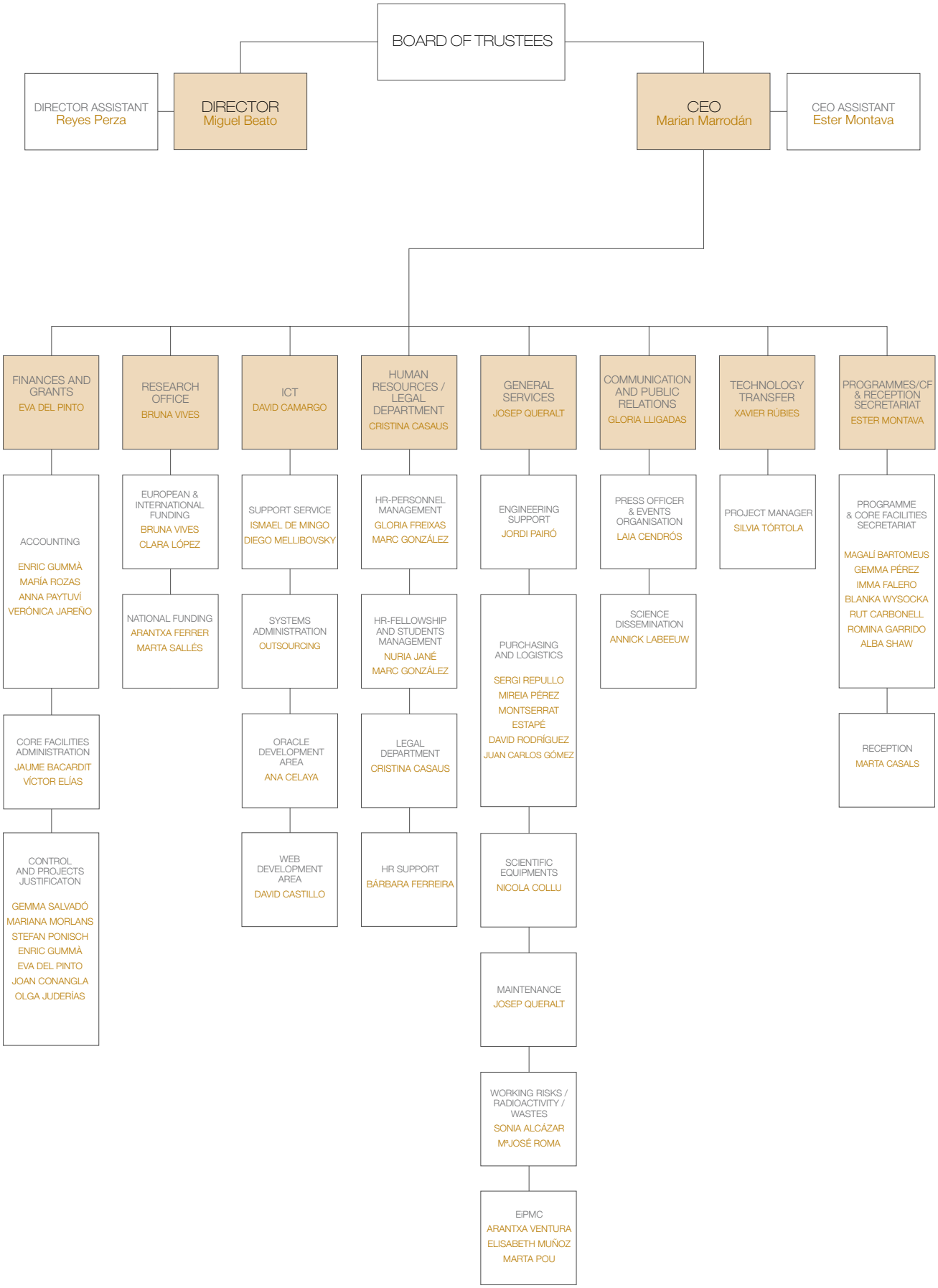

BOARD OF TRUSTEES
DIRECTOR
Miguel Beato
DEPUTY DIRECTOR
Luis Serrano
SCIENTIFIC ADVISORY BOARD
FACULTY BOARD
EXECUTIVE BOARD
SCIENTIFIC COMMITTEE
BUSINESS BOARD
GENE REGULATION
MIGUEL BEATO
Coordinator
BIOINFORMATICS AND GENOMICS
RODERIC GUIGÓ
Coordinator
CELL AND DEV. BIOLOGY
VIVEK MALHOTRA
Coordinator
SYSTEMS BIOLOGY
LUIS SERRANO
Coordinator
DIFFERENTIATION AND CANCER
THOMAS GRAF
Coordinator
GENES AND DISEASE
XAVIER ESTIVILL
Coordinator
CHROMATIN AND GENE EXPRESSION
MIGUEL BEATO
Group Leader
REGULATION OF ALTERNATIVE PRE-MRNA SPLICING
JUAN VALCÁRCEL
Group Leader
RNA INTERFERENCE AND CHROMATIN REGULATION
RAMIN SHIEKHATTAR
Group Leader
TRANSLATIONAL CONTROL OF GENE EXPRESSION
RAÚL MÉNDEZ
Group Leader
REGULATION OF PROTEIN SYNTHESIS IN EUKARYOTES
FÁTIMA GEBAUER
Group Leader
RNA-PROTEIN INTERACTIONS AND REGULATION
JOSEP VILARDELL
Group Leader
BIOINFORMATICS AND GENOMICS
RODERIC GUIGÓ
Group Leader
COMPARATIVE BIOINFORMATICS
CEDRIC NOTREDAME
Group Leader
COMPARATIVE GENOMICS
TONI GABALDÓN
Group Leader
EVOLUTIONARY GENOMICS
FYODOR KONDRASHOV
Group Leader
INTRACELLULAR COMPARTMENTATION
VIVEK MALHOTRA
Group Leader
MICROTUBULES FUNCTION AND CELL DIVISION
ISABELLE VERNOS
Group Leader
SENSORY CELL BIOLOGY AND ORGANOGENESIS
HERNAN LOPEZ-SCHIER
Group Leader
CYTOKINESIS AND CHROMOSOME SEGREGATION
MANUEL MENDOZA
Group Leader
BIOMECHANICS OF MORPHOGENESIS
JÉRÔME SOLON
Group Leader
DESIGN OF BIOLOGICAL SYSTEMS
LUIS SERRANO
Group Leader
SYSTEMS ANALYSIS OF DEVELOPMENT
JAMES SHARPE
Group Leader
GENE NETWORK ENGINEERING
MARK ISALAN
Group Leader
GENETIC SYSTEMS
BEN LEHNER
Group Leader
SENSORY SYSTEMS AND BEHAVIOUR
MATTHIEU LOUIS
Group Leader
COMPARATIVE ANALYSIS OF DEVELOPMENTAL SYSTEMS
JOHANNES JÄGER
Group Leader
HEMATOPOIETIC DIFFERENTIATION AND STEM CELL BIOLOGY
THOMAS GRAF
Group Leader
EPIGENETICS EVENTS IN CANCER
LUCIANO DI CROCE
Group Leader
EPITHELIAL HOMEOSTASIS AND CANCER
SALVADOR AZNAR-BENITAH
Group Leader
MECHANISMS OF CANCER AND AGING
BILL KEYES
Group Leader
GENETIC CAUSES OF DISEASE
XAVIER ESTIVILL
Group Leader
FUNCTIONAL GENOMICS OF NEUROLOGICAL DISORDERS
EULÀLIA MARTÍ
Staff Scientist
GENE THERAPY
CRISTINA FILLAT
Group Leader
NEUROBEHAVIORAL PHENOTYPING OF MOUSE MODELS OF DISEASE
MARA DIERSSEN
Group Leader
GENE FUNCTION AND MURINE MODELS OF DISEASE
SUSANA DE LA LUNA
MARIONA ARBONÉS
Group Leaders

CORE FACILITIES STRUCTURE




BOARD OF TRUSTEES
DIRECTOR
Miguel Beato
CEO
Marian Marrodan
CORE FACILITIES
Doris Meder
Director
PROTEOMICS UNIT (CRG/UPF)
HENRIK MOLINA
Unit Leader
FACS UNIT (CRG/UPF)
ÒSCAR FORNAS (UPF)
Unit Leader
ADVANCED LIGHT MICROSCOPY UNIT (CRG/UPF)
TIMO ZIMMERMANN
Unit Leader
BIOINFORMATICS UNIT
GUGLIELMO ROMA
Unit Leader
HTS/ROBOTICS UNIT
RAÚL GÓMEZ
Unit Responsible
GENOMICS UNIT
GENOTYPING UNIT
MÒNICA BAYÉS
Unit Leader
MICROARRAYS UNIT
MÒNICA BAYÉS
Unit Leader
ULTRASEQUEN-CING UNIT
HEINZ HIMMELBAUER
Unit Leader

MANAGEMENT STRUCTURE


BOARD OF TRUSTEES
DIRECTOR ASSISTANT
Reyes Perza
DIRECTOR
Miguel Beato
CEO
Marian Marrodán
CEO ASSISTANT
Ester Montava
FINANCES AND GRANTS
EVA DEL PINTO
RESEARCH OFFICE
BRUNA VIVES
ICT
DAVID CAMARGO
HUMAN RESOURCES / LEGAL DEPARTMENT
CRISTINA CASAUS
GENERAL SERVICES
JOSEP QUERALT
COMMUNICATION AND PUBLIC RELATIONS
GLORIA LLIGADAS
TECHNOLOGY TRANSFER
XAVIER RÚBIES
PROGRAMMES/CF & RECEPTION SECRETARIAT
ESTER MONTAVA
ACCOUNTING
ENRIC GUMMÀ
MARÍA ROZAS
ANNA PAYTUVÍ
VERÓNICA JAREÑO
EUROPEAN & INTERNATIONAL FUNDING
BRUNA VIVES
CLARA LÓPEZ
SUPPORT SERVICE
ISMAEL DE MINGO
DIEGO MELLIBOVSKY
HR-PERSONNEL MANAGEMENT
GLORIA FREIXAS
MARC GONZÁLEZ
ENGINEERING SUPPORT
JORDI PAIRÓ
PRESS OFFICER & EVENTS ORGANISATION
LAIA CENDRÓS
PROJECT MANAGER
SILVIA TÓRTOLA
PROGRAMME & CORE FACILITIES SECRETARIAT
MAGALÍ BARTOMEUS
GEMMA PÉREZ
IMMA FALERO
BLANKA WYSOCKA
RUT CARBONELL
ROMINA GARRIDO
ALBA SHAW
CORE FACILITIES ADMINISTRATION
JAUME BACARDIT
VÍCTOR ELÍAS
NATIONAL FUNDING
ARANTXA FERRER
MARTA SALLÉS
SYSTEMS ADMINISTRATION
OUTSOURCING
HR-FELLOWSHIP AND STUDENTS MANAGEMENT
NURIA JANÉ
MARC GONZÁLEZ
PURCHASING AND LOGISTICS
SERGI REPULLO
MIREIA PÉREZ
MONTSERRAT ESTAPÉ
DAVID RODRÍGUEZ
JUAN CARLOS GÓMEZ
SCIENCE DISSEMINATION
ANNICK LABEEUW
RECEPTION
MARTA CASALS
CONTROL AND PROJECTS JUSTIFICATON
GEMMA SALVADÓ
MARIANA MORLANS
STEFAN PONISCH
ENRIC GUMMÀ
EVA DEL PINTO
JOAN CONANGLA
OLGA JUDERÍAS
ORACLE DEVELOPMENT AREA
ANA CELAYA
LEGAL DEPARTMENT
CRISTINA CASAUS
SCIENTIFIC EQUIPMENTS
NICOLA COLLU
WEB DEVELOPMENT AREA
DAVID CASTILLO
HR SUPPORT
BÁRBARA FERREIRA
MAINTENANCE
JOSEP QUERALT
WORKING RISKS / RADIOACTIVITY / WASTES
SONIA ALCÁZAR
Mª JOSÉ ROMA
EIPMC
ARANTXA VENTURA
ELISABETH MUÑOZ
MARTA POU

SCIENTIFIC ADVISORY BOARD (SAB)

Chairman

Dr. Kai Simons
Max Planck Institute of Molecular Cell Biology and Genetics,
Dresden, Germany

Members

Dr. Stylianos Emmanuel Antonarakis
Medical Genetics, University of Geneva,
Geneva, Switzerland

Dr. Michael Ashburner
European Bioinformatics Institute-EMBL Outstation, Department of Genetics,
University of Cambridge,
Cambridge, United Kingdom

Dr. Pierre Chambon
Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC),
Strasbourg, France

Dr. Iain Mattaj
European Molecular Biology Laboratory (EMBL),
Heidelberg, Germany

Dr. Joan Modolell
Center of Molecular Biology “Severo Ochoa”,
CSIC & Autonomous University of Madrid,
Madrid, Spain

Dr. Arnold Munnich
Det. Génétique, Hôpital des Enfants Malades,
Paris, France

Dr. Christiane Nüsslein-Volhard
Abt. Genetik, Max Planck Institut für Entwicklungsbiologie,
Tübingen, Germany

Dr. Marc Vidal
Dana Farber Cancer Institute,
Boston, USA

Dr. Erwin Wagner
Spanish National Cancer Research Centre,
Madrid, Spain

BUSINESS BOARD

President

Dr. Josep M. Taboada
Medical Director,
Sanofi-Aventis

Vice-president

Dr. Antonio Parente
President,
Grup Lipotec

Members

Dr. Pere Berga
R+D Management Director,
Almirall Prodesfarma, S.A.

Dr. Clara Campàs
Executive Vice-President,
Advancell Therapeutics

Dr. Antoni Esteve
President,
Esteve

Dr. Josep Prous, Jr.
Executive Vice President,
Prous Science

Sr. Jordi Ramentol
General Manager,
Grupo Ferrer Internacional, S.A.

Mr. Josep Lluís Sanfeliu
Co-founder partner,
Ysios Capital Partners

Sr. Fernando Turró
General Manager,
Contratas y Obras

YEAR RETROSPECT

by the Director of the CRG: Miguel Beato

Our environment in 2009 has been marked by the financial and economic crisis and by fears about cuts in the science budget of the public administration. Fortunately, the Catalan Government maintained the commitments as agreed in the strategic plan 2007-2012, which was signed in February, and the CRG remained relatively unaffected by the contraction of the Spanish economy.

Significant changes took place in our advisory organs. After serving six years four members of the Scientific Advisory Board will be replaced at the beginning of next year and two new members were nominated for the expanding computational programmes. Our Business Board was reorganized to include leaders of biotech and venture capital companies in order to make it more efficient in supporting our translational research and more attractive for private sponsors.

Long deliberations on the use of the additional 600 m^2 in the ground floor of the PRBB originally planned for the PRBB and the difficulties encountered in converting these spaces into wet labs, lead to the decision to use them for seminar rooms, bioinformatics groups and administrative offices for the CEXS/UPF and the CRG. The liberated spaces in the main building will be used to concentrate the CRG/UPF core facilities in the 4^{th} floor.

In terms of the development of the Scientific Programmes, two new junior group leaders, Jerome Solon and Bill Keyes, joined the Cell and Developmental Biology and the Differentiation and Cancer programmes, respectively, while a senior scientist Ramin Shiekhattar left the Gene Regulation programme and returned to The Wistar Institute in Philadelphia.

The Core Facilities have been strengthened by the recruitment in January of Doris Meder from the MPI in Dresden as Director of all our facilities. The genomic facility has been reorganized and placed under the leadership of Monica Bayes, with the deep sequencing headed by Heinz Himmelbauer as its most advanced unit. We have continued to improve the equipment of all facilities, in particular the advanced light microscopy facility, for which a two-photon system has been installed that is unique in Europe. Moreover, the first successful siRNA screening was performed in our robotics facility and we have signed an agreement with the UPF for the operation of the common Proteomics Facility. The CRG and the UPF also merged their FACS instruments into a new common facility. A new Bioinformatics Unit was established under the leadership of Guglielmo Roma.

With the objective to serve the growing needs of the CRG the management team has continued to evolve and has updated to the new version of Oracle E-Business Suite under the leadership of David Camargo, in a great effort to reach a more efficient and easy administration process.

In the spring we interviewed 30 candidates for the second 10 positions of the International PhD Programme generously financed by the Obra Social “la Caixa”. The positions were filled with excellent candidates from many different nationalities, who joined the program in the fall. The CRG has initiated activities to support its postdoctoral community. In 2009 a Postdoctoral Committee was created and a novel programme for postdoctoral fellows was launched, the INTERPOD Program, co-financed by the Marie Curie COFUND project of the EU, and directed to interdisciplinary postdocs, who will work bridging groups from different programmes.

In March the second Spanish Worm meeting took place in the CRG and was organized in collaboration with the ICFO. As in the past years, the CRG organized its Annual Symposium in October 2009. This 9th edition was coordinated by the Differentiation and Cancer Programme, focused on “Stem Cells, Differentiation and Cancer” and was a great success of attendance and scientific quality. In November the graduate students organized the III CRG PhD Students Symposium including 20 talks and many posters by second and third year students. In December 2009 members of the core facilities started a new series of short technology symposia. The first took place and concentrated in “Next-Generation Sequencing”.

In collaboration with other institutions of the PRBB, our Communication & PR Dept. team organized an Open Doors Day with multiple events and guided visits. They also organized six scientific coffees on timely topics, workshops for school kids and high school students, a fair on illusions of the brain, a scientific summer camp in the mountains and 5 summer internships at the CRG financed by the Caixa Catalunya Foundation, dissemination talks by group leaders and participated in the Science Week 2009. Also, the first DNA model contest was successfully organized, with 52 models received from schools and kids from all around Catalonia.

2009 has been a good year in terms of competitive resources attracted to the CRG. We obtained a total of 15.7 M€, 9 M€ from national and 6.7 M€ from international calls. Another indicator is that CRG scientists coordinate two of the 13 CONSOLIDER programs assigned to biomedicine in 2009.

At the end of 2009 there were 310 scientists working at the CRG; 29 group leaders, 7 unit heads, 6 staff scientists, 78 postdoctoral fellows, 89 PhD students and 101 technicians. Over 71% of the group leaders, postdocs and graduate students are foreigners.

All together researchers and the management team are contributing to make the CRG an attractive place for young scientists around the world. The SCImago Institution Rankings (SIR) 2009 World Report, based on the citations of publications from 2003-2007, classifies the CRG in position 20 from over 2,000 research institutions around the world. In Europe only four other research centres have a higher citation index. Given that the CRG was officially inaugurated in the fall of 2002, this is a remarkable success. The number and quality of the papers published by the CRG scientists continue to increase as do the number and quality of the seminars by external speakers. In 2009 193 papers were published in international journals with an average impact factor of 8.6, and 141 seminars were held by invited speakers of very high quality. Many of these activities found their way to the public media and the CRG was on the news (newspapers, radio, TV) in 515 occasions. We also elaborated an institutional video that can be seen in our home page and on YouTube.

The future of the Spanish economy is crowded with dark clouds and we will face a few storms next year, but we are confident that the quality and strength of both our scientific community and our management team will manage to cross them without serious damage.

GENE REGULATION

Coordinator: Miguel Beato

In second half of 2009 the programme experienced the departure of one of his senior scientists, Ramin Shiekhattar, who moved back to the Wistar Research Institute in Philadelphia. A call for filling this position was announced in the fall and the seminars and interviews with candidates will take place early next year. For 3 years Ramin has been a very stimulating and productive member of our community and leaves behind several active collaborations with scientists from the Gene Regulation programme and from other programmes.

During the spring we also had seminars and interviews to recruit a colleague to fill the vacant junior group leader position. We are in negotiations with a serious candidate, who will reinforce our research on non-coding RNAs.

In 2009 we also recruited a technician, Marc Batalle, who takes care of the cold room and all the common instruments of the programme. Since his arrival the number of incidences and the general performance of the programme´s infrastructure have improved considerably.

Finally, in the context of the general reorganization of the Core Facilities, the proteomics facility has passed under the control of Doris Meder, the head of all CRG facilities. During the year we also signed a collaboration agreement with the Department of Experimental and Health Sciences (CEXS) of the UPF to regulate management and function of the proteomics facility headed by Henrik Molina.

A common feature of the scientific research in the programme during 2009 has been the increase in the number and volume of genome wide studies, which involved a number of close collaborations with members of the Bioinformatics and Genomics programme. Moreover, groups working on splicing and chromatin in the programme have started a collaboration that includes co-mentoring of a graduate student. A sign of these multiple collaborations is the highlighted paper summarized at the end of this report in which two groups from the programme and a bioinformatics group are co-authors.

An important event in 2009 was the granting of the “Premio Ciudad de Barcelona 2009” in the category of research to Raúl Méndez for his achievement in decoding the rules governing cytoplasmic polyadenylation during meiosis (published in Cell and Nature Cell Biology in 2008).

The structure of the programme at the end of 2009 was

5 Research groups:

> Chromatin and Gene Expression (Miguel Beato, coordinator)
> Regulation of Alternative Pre-mRNA Splicing (Juan Valcárcel)
> RNA-Protein Interactions and Regulation (Josep Vilardell)
> Translational Control of Gene Expression (Raúl Méndez)
> Regulation of Protein Synthesis in Eukaryotes (Fátima Gebauer)

Programme Secretary:

Imma Falero

Numbers:

4 group leaders, 2 staff scientists, 12 postdocs, 20 students,
12 technicians and support personnel.

GENE REGULATION

Group: Chromatin and Gene Expression

Group structure:

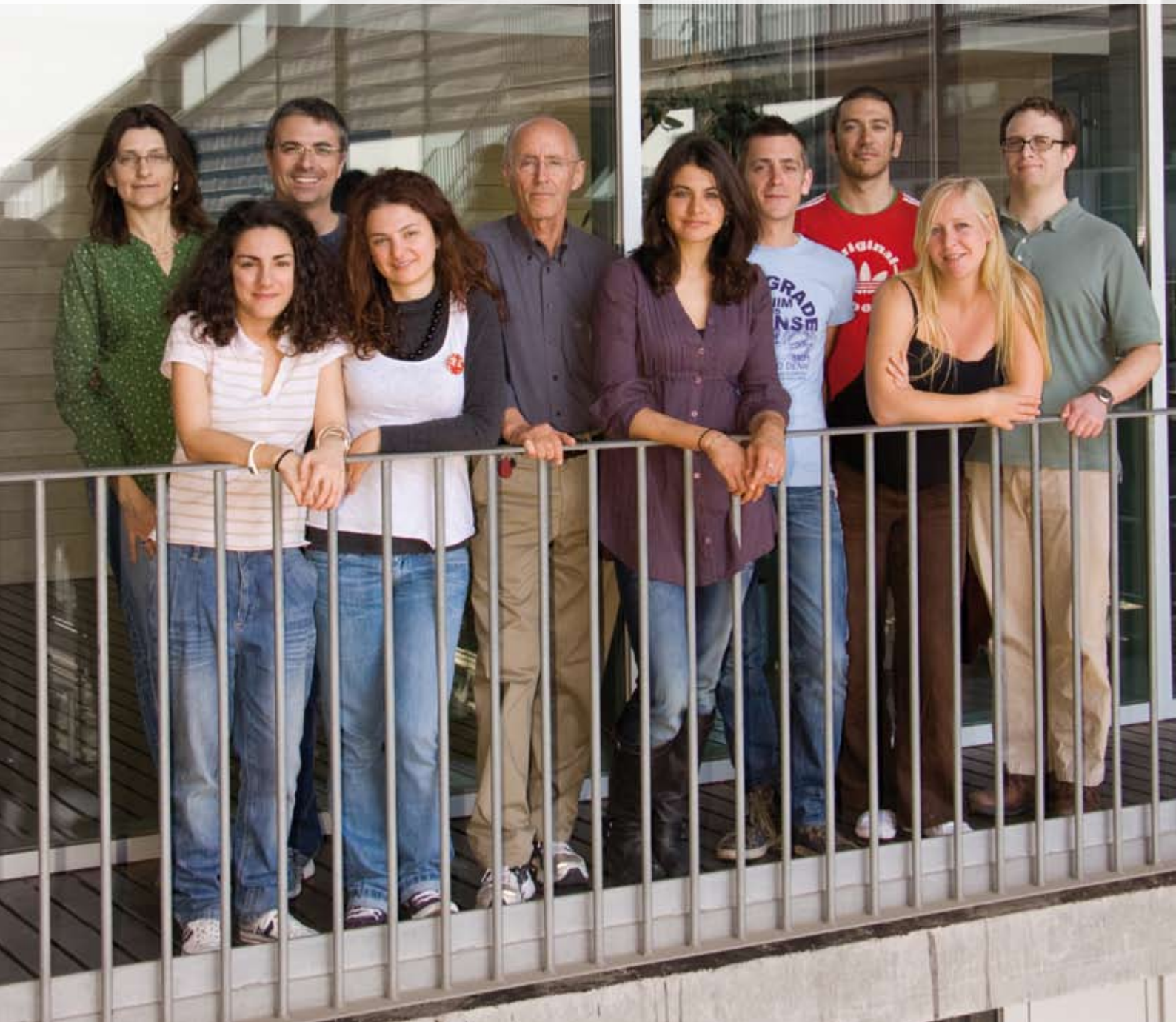
Group Leader: Miguel Beato

Postdoctoral Fellows: Cecilia Ballaré, Alessandra Ciociola, François Le Dily, Guillermo Vicent (Staff Scientist), Roni Wright , Marija Kundakovic

PhD Students: Laura Gaveglia, Andy Pohl (since December), Diana Reyes, Michael Wierer, Roser Zaurin (till September)

Technician/s: Giancarlo Castellano (Consolider), Jofre Font, Silvina Nacht

Visitors: Stefania Mai, Milan (till November), Paola Bertucci, Buenos Aires (August/September 2009)



SUMMARY

The group is interested in understanding how eukaryotic cells respond to external signals, in particular how different signals are transduced to the nucleus to modulate gene expression. Given the relevance of chromatin structure and remodelling on gene regulation, we are studying the topological information that determines the position of nucleosomes and the outcome of the remodelling process. Steroid hormones signal to chromatin no only directly via binding of their receptors to DNA, but also indirectly via crosstalk with kinase signalling pathways. We are trying to unravel how this activated kinases impinge on chromatin structure and dynamics using model systems and genome wide studies. We also try to use this knowledge to decipher the role of steroid hormones in breast and endometrial cancer cell proliferation and differentiation.

RESEARCH PROJECTS

1. Global analysis of chromatin structure and dynamics during hormonal gene regulation

R. Zaurin, L. Gaveglia, R. Wright, F. Le Dily, A. Pohl, G. Castellano (collaboration with Christophoros Nikolau and Roderic Guigo from the Bioinformatics and Genomic programme)

We have developped an algorithm for predicting nucleosome positioning in the human genome that we are currently validating using tiling microarray and massive nucleosome sequencing. We have analyzed a selection of 40 hormone responsive promoters by high resolution tiling arrays to define their nucleosomal structure before and after hormone induction, and have identified several progesterone receptor (PR) binding sites within 10 kb flanking the transcription initiation site. We are correlating these findings with posttranslational modifications of core histones and binding of other transcription factors and chromatin remodelling complexes in selected genes. Currently these results are being extended to the whole genome of breast cancer cells by massive parallel sequencing to determine the positions of nucleosomes before and after hormone addition. An inverse correlation between nucleosome occupancy and the strength of splice signals in internal exons has been identified, indicating a role of nucleosomes in alternative splicing (Tilgner et al, 2009).

2. Signalling by progesterone to kinase cascades and to chromatin

C. Ballaré, A. Ciociola, A .Mai, D. Reyes, M. Wierer, R. Wright, M. Kundacovic

Progesterone controls proliferation and gene expression in breast cancer cells via transient activation of the Src/Ras/Erk pathways mediated by an interaction of PR with estrogen receptor alpha, ER□ (Quiles et al, 2009). Activated Erk in the cell nucleus phosphorylates PR and Msk1, a process essential for the proliferative response of breast cancer cell lines as well as for progestin induction of MMTV and other target genes. We are now studying the mechanism by which Msk1 controls breast cancer cell proliferation in response to estrogens or progesterone in order to develop selective strategies to interfere with the proliferative effect of ovarian hormones.

We have performed gene profiling studies and ChIP-on-chip experiments in breast cancer cell lines to study the response to estrogens and progesterone of hormone-sensitive promoters. The focus is on the identification of the signalling pathways mediating the regulation of different clusters of genes with the aim of modulating specific aspects of hormone action, such as cell proliferation, apoptosis or cell differentiation. We use selective chemical inhibitors and siRNA specific for various kinases (Erk, Msk, Akt, JAK/STAT, CDK2) combined with transcriptome analysis in response to hormones. In parallel we have identified all genomic binding sites for PR in breast cancer cells using ChIP-seq. Finally, quantitative proteomics (SILAC) in breast cancer cells carrying tagged PR is used to delineate the PR interactome. These omics results, along with results from inihibiting ER activation and protein synthesis, are now being integrated in a dynamic network, which should serve to identify relevant nodes connecting various signalling pathways with regulation of different gene cohorts.

3. Regulation of MMTV transcription in the chromatin context

G. Vicent, R. Zaurin, F. Le Dily, S. Nacht, J. Font

The group has studied the structural changes accompanying activation of MMTV promoter chromatin and how they are catalyzed. Within 5 minutes of progestin addition to cells carrying an integrated copy of the MMTV promoter, a ternary complex of activated PR and two activated kinases, pErk1/2 and pMsk1, is recruited to the promoter and phosphorylates histone H3 at S10. This leads to dissociation of a repressive complex containing HP1g as a prerequisite for the recruitment of ATP-dependent chromatin remodelling complexes, co-regulators and RNA polymerase II (Figure 1). Shortly thereafter we detect the displacement of histones H2A and H2B from the promoter nucleosome containing the HREs but not from the adjacent nucleosomes.

Recently, we have identified a key role of the BAF chromatin remodelling complex and the PCAF histone acetyltransferase in hormonal activation of MMTV. Upon hormone addition, BAF is recruited to the promoter by PR, and depleting the BAF complex leads to inhibition of hormonal induction. BAF binding to the MMTV promoter is stabilized by PCAF mediated acetylation of lysine 14 of histone H3 (Vicent et al, 2009). BAF mediated remodelling leads to displacement of H2/H2B dimers and facilitates NF1 access and binding of additional PR to the central HREs on the remodelled nucleosome (Vicent et al, submitted). The new PR molecules bring additional BAF complexes to the promoter and lead to full activation (Figure 1).

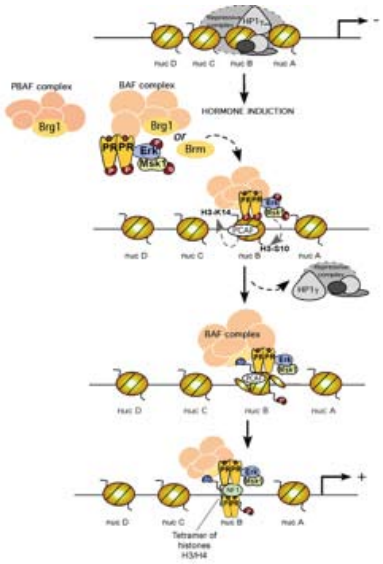


Figure 1.
The MMTV promoter nucleosome B is maintained in a repressive state by bindng of a repressive complex containing HP1g. A complex pPR/pErk/pMsk1 including likely PCAF and BAF, binds to the exposed HRE1 on the MMTV nucleosome and phosphorylates H3 at S10 and acetylates H3 at K14. This leads to displacement of the repressive complex, as a prerequisite for the stable binding of BAF via an interaction with H3K14ac. The BAF complex catalyzes the ejection of H2A/H2B dimers, enabling NF1 binding. In the absence of NF1 this remodelled nucleosome may revert to the closed conformation, but binding of NF1 stabilizes the open nucleosome conformation and facilitates binding of further PR molecules and BAF remodelling complexes to the internal HREs. Subsequently, PR recruits transcriptional co-activators and the transcription initiation complex, leading to promoter activation.

4. Role of steroid hormones in breast cancer and endometrial physiology

C. Ballaré, A. Ciociola, A.Mai, D. Reyes, M. Wierer (collaboration with Belen Miñana, Juan Valcárcel's lab, and the Departments of Pathology and Oncology of the Hospital del Mar).

We have studied the gene networks regulated by estrogens and progestins in primary tumour material from breast cancer patients. In the meantime 108 samples have been analyzed on a proprietary microarray platform containing over 800 cDNAs relevant for breast cancer and hormone action. A comparison of the gene clusters affected in each tumour sample with the results obtained in breast cancer cell lines (see point 2 above) will be used to classify the tumours according to the perturbed signalling pathways. The correlation of this molecular description with clinical and histological data will be used to establish diagnostic and prognostic markers for future cancer management. We are also directly investigating the mechanism of the mutual inhibitory relationship between BRCA1 and PR function.

In collaboration with the group of Patricia Saragüeta, CONICET Buenos Aires, we found that picomolar concentrations of progestins induce proliferation of endometrial stromal cells via activation of a crosstalk of PR with ERß and the mitogenic kinase cascades. We have now studied how blood serum induces the decidual differentiation of these cells, a physiological example of trans-differentiation (Vallejo et al, 2009). We are now comparing these results with the changes in gene expression detected during decidualization of UIII cells in response to the combination of estrogens and progestins.

5. Role of linker histone H1 subtypes in chromatin and transcription

R. Wright, F. Le Dily, A. Pohl

We found that histone H1 enhances the activation of the MMTV promoter by PR and NF1 and are now studying the role of various H1 subtypes and their phosphorylation by Cdk2 on the remodelling and transcription of MMTV chromatin. At least six histone H1 subtypes exist in somatic mammalian cells that bind to the linker DNA and stabilize the nucleosome particle contributing to higher order chromatin compaction. It is not clear whether the different subtypes have distinct roles or if they regulate specific promoters. We have shown that in breast cancer cells depletion of H1.2 caused cell cycle arrest and decreased global nucleosome spacing, whereas H1.4 depletion caused cell death. Using in vivo and in vitro approaches, including atomic force microscopy (AFM), we have investigated the role of histone H1 subtypes and their modifications in the regulation of chromatin spacing and chromatin compaction as well as in the control of chromatin remodelling by ATP-dependent complexes. We find they exhibit different affinity for chromatin and very different capacity to compact chromatin, due mainly to their different C-terminal domains. Minichromosomes containing physiological levels of the different H1 subtypes can be efficiently remodelled by two ATP-dependent remodelling complex, SWI/SNF and NURF (Clausell et al, 2009).

PUBLICATIONS

Clausell J, Happel N, Hale TK, Doenecke D, Beato, M.
"Histone H1 subtypes differentially modulate chromatin condensation without preventing ATP-dependent remodeling by SWI/SNF or NURF."
PLoS One 4, e0007243 (2009).

Quiles I, Millan-Arino L, Subtil-Rodriguez A, Minana B, Spinedi N, Ballare C, Beato M, Jordan A.
"Mutational analysis of progesterone receptor functional domains in stable cell lines delineates sets of genes regulated by different mechanisms."
Mol Endocrinol 23, 809-826 (2009).

Tilgner H, Nikolaou C, Althammer S, Sammeth M, Beato M, Valcarcel J, Guigo R.
"Nucleosome positioning as a determinant of exon recognition."
Nat Struct Mol Biol 16, 996-1001 (2009). Cover picture.

News and Views by Kornblihtt AR et al, titled
"When chromatin meets splicing."
in Nat Struct Mol Biol 16(9):902-3 (2009).

Vicent GP, Zaurin R, Nacht AS, Li A, Font-Mateu J, Le Dily F, Vermeulen M, Mann M, Beato M.
"Two chromatin remodeling activities cooperate during activation of hormone responsive promoters."
PLoS Genet 5, e1000567 (2009).

Vallejo G, Maschi D, Mestre-Citrinovitz AC, Aiba K, Maronna R, Yohai V, Ko MSH, Beato M, Saragüeta P.
"Changes in global gene expression during in vitro decidualization of rat endometrial stromal cells."
J Cell Physiol 222, 127-137 (2009).

Vicent GP, Zaurin R, Nacht As, Font-Mateu J, Le Dily F, Beato M.
"NF1 synergizes with progesterone receptor on the MMTV promoter wrapped around a histone H3/H4 tetrame by facilitatinbg access to the central hormone responsive elements."
J Biol Chem 285(4):2622-31. Epub 2009 Nov 24.

Vicent, GP, Zaurin R, Ballare C, Nachr SA, Beato M.
"Erk signaling and chromatin remodeling in MMTV promoter activation by progestins."
Nuclear Receptor Signaling 7:e008 (2009).

GENE REGULATION

Group: Regulation of Alternative pre-mRNA Splicing during Cell Differentiation, Development and Disease

Juan Valcárcel has a Senior ICREA Group Leader position.

Group Structure:

Group Leader: Juan Valcárcel

Staff scientist: Sophie Bonnal

Postdoctoral Fellows: Sergio Barberán (since December 2009), Elias Bechara, Maria Paola Paronetto, Veronica Raker (until March 2009), Joao Tavanez

Students: Anna Corrionero, Camilla Ianonne, Juan Ramón Tejedor

Technicians: Cecilia Albor (since October 2009), Belén Miñana, Anna Ribó

Diploma Students: Davide Rubbini (from February to November 2009), Mara Bergemann (since May 2009)

SUMMARY

Our group studies molecular mechanisms of alternative pre-mRNA splicing regulation. Alternative splicing of mRNA precursors allows the production of multiple mRNAs from a single gene, greatly expanding the information content of the genomes of multicellular organisms. During 2009, we have made progress in understanding the impact on splice site selection of RNA secondary structures, of nucleosome positioning on exonic DNA sequences and of the activation of signaling pathways triggered by exposure to UV irradiation. In addition, we have continued genomic and mechanistic analyses of splicing regulation of genes important for cellular proliferation and apoptosis.

RESEARCH PROJECTS

1. Impact of RNA secondary structures in splice site selection

Despite the need for high accuracy in the splicing process, the RNA sequences that determine the boundary between exons and introns are rather short and degenerate, raising the question of how splice sites are identified, paired and regulated. In collaboration with D. Pervouchine (Moscow), our group has used phylogenetic comparisons between 12 Drosophila genomes to identify highly conserved pairs of intronic sequences capable of forming base pairing interactions over long RNA distances, which help to bring together and/or select splice sites. The functional relevance of these sequences was confirmed experimentally in cells in culture. Thus, long-range secondary structures are frequent modulators of alternative splicing.

2. Impact of chromatin structure on splice site identification and selection

In collaboration with the groups of Roderic Guigó and Miguel Beato, we are exploring the possibility that chromatin structure influences splice site recognition in pre-mRNAs that are processed co-transcriptionally. In particular, we found that nucleosome positioning correlates with the distribution of exonic sequences in the DNA, an effect which is more evident for exons flanked by weak splice site signals. We are particularly interested in chromatin modifications that influence alternative splicing regulation induced by steroid hormones and by polymorphic changes in DNA sequence relevant for human disease.

3. Genome-wide analyses of alternative splicing changes induced by UV light

In collaboration with the group of Alberto Kornblihtt (University of Buenos Aires), we have analyzed changes in alternative splicing induced by irradiation of cells in culture with UV light. Our results document hundreds of changes in alternative processing, frequently linked to decreased expression levels associated to changes in chromatin compaction triggered by UV light irradiation, suggesting coupling between transcription elongation rates and splice site selection. We are currently exploring the mechanistic basis for these effects, including alterations in the function of key splicing regulators induced by UV light.

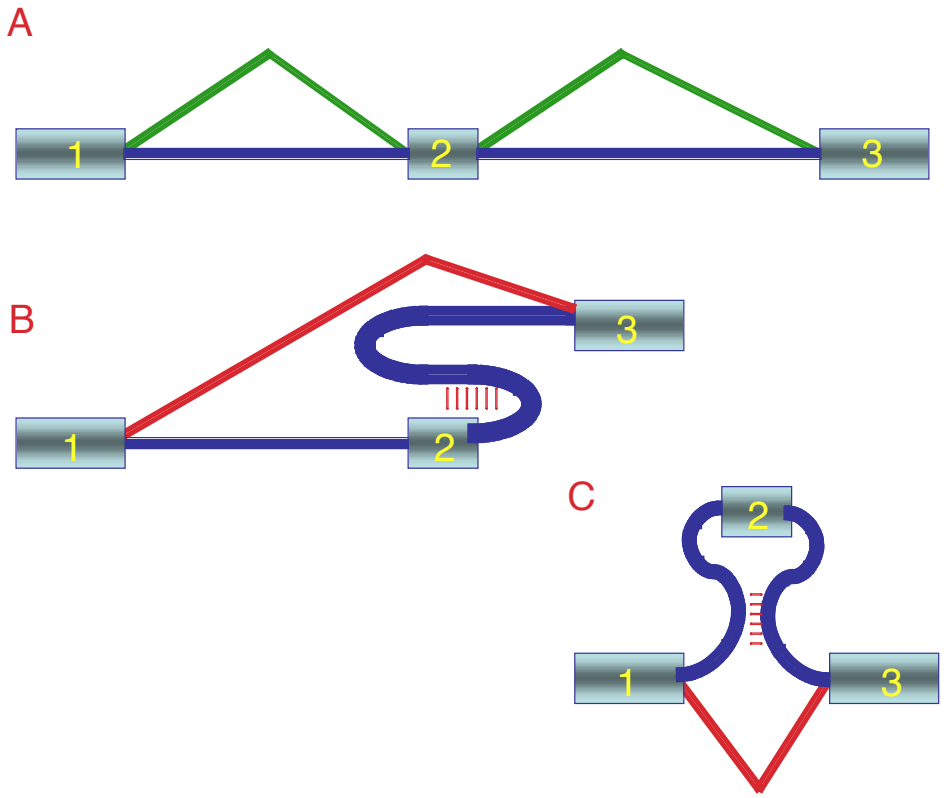


Figure 1. *Mechanism by which secondary RNA structures can modulate splice site selection. A. In this three exon / 2 intron model pre-mRNA, exon 2 is included (green lines) in the absence of RNA secondary structures. B. Base pairing interactions between intronic and exonic regions sequester splice sites and prevent exon 2 inclusion. C. Base pairing interactions between intronic regions flanking exon 2 bring exons 1 and 3 closer to each other, facilitating splicing between them.*

4. Alternative splicing regulation and the control of cell proliferation and apoptosis

We have continued to study the function of splicing factors RBM5 and EWS, which have properties of tumor suppressors, in alternative splicing regulation. Our approach combines genome-wide analysis of their binding sites in pre-mRNAs, using crosslinking/immunoprecipitation (CLIP), with analysis of changes induced by their depletion, using splicing-sensitive microarrays. These datasets are allowing us to draw regulatory maps that can help to predict the targets and functions of these regulatory factors. Consistent with their roles in the control of cell proliferation, genes regulated by RBM5 and EWS play important roles in the control of cell cycle and of programmed cell death.

PUBLICATIONS

Corrionero, A and Valcárcel, J.
"RNA Processing: redrawing the map of charted territory."
Molecular Cell, 36:918-919 (2009).

Hartmann B, Castelo R, Blanchette M, Boue S, Rio DC and Valcárcel J.
"Global analysis of alternative splicing regulation by insulin and wingless signalling in Drosophila cells."
Genome Biology, 10:R11 (2009).

Hartmann B and Valcárcel J.
"Decrypting the genome's alternative messages."
Current Opinion in Cell Biology, 21:377-386 (2009).

Martins-Araujo M, Bonnal S, Hastings ML, Krainer AR and Valcárcel J.
"Differential 3' splice site recognition of SMN1 and SMN2 transcripts by U2AF and U2 snRNP."
RNA, 15:515-523 (2009).

Muñoz MJ, Pérez Santangelo MS, Paronetto MP, de la Mata M, Pelisch F, Boireau S, Glover-Cutter K, Ben-Dov C, Blaustein M, Lozano JJ, Bird G, Bentley D, Bertrand E, Kornblihtt AR.
"DNA damage regulates alternative splicing through inhibition of RNA polymerase II elongation."
Cell, 137:708-20 (2009).

Raker VA, Mironov AA, Gelfand MS and Pervouchine DD.
"Modulation of alternative splicing by long-range RNA structures in Drosophila."
Nucleic Acids Res, 37:4533-44 (2009).

Tilgner H, Nikolau C, Althammer S, Sammeth M, Beato M, Valcárcel J and Guigó R.
"Nucleosome positioning as a determinant of exon recognition."
Nature Structural & Molecular Biology, 16:996-1001 (2009).

GENE REGULATION

Group: RNA interference and Chromatin Regulation
Ramin Shiekhattar has a Senior ICREA Group Leader position.

Group Structure:

Group Leader: Ramin Shiekhattar

Staff Scientist: Glenda Harris

Postdoctoral Researchers: Ildem Akerman, David Baillat, Malte Beringer, Clement Carre, Klaus Fortschegger Alessandro Gardini, Fan Lai, Frederic Tort, Laure Weill

Predoctoral Researcher: Dimitra Saviolaki

Technician: Albane Lamy de la Chapelle, Leonor Avila

SUMMARY

We are interested to understand the mechanisms by which protein-coding genes are silenced and how disruption of negative regulation may lead to neoplasia. Such genomic repression is achieved at both transcriptional and post-transcriptional levels. Recent evidence has implicated changes in chromatin structure as an important mechanism in gene regulation. Moreover, besides a classic role for proteins in mediating transcriptional and post-transcriptional effects, it has become clear that non-coding RNAs play an integral role in silencing through fine-tuning gene expression patterns during development and differentiation. We believe that this "signaling by RNA" is an emerging field of research that will uncover novel pathways in genome organization and regulation. In the past few years my laboratory has employed a biochemical approach to uncover a number of key factors that mediate transcriptional repression through regulation of chromatin structure, or post-transcriptional silencing via non-coding RNAs. We will continue our studies through a detailed structure/function analysis of these factors in regulation of genomic silencing and extend our experiments to elucidate the role these factors in the genesis of cancer.

RESEARCH PROJECTS

1. Transcriptional regulation through chromatin modifying complexes

Using biochemical approaches my laboratory has identified novel co-repressor complexes (SMRT-complex; LSD1-containing complexes; JARID1d/Ring6a) that mediate tissue and gene-specific transcriptional repression through modification of chromatin structure. These studies have been instrumental in understand mechanisms by which nuclear hormone receptors and tissue-specific transcriptional repressors (such as the neuronal silencer, REST), mediate their biological functions – these activities include remodelling of nucleosomes and lysine demethylation of histones. More recently, our studies have identified an intimate connection between the polycomb repressive group of proteins and histone demethylases (Figure 1) that underscore their cooperation in maintaining gene expression patterns.

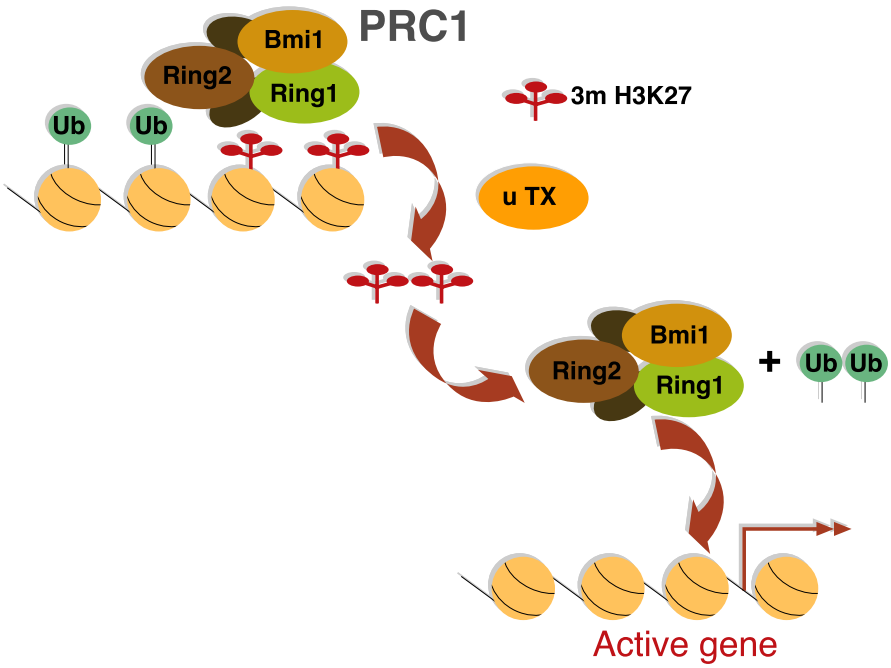


Figure 1.
UTX a histone demethylase specific for histone H3 lysine 27 increase the occupancy of polycomb complex PRC1 on target promoters

2. Transcriptional and post-transcriptional regulation by non-coding RNAs

My laboratory has also identified a pathway by which small non-coding RNAs, termed microRNAs (miRNAs), are synthesized and their mechanism in mediating post-transcriptional silencing in human cells. We identified two distinct RNase III-containing complexes, Drosha/DGCR8 and Dicer/TRBP, which mediate the stepwise processing of primary miRNAs to mature miRNAs. The last step of miRNA processing by Dicer/TRBP involves the incorporation of mature miRNA into a larger complex composed of Argonaute 2 (Figure 2), the subunits of 60S ribosome and the ribosome anti-association factor eIF6. Our studies demonstrate a key role for eIF6 (through ribosome disruption) as the mediator of miRNA inhibitory function.

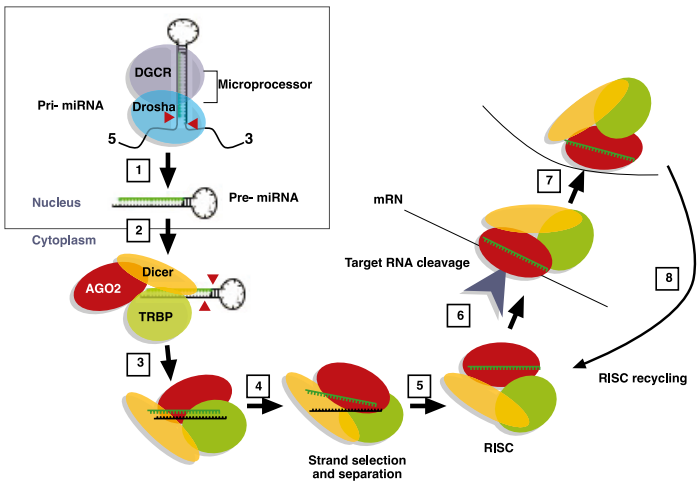


Figure 2.
Steps of microRNA biogenesis and RNA-induced silencing complex (RISC) loading designated by numbers 1 through 8.

PUBLICATIONS

Baillat D, Shiekhattar R.
"Functional Dissection of Human TNRC6, GW182-related, Family of proteins."
Mol Cell Biol, 29(15):4144-55. Epub 2009 May 26 (2009).

Berger SL, Kouzarides T, Shiekhattar R, Shilatifard A.
"An operational definition of epigenetics."
Genes Dev, 23(7):781-3 (2009). (*)

Melo SA, Ropero S, Moutinho C, Aaltonen LA, Yamamoto H, Calin GA, Rossi S, Fernandez AF, Carneiro F, Oliverira C, Ferreira B, liu CG, Villaneva A, Capella G, Schwartz Jr S, Shiekhattar, R, Esteller M.
"A TARBP2 mutation in human cancer impairs microRNA processing and DICER1 function."
Nature Genetics, 41(3):365-70 (2009).

(*) This publication is the result of the work of Dr. Ramin Shiekhattar at the Wistar Institute, Philadelphia, USA.

GENE REGULATION

Group: RNA-Protein Interactions and Regulation

Group Structure:

Group Leader:	Josep Vilardell
Postdoctoral Researcher:	Maria Camats
PhD Student:	Markus Meyer
“Leonardo” Student:	Valeria Giangarrà
Practice Student:	Mercè Guzman
Visiting Student:	Bernardo Cuenca

SUMMARY

We are focused on the study of the molecular mechanisms of regulation of gene expression by RNA-protein interactions, using the model organism ***Saccharomyces cerevisiae***. We have two main research goals in our laboratory. First, to dissect the molecular interactions involved on ***RPL30*** regulation of splicing. Second, to study the mecanisms that control the selection of 3' splice sites in yeast. A third project is a joint effort with Susana-Rodriguez Navarro (Principe Felipe, Valencia), centered on the possibility of exon-definition mechanisms in yeast, which would offer a powerful tool to study a complex problem.
In addition, in collaboration with other groups, we are undertaking bioinformatic and genomic approaches to uncover new instances of regulation, and to investigate how spread is this control of gene expression and its degree of coordination.

RESEARCH PROJECTS

1. Genetic screen to select mutants in regulation of splicing

Mercè Guzman, Maria Camats

RPL30, one of the best understood models of splicing regulation in ***Saccharomyces***, encodes the ribosomal protein L30. Through binding to a structure present in its own transcript L30 can regulate RNA processing at several steps. Our main interest is on control of splicing (see Fig 1), and the L30 system of regulation should provide insights on how RNA sequences and structures near the 5' splice site (ss) can affect spliceosome assembly and splicing.

Employing refined screening methods we have isolated mutants that alter L30 regulation of splicing in either way: some mutants display a phenotype in which L30 can not regulate splicing of a target transcript; while others show the opposite behavior, in which L30 can regulate splicing of a transcript bearing a mutation that blocks L30 inhibition in wild type cells. One of the latter mutations is located in the gen STO1, encoding the large subunit of the cap binding complex (CBC), Cbp80. CBC has been shown to be required for the proper stability of U1 snRNP bound to the pre-mRNA, and we are investigating how this can affect L30 regulation. Our data indicate that Cbp80 plays a role in a proposed remodelling of the interaction between U1 snRNP and the intron. This role becomes more apparent in introns where this interaction is strong, such as RPL30. The intron-U1 snRNP remodeling is the likely target of L30, which would explain the genetic interaction between L30 and STO1 (Bragulat et al., 2nd submission to *Molecular Cell*).

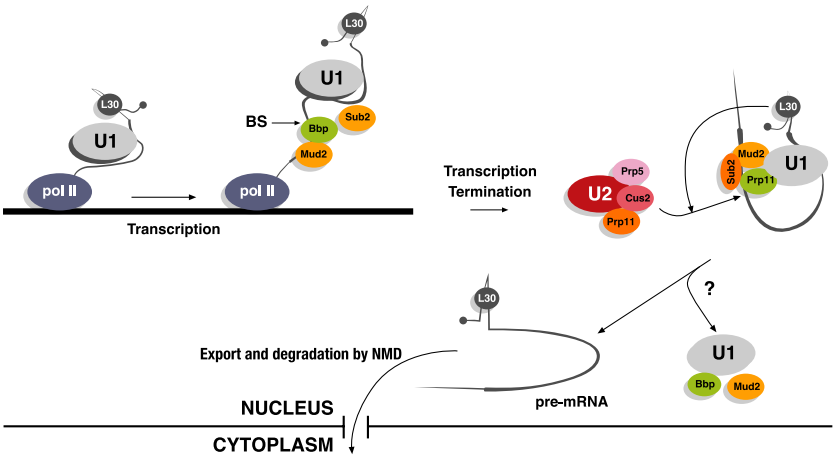
The relevance of this system of regulation, evidenced by our biochemical dissection, makes it relevant the continuation of this genetic approach. In addition to new screenings, we are asking the possible role of mutations in factors like BBP, Mud2, Prp5, or Sub2.

2. Biochemical analysis of the mechanism of *RPL30* control of splicing.

Maria Camats, Markus Meyer

As shown in Fig. 1, during L30 regulation of splicing a new complex is formed. We are interested on the nature of this complex, its components and their interactions. Several ***in vitro*** and ***in vivo*** approaches produced data suggesting that this system of regulation follows a novel strategy, based in interfering with the rearrangements that occur during spliceosome assembly. We have been able to link the role of L30 as a splicing factor with transcription, providing novel evidence about the connection of regulated splicing and pre-mRNA synthesis. Some of the pre-spliceosomal components show altered ChIP profiles during regulation, and we are investigating the relevance of these findings.

Figure 1.
Regulation of RPL30 splicing by L30. L30 binds to the RPL30 nascent transcript; however, initial intron recognition is not affected. Inhibition of splicing occurs because the L30-containing nascent spliceosome fails to undergo a necessary remodelling that is required for the joining of U2 snRNP. Eventually the "failed" pre-spliceosome will be disassembled and the pre-mRNA exported to the cytoplasm. There is will be degraded by NMD since it contains premature stop codons.



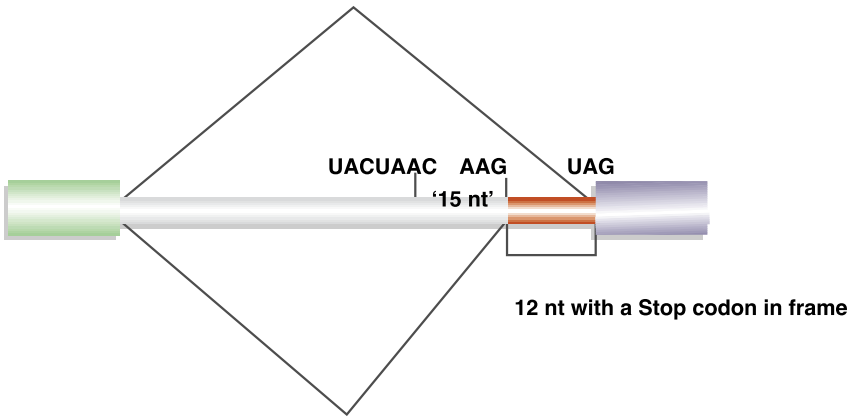
3. Other *S. cerevisiae* genes with regulated splicing

Markus Meyer, Valeria Giangarrà

To what extent splicing regulation plays a role in the control of gene expression in *Saccharomyces*? Using bioinformatics, several studies are being performed. We are looking at (1) the possible folding around 3' splice site regions; and (2), putative alternate 3'SS. In both cases we include phylogenetic comparisons in our studies. We are collaborating in these efforts with the group of Dr. Eduardo Eyras (UPF).

As a fruit of these efforts a novel case of dual use of 3'SS has been uncovered (Fig. 2), and other cases show discrepancies with the annotated sequences in our experimental conditions.

Figure 2.
DMC1 shows dual use of its 3'SS. The shorter version of the intron produces an mRNA with a premature stop codon and it's degraded by the non-sense mediated decay pathway.



In addition, we are collaborating with the group of Dr. Jean Beggs (University of Edinburgh) to apply genomic approaches to this question. We have supplemented their Saccharomyces splicing microarray with additional probes that can detect new, predicted, 3'SS splicing events, and we are studying splicing and 3'SS usage in splicing in several conditions, including meiosis, mutations in splicing factors, and several stresses (collaboration with Francesc Posas, UPF).

In addition to the microarray approaches, we have started using deep-sequencing techniques, in collaboration with the group of Roderic Guigó. This new approach offers us considerable advantages, especially regarding the identification of new spliced forms.

The combination of all these efforts has led us to a deep understanding of the definition of 3' splice sites in yeast. We have uncovered the rules followed by the yeast spliceosome in the selection of 3' splice sites, and currently we can explain all but one of the ***Saccharomyces*** introns (Meyer et al., in preparation). This provides an outstanding base to further study splicing mechanics and its regulation. Remarkably, intron folding plays a crucial role in the splicing of at least one third of introns, and there is notably little selective pressure in keeping the introns small. Given the evolutionary conservation of the splicing process, these findings are likely to be relevant to human, where these studies are not practical.

4. Exon definition in yeast?

Bernardo Cuenca

In a joint effort with Susana-Rodríguez-Navarro (Instituto Príncipe Felipe, Valencia), we are studying the splicing of the gene SUS1 (Fig. 3), involved in transcription and RNA export. This gene contains two introns, and some spliced forms show exon 2 skipping. We are investigating the mechanisms involved in the splicing of this gene; with an interest in asking, in vitro and in vivo, whether the mechanisms involved in exon skipping in metazoans are also having a role in this phenomenon.

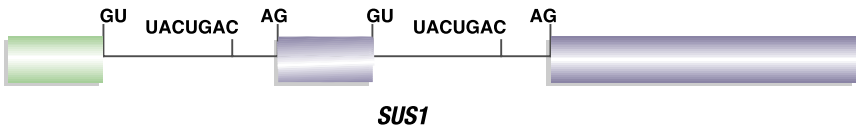


Figure 3.
Genomic structure of the SUS1 gene.

PUBLICATIONS

Meyer M and Vilardell J.
"The quest for a message: budding yeast, a model organism to study the control of pre-mRNA splicing."
Briefings in Functional Genomics & Proteomics, 8(1):60-7 (2009).

GENE REGULATION

Group: Regulation of Protein Synthesis in Eukaryotes

Group Structure:

Group Leader:	Fátima Gebauer
Lab manager:	Olga Coll
Postdoctoral Researchers:	Marija Mihailovic, Antoine Graindorge
PhD Students:	Aida Martínez, Ana Villalba, Cristina Militti, Emilia Szostak
Master Student:	Nicole Busche
Technician:	Anna Ribó

SUMMARY

We are interested in the regulation of mRNA translation during embryonic development and cell homeostasis. We wish to understand the molecular mechanisms of translational control exerted by RNA-binding proteins, and the way translational regulators are wired in the cell for appropriate cell function. Our lab focuses on translational control events that are essential for early embryonic development and viability in the fruitfly ***Drosophila melanogaster***. By studying conserved regulators in ***Drosophila***, we expect to extract general principles used also in mammalian cells which could go awry in disease.

RESEARCH PROJECTS

1. Translational control of dosage compensation

X-chromosome dosage compensation is the process that equalizes the expression of X-linked genes in males (XY) and females (XX). Dosage compensation is essential for life and is initiated early during embryonic development. In ***Drosophila***, dosage compensation is achieved by hypertranscription of the male X chromosome as a consequence of the binding of the dosage compensation complex (DCC) to hundreds of sites on the X (Figure 1). In females, dosage compensation is repressed via the translational inhibition of the rate-limiting DCC component MSL-2. At least two RNA-binding proteins are involved in this repression: the female-specific protein Sex-lethal (SXL) and the ubiquitous protein Upstream of N-ras (UNR). SXL and UNR form a complex that binds to the 3' UTR of msl-2 mRNA, but they are not sufficient to repress translation. Part of our current research focuses on the isolation and characterization of other co-factors involved in msl-2 translational repression, and the investigation of the molecular mechanisms of msl-2 translational control.

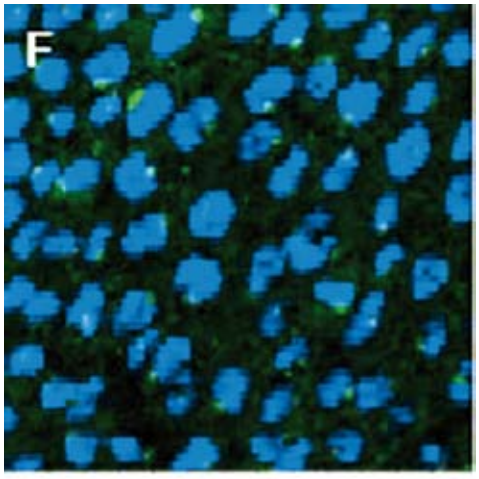


Figure 1

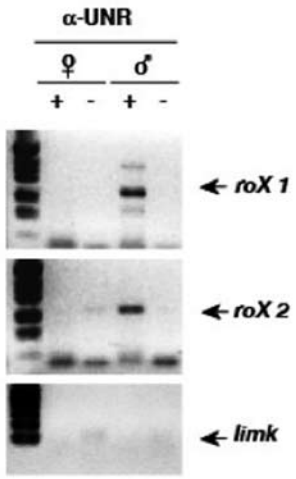


Figure 2

Figure 1.
The DCC complex binds to the male X chromosome. Staining with anti-MSL2 antibodies (green) reveals the X chromosome territory in male cells. DNA is stained with DAPI (blue). Image obtained in collaboration with Nuria Paricio (University of Valencia).

Figure 2.
UNR binds to the non-coding roX RNAs. IP of UNR followed by RT-PCR to detect the indicated RNAs.

Using genetics we have observed that UNR performs opposite, sex-specific functions in dosage compensation: while UNR inhibits DCC formation in females, it promotes DCC binding to the X chromosome in males. Curiously, UNR binds to the non-coding RNA components of the DCC (Figure 2). We are testing the possibility that this binding underlies the mechanism of UNR function in males.

Dosage compensation is an example of how a general RNA binding protein (UNR) displays sex-specific functions by binding to different targets in the two sexes. UNR contains multiple cold shock domains, which in principle could serve to bind and regulate many other targets. We are investigating the sex-specific RNA regulatory networks that are controlled by UNR.

2. Translational regulation by cytoplasmic polyadenylation

Embryonic axis formation in ***Drosophila*** heavily depends on translational control. Expression of the key morphogens Bicoid and Toll is activated at precise times in development by a process called cytoplasmic polyadenylation. The sequences and factors regulating cytoplasmic polyadenylation in ***Drosophila*** are largely unknown. Using a cell-free cytoplasmic polyadenylation system obtained from early embryos we have found that Toll mRNA is polyadenylated by a novel, non-canonical mechanism (Figure 3). Other transcripts follow the canonical pathway, which requires the cytoplasmic polyadenylation element (CPE) and the polyadenylation hexanucleotide. We are currently trying to isolate the machineries responsible for both canonical and non-canonical polyadenylation.

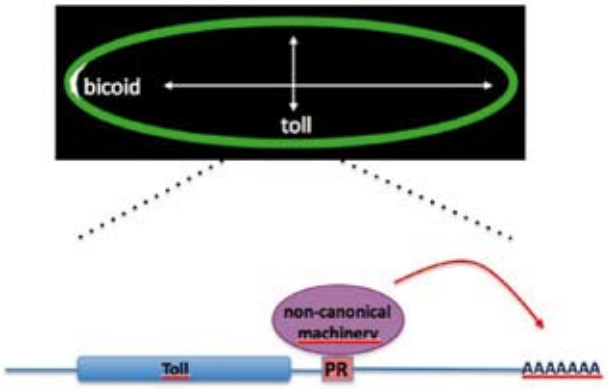


Figure 3.
Toll mRNA is polyadenylated by a non-canonical machinery.

PUBLICATIONS

Coll O, Villalba A, Bussotti G, Notredame C and Gebauer F.
"A novel, non-canonical mechanism of cytoplasmic polyadenylation operates in Drosophila embryogenesis."
Genes Dev, in press.

Mihailovich M, Militti C, Gabaldón T and Gebauer F.
"Eukaryotic cold shock domain proteins: versatile regulators of gene expression."
BioEssays, in press.

Cuesta R, Martínez-Sánchez A and Gebauer F.
"miR-181a regulates cap-dependent translation of p27kip1 mRNA in myeloid cells."
Mol Cell Biol, 29:2841-2851 (2009).

Patalano S, Mihailovich M, Belacortu Y, Paricio N and Gebauer F.
"Dual sex-specific functions of Drosophila Upstream of N-Ras in the control of X chromosome dosage compensation."
Development, 136:689-698 (2009).

GENE REGULATION

Group: Translational Control of Gene Expression

Group structure:

Group Leader: Raul Méndez

Postdoctoral Researchers: Isabel Novoa (Ramón y Cajal-awarded, till November 2009), Maria Piqué , David Pineda (Juan de la Cierva awarded), Carolina Segura, Laure Weill (from March 2009)

Students: Ana Igea (Graduate Student, PhD, November 2009), Alessio Bava ("la Caixa" Foundation Fellow, Graduate Student), Valeria Giangarra (Graduate Student, September 2009), Vittorio Calderone (Graduate Student, September 2009), Jordina Guillen (Graduate Student, September 2009)

SUMMARY

The primary interest of our group is to understand the molecular mechanisms that dictate the temporal and spatial translational control of specific mRNAs during cell cycle progression and early embryonic development. Meiotic progression and early development are programmed, at least in part, by maternally inherited mRNAs. These mRNAs are not translated en masse at any one time, or even at any one place; rather, their translation is specifically regulated by sequences located at the 3´-untranslated region (3´-UTR) of the mRNA and their binding proteins. Cytoplasmic polyadenylation is one of the most important mechanisms for regulating translation during meiotic progression and is directly controlled by the RNA-binding protein CPEB. The work of our group focuses on four questions in the area of translational regulation by cytoplasmic changes in the poly(A) tail length of mRNAs encoding for factors that drive cell cycle progression: 1) Genome-wide identification of the mRNAs that are regulated by cytoplasmic changes in their poly(A) tail length; 2) Determination of the configuration of cis-acting elements that define the temporal and spatial translational regulation by CPEB; 3) Role of the localized CPE-mediated translational regulation in meiotic progression; and 4) Identification of the cell cycle-related events regulated by the CPEB family of proteins.

RESEARCH PROJECTS

1. Cytoplasmic polyadenylation and the CPEB familly of proteins

Early animal development is directed by maternal mRNAs that are stored in the developing oocyte, until subsequent use during the late stages of meiosis or after fertilization. These mRNAs contain short poly(A) tails (~20-40 nts), and it is only when these tails are elongated in response to progesterone stimulation (to ~150 nts) that translation take place. Polyadenylation requires an element in the 3'UTR named cytoplasmic polyadenylation element (CPE). The CPE is bound by a RNA-binding protein named CPE Binding protein 1 (CPEB1). CPEB1 is the best characterized and founder member of a family of four proteins conserved in their RNA-recognition domain but distinct in their regulatory motifs (Reviewed in Mendez, 2001). In Xenopus oocytes, CPEB-1 can assemble two different functional complexes to either repress translation or to mediate cytoplasmic polyadenylation and activate translation. The switch from the «repressing-state» to the «activating-state» is driven by phosphorylation by Aurora-A and Cdc2/Plk1 kinases (Mendez, 2000a, Mendez 2000b, Mendez 2002). The unphosphorylated CPEB1 recruits a set of factors that shorten the poly(A) tail and blocks the recognition of the cap by the translational machinery. Once phosphorylated, CPEB1 changes its co-factors to recruit the cytoplasmic polyadenylation machinery and enhance the association of the translation initiation factors to the CPE-regulated mRNA. In two recent works (Pique 2008, Belloc and Mendez 2008) we have performed a systematic analysis of the combinations of cis-acting elements that define, qualitatively and quantitatively, the differential translational control of CPE-regulated mRNAs during meiosis. Overlaying this temporal control of translational activation, the same combinatorial pattern of CPEs contains the spatial information defining the oocyte subcellular localization where the mRNAs will be activated (Eliscovich 2008). All together our results indicate that the CPEBs control, in time and space, the translation of hundred of mRNAs encoding factors implicated in cell cycle, cell differentiation, angiogenesis, inflammation and other cellular functions. The CPEBs act, therefore, as hubs controlling the coordinated expression of up to 20% of the genome (Belloc et al. 2008).

Although the function of the CPEBs has been elucidated mainly in the xenopus oocytes, we have recently found that CPEB-mediated post-transcriptional regulation by phase-specific changes in poly(A) tail length is not a meiotic specific mechanism, but also required for cell proliferation and specifically for M-phase entry in mitotically dividing cells, where CPEB1 and CPEB4 regulate the polyadenylation of mRNAs encoding factors differentially expressed during cell cycle (Novoa, In press).

Defining the translational control circuit that regulates progression through the two meiotic divisions.

Recent works from our group have contributed to define a molecular circuit to temporally regulate translational activation of mRNAs encoding the factors that mediate phase transitions during meiosis. This circuit is stabilized by a number of positive and negative feed back loops that ensure the irreversibility of the process, produce hysteric responses (switch-like transitions) and, yet, allow the key kinase activities that drive meiotic progression to oscillate between the two meiotic divisions (figure 1). One essential

negative feed-back loop is mediated by the CPEB1-driven translational activation of mRNA encoding C3H-4, an ARE-binding protein that we find to accumulate in MI and whose ablation induces meiotic arrest. C3H-4, in turn, opposes CPEB activity on mRNAs containing both CPEs and AREs by recruiting the CCR4-deadenylase complex to the ARE-containing mRNAs (Belloc and Mendez 2008). However, this negative feed-back loop, required to exit the first meiotic metaphase, has to be inactivated to start the second meiotic division. We are studying how this inactivation is accomplished.

We have also defined an essential positive feed back loop where CPEB1 regulates the translation of CPEB4 mRNA. We have shown that CPEB4 can regulate the translation and mediate cytoplasmic polyadenylation of the same mRNAs that CPEB1, but at different meiotic times. Thus, while CPEB1 regulates the early maternal mRNAs during the first meiotic division CPEB4 regulates the late mRNAs during the second meiotic division. Altogether our work shows that CPEB4 expression is controlled by CPEB1 and, in turn CPEB4 replaces CPEB1 when it is degraded in anaphase I. But CPEB1 and CPEB4 are activated by different kinases and at different meiotic phases (Igea and Mendez, under review). We are currently studying the posttranslational regulation of CPEB4, which, contrary to CPEB1, is not activated by Aurora A kinase but rather by an unknown kinase activated latter during cell cycle. Mapping the phosphorylated residues in CPEB4, characterizing how that regulates cytoplasmic polyadenylation and identifying the regulatory kinase are part of our future goals. Delineating the signal transduction pathways that regulate each CPEB will not only contribute to better define the meiotic circuit, but should be extrapolable to somatic cells and different extracellular signals.

Regulation of gene expression by CPEBs in somatic tissues.

2. Translational control of mitotic cell cycle

We have shown that CPEB-mediated post-transcriptional regulation by phase-specific changes in poly(A) tail length is required for M-phase entry and cell proliferation in mitotically dividing cells. This translational control is largely mediated by two members of the CPEB-family of proteins, CPEB1 and CPEB4, which regulate the polyadenylation of mRNAs encoding factors differentially expressed during cell cycle. We conclude that regulation of poly(A) tail length is not only required to compensate for the lack of transcription in specialized cell divisions, but acts as a general mechanism to control mitosis (Novoa et al., in press; Figure 2). We continue to investigate the differential functions of CPEB1 and CPEB4 during mitotic divisions and their phase specific regulation by different signal transduction pathways. We are also studying in great detail their functions in chromosome segregation. In collaboration with the Isabelle Vernos group (CRG), we have shown that spindle-localized translational activation, by cytoplasmic polyadenylation, of the mRNAs encoding for TPX2 (Targeting Protein for *Xenopus* kinesin-like protein 2) and XKid (*Xenopus* Kinesin-like DNA binding protein) is essential to complete the first meiotic division and also for chromosome segregation in *Xenopus* oocytes (Eliscovich et al. 2008; Figure 3). We have also found that CPEB and co-factors in the translational repression and activation complexes are asymmetrically distributed within the mitotic spindle in somatic cells. We are analyzing whether these factors generate translational gradients that define the polarity of the spindle and performing a Genome-wide identification of spindle translated mRNAs.

3. CPEB-mediated translational control in pancreatic ductal adenocarcinomas (PDAs)

In collaboration with Pilar Navarro (IMIM/PRBB) and Paco Real (CNIO), we have found that members of the CPEB family of proteins are overexpressed in tumors, resulting in abnormal regulation of mRNAs encoding pro-tumoral proteins. Moreover, nude mice injection of human pancreatic tumoral cells in which the levels of CPEB4 have been stably knocked down results in much reduced tumor growth, invasion and vascularization. Our results indicate that overexpression of CPEB1 and CPEB4 could reprogram gene expression in tumoral cells causing abnormal expression of growth, invasion and angiogenic factors, which will promote tumor growth and vascularization in PDA (Ortiz et al. Manuscript in preparation).

Figure 1.

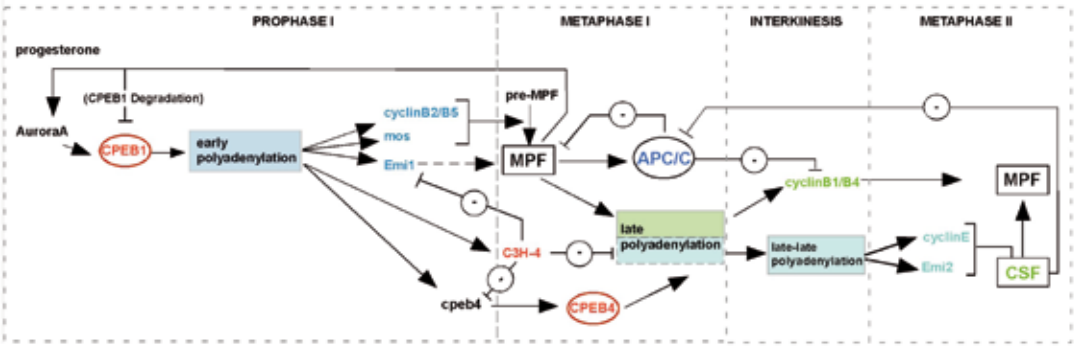


Figure 2.

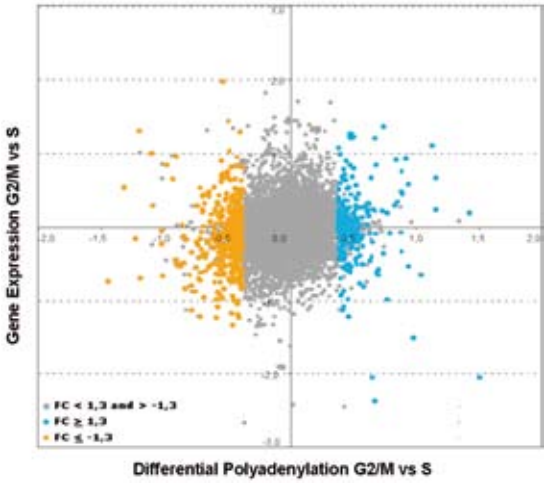
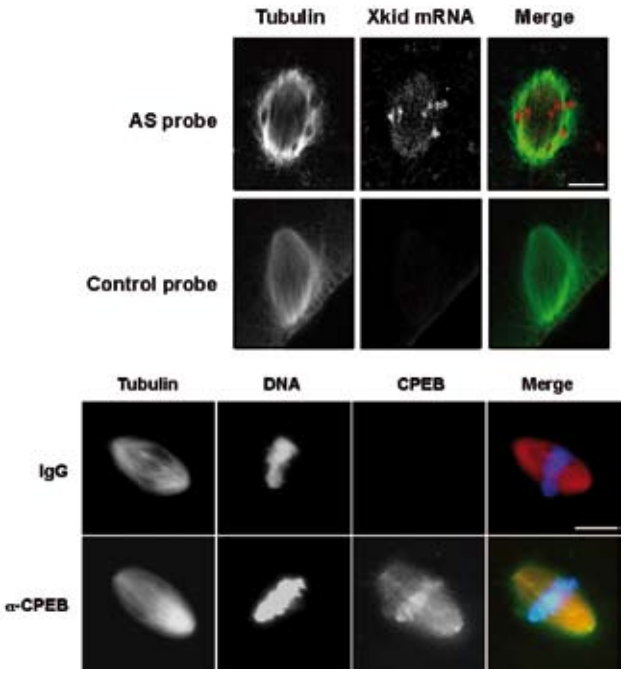


Figure 3.



PUBLICATIONS

Mejías M, García-Pras E, Gallego J, Méndez R, Bosch J and Fernández M.
"Relevance of the mTOR signaling pathway in the pathophysiology of splenomegaly in rats with chronic portal hypertension."
Journal of Hepatology, in press.

Novoa I, Gallego J, Ferreira PG, Méndez R.
"Mitotic cell cycle progression is regulated by CPEB-dependent translational control."
Nature Cell Biology, in press.

Book chapter

Novoa I, Eliscovich C, Belloc E and Mendez R.
"Oocyte-specific translational control mechanisms."
In: Oogenesis: The Universal Process Marie-Helene Verlhac and Anne Villeneuve, 2010 John Wiley & Sons, Ltd., in press.

DIFFERENTIATION AND CANCER

Coordinator: Thomas Graf

The areas of research covered within the programme are adult stem cell biology, mechanisms of cell fate instruction and cancer. The programme currently consists of four groups:

> Luciano Dicroce: Epigenetic mechanisms during PML-RAR induced leukemia
> Thomas Graf: Hematopoietic stem cells, differentiation and reprogramming
> Salvador Aznar-Benitah: Epithelial stem cells in the skin and cancer
> Bill Keyes: Skin stem cells, cancer and senescence

In the spring of 2009 a new Group Leader, Bill Keyes, joined the programme, moving from Cold Spring Harbor Laboratory. He has increased the programme's critical mass in adult stem cell research and also shares with the other Group Leaders various technologies, including mouse genetics, epigenetic analyses, fluorescence microscopy, histopathology and flow cytometry.

In the course of the year we were also able to attract a new Senior Group Leader, Maria Pia Cosma. She will move to the CRG in April 2010 from Naples, where she currently works at the Telethon Institute. Pia has a broad background in transcription factor mechanisms and synthetic biology in yeast and has also worked on the role of sulfatase modifying factors in hematopoietic stem cells. In the last few years she has been working on the reprogramming of somatic cells into induced pluripotent (iPS) cells, using both enforced transcription factor expression and cell fusion. The mechanism of cell reprogramming by cell fusions, leading to the dominant induction of new phenotypes, will be the main focus of her research at the CRG, which is supported by an ERC grant.

The Histology Unit remains a service associated with the Programme, and is now operated by two technicians.

DIFFERENTIATION AND CANCER

Group: Hematopoietic Differentiation and Stem Cell Biology
Thomas Graf is an ICREA Professor.

Group structure:

Group Leader: Thomas Graf

Staff Scientist: Chris van Oevelen; joined in October

Postdoctoral Fellows: Maribel Parra Bola (Ramon y Cajal) ; left in March, Sabrina Desbordes (Juan de la Cierva); left in November, Thien-Phong Vu Manh (HEROIC); left in October, Christos Gekas (EMBO fellow), Eric Kallin (EMBO fellow); joined in September

PhD Students: Lars Bussmann (DAAD, CRG), Alessandro DiTullio (Leonardo, CRG), Francesca Rapino ("la Caixa"); joined in March

Technicians: Luisa Irene de Andrés (CRG), Vanessa Chiganças (Plan Nacional), Jose Francisco Infante (CRG)

SUMMARY

The laboratory's main interests are mechanisms of transcription factor-induced reprogramming of hematopoietic cells and genes that are required to establish and maintain the hematopietic stem cell phenotype.

RESEARCH PROJECTS

1. Pre-B cells do not retro-differentiate before becoming macrophages, but transiently express a few progenitor markers

During hematopoiesis, the earliest multipotent progenitors branch into a common lymphoid and a common myeloid precursor, which then give rise to differentiated cells of the adaptive and innate immune systems. However, it is largely unknown which transcription factors determine the formation of lymphoid versus myeloid cells. To answer this question, we perturb regulatory networks in committed hematopoietic cells by enforced transcription factor expression. In earlier work we found that the bZip type transcription factor C/EBP alpha, which is expressed in macrophages but not in B cells or committed B cell precursors, effectively induces a switch of B cell precursors towards functional macrophages. The activation of myeloid genes requires the collaboration between C/EBPa and the transcription factor PU.1, which is already expressed in B cell precursors. In contrast, the extinction by C/EBPa of the late B cell marker CD19 is PU.1 independent, and mediated by the downregulation of Pax5.

A question of fundamental interest about the mechanism of the induced reprogramming of differentiated cells into another cell type is whether the cells de-differentiate before they re-differentiate into a new cell type. We have addressed this question by treating pre-B cells infected with an C/EBPaER virus for 3, 12, 48 and 120 hours with the inducer and analyzing the expression of stem cell and progenitor cell surface antigens, as well as gene expression profiles.

Our FACS data show that the cells undergoing phenotypic conversion do not reactivate the stem cell/progenitor markers Slamf1, Kit, CD34, Flt3 (Flk2), Itga2b (CD41) and Il7r, but upregulate Sca1 and transiently express Itgb3 (CD61). Analysis of gene expression by Affymetrix 430.2 expression arrays and real time RT-PCR revealed similar changes and in addition a transient upregulation of Flt3, which was not translated into protein. The upregulation of Sca1 probably does not reflect an earlier phenotype, as this antigen can be expressed on myeloid cells. Then we compared the data to the expression profiles of various normal progenitors. The rationale was that if the cells retro-differentiate they should transiently express markers of earlier progenitors, such as that of lymphoid myeloid precursors, hematopoietic stem cells or even embryonic stem cells. Hierachical clustering did not group pre-B cells in the process of transdifferentiation and the various hematopoietic progenitors, but in the expected order of corresponding to the timing of induction. We also did not observe the reactivation of T cell or erythroid/megakaryocytic genes with the exception of Nfe2, an erythroid/megakaryocytic transcription factor. Overexpression of this factor in Pre-B cells activated expression of Itgb3, explaining its transient activation. We then used an unbiased bioinformatic analysis to determine the gene expression signature for the different progenitors. Our results showed that there was no significant upregulation of hematopoietic stem cells and multipotent progenitor genes, but that a subset of granulocyte/macrophage precursor specific genes was upregulated transiently 2 days after induction. From these observations we conclude that during the C/EBPa induced transdifferentiation of pre-B cells into macrophages no cells with the expression pattern of earlier cells are transiently produced. This suggests that the cell conversion largely reflects the direct transformation of the transcription factor network that specifies one immune cell type into that of another.

2. A cell line system to study the mechanism of B cell to macrophage reprogramming

A disadvantage of the system is that it relies on primary cells that have to be iolsated anew for every experiment, that B lineage cells only grow on stromal cells supplemented by cytokines, and that the proportion of infected cells is variable. To circumvent these problems, we have developed a pre-B cell line (C10) derived from the fetal Pre-B cell line HAFTL, and which contains a fusion between C/EBPalpha and the estrogen hormone binding domain. Cells treated with the inducer (either beta estradiol or tamoxifen) become reprogrammed into cells closely resembling macrophages at 100% efficiency and within about 3 days. This permits to study the process of reprogramming in homogeneous and synchronously induced populations of cells.

3. Time-lapse microscopy studies of cells undergoing reprogramming

Flow cytometry analyses of the C10 cells in the process of reprogramming suggests that the process is rapid, with about 90% of the cells expressing Mac-1 after 24 hours of induction. However, these experiments do not give any clues about the onset of differentiation at the single cell level, whether it is accompanied by apoptosis and whether the cells divide before acquiring a new fate. To address these questions, we have analyzed the process of induced transdifferentiation of C10 cells by time -lapse microscopy, in collaboration with Timo Zimmermann. We found that the first cells that change shape and become adherent, showing membrane ruffling, appear already within 10 hours. In addition, the reprogrammed cells show a substantial increase in cell motility, as can be visualized by overlaying frames of pictures taken within different time windows (Fig.1). Finally, we observed no apoptosis during the reprogramming. Current experiments address the question whether cell disvison is required for lineage reprogramming.

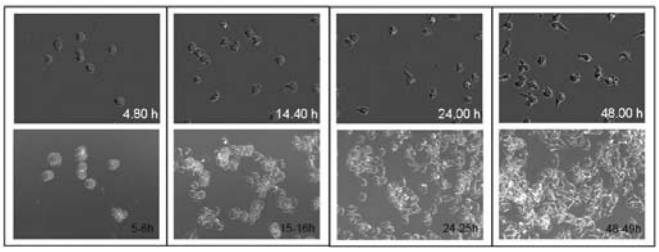


Fig.1.
Imaging of cells at different stages of induced reprogramming. C10 cells treated with b-estradiol were recorded by time-lapse microscopy at 6 min internals for 55 hours. The upper panels show images of a video of cells at the indicated time points; the lower panels represent stacks of 10 superimposed pictures corresponding to 1-hour intervals, as indicated. As can be seen, the cells change morphology and become highly motile at around 15 hours after induction.

4. A novel function of platelet integrin allb (CD41) in hematopoietic stem cell function

The maintenance and repair of the hematopoietic system requires continuous differentiation and expansion of self-renewing hematopoietic stem cells (HSCs) throughout life. In the bone marrow HSC maintenance and lineage differentiation are regulated mainly within two distinct hematopoietic niches; the osteoblast niche thought to maintain HSCs in a quiescent state, and the endothelial niche, which is involved in HSC differentiation and proliferation. Whereas a lot is known about the molecular apparatus responsible for HSC self-renewal and differentiation, key components that regulate the interaction of HSCs and downstream progenitors to the bone marrow hematopoietic niches remain poorly defined. CD41 was recently shown by us and others to be expressed also on the earliest definitive hematopoietic progenitors in the developing embryo as well as on adult HSCs. This raised the question whether the integrin CD41 is required for HSC function.

To address this question, we analyzed the hematopoietic system of the CD41 knock-out mouse model that we described earlier (Zhang et al. , Expl. Hematology, 2007). The knockout mice are viable but show a reduction in the number of cells from megakaryocyte/erythroid as well as myeloid and lymphoid lineages. However, whereas the number of multipotent and intermediate progenitors were reduced the total number of HSCs was unaltered. This suggests that the decreased bone marrow cellularity is due to an impairment of integrin-defective HSCs to access their endothelial differentiation niches. Importantly, the defects of $CD41^{-/-}$ HSCs are cell-autonomous as non-competitively transplanted recipients are capable of reconstituting the irradiated recipients and phenocopy the $CD41^{-/-}$ donor mouse defects. However, strikingly, when transplanted together with wild-type bone marrow competitor cells, $CD41^{-/-}$ HSCs exhibited a drastically reduced functionality. In addition, pre-treating wild-type HSCs with a blocking antibody against CD41 led to a similar impairment of competitive repopulation ability. The requirement of CD41 for efficient transplantation of HSCs raises the possibility that agonist activation of the integrin into a high-affinity state could lead to an increased reconstitution ability of HSCs, a possibility that we are currently testing. Taken together, we have identified integrin CD41 as a novel regulatory component for the interaction between HSCs and its niche (s) in the bone marrow.

5. A role of the RNA binding protein musashi 2 (Msi2) in hematopoietic stem cells

Another stem cell project in the laboratory was initiated by Florencio Varas. Here, we have used retroviral insertions to activate adjacent genes with a possible role in HSC expansion. In brief, bone marrow was infected with a GFP retrovirus and the cells transplanted into irradiated primary mouse recipients. After

4 months, the bone marrow of one reconstituted mouse was injected into five secondary recipients and their bone marrow analyzed several months later for the presence of common retroviral integration sites. The rationale of the experiment was that, since retroviruses integrate randomly near gene transcription start sites, any enrichment of common integration sites is a strong indication of an in vivo selection of a dominant stem cell clone due to the activation of a gene with a function in HSCs. Several retrovirus integration sites were found to be enriched, two of them corresponding to genes that are known or suspected to have a function in HSCs. The third one, however, landed close to the Musashi 2 gene, which so far has not been implicated in HSC expansion or leukemia. This gene encodes an RNA binding protein whose homologue in Drosophila shows a germ stem cell defect when inactivated.

If indeed Msi2 is important for stem cell function it would predict that its loss would induce an HSC defect. To study this, we collaborated with Thomas Floss from the GSH in Muenchen, who provided us with Msi2 knockout mice generated after an enhancer trap screen with ES cells. The mutant mice are viable and have a functioning hematopoietic system. However, they show a decrease in the numbers of multipotent progenitors in the bone marrow. More importantly, when bone marrow from these mice, which contain HSCs, was transplanted competitively with bone marrow from wild type animals, they were dramatically outcompeted. This indicates that the Msi2 genes has an important function in the formation of multipotent progenitors and probably also in the maintenance of functional HSCs. Current studies are directed towards understanding the mechanism of action of this gene in hematopoietic stem and progenitor cells.

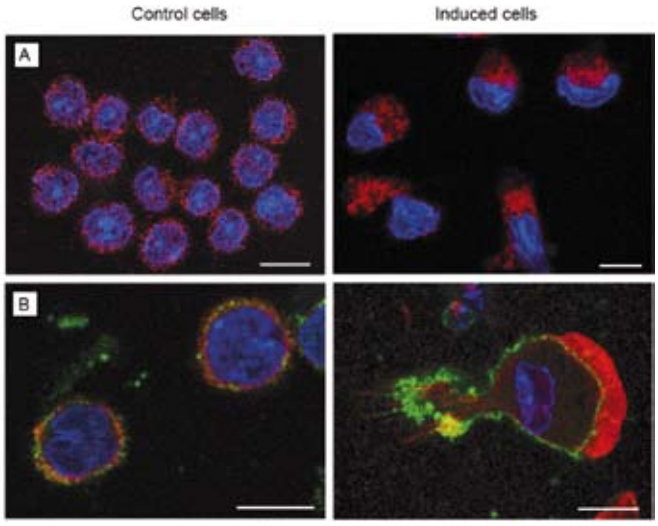


Fig.2.
Changes in cell morphology during C/EBPa induced transdifferentiation of C10 A, staining of Golgi apparatus with antibodies against mannosidase II (red) and nuclei with DAPI (blue). B, staining of cells with antibodies to cell surface antigens (green), with phalloidin-Alexa Fluor 568 for F actin (red) and with DAPI for nuclei (blue). The green fluorescence in the control cells represents CD19 antigen; the green fluorescence in the induced cell represents Mac-1 antigen. Left panels: uninduced cells; right panels cells induced for 3 days. Bars, 10 um.

PUBLICATIONS

Varas F, Stadtfeld M, De Andres-Aguayo L, Maherali N, di Tullio A, Pantano L, Notredame C, Hochedlinger K, Graf T.
"Fibroblast derived induced pluripotent stem cells show no common retroviral vector insertions."
Stem Cells, 27, 300-306 (2009).

Bussmann L, Schubert A, Vu Manh TP, De Andres L, Desbordes S, Parra M, Zimmermann T, Rapino F, Rodriguez-Ubreva J, Ballestar E and Graf T.
"A robust and highly efficient immune cell reprogramming system."
Cell Stem Cell, 5, 554- 566 (2009).

Graf T.
"An uphill battle towards pluripotency."
Nature Genetics, 41, 960-961 (2009).

Graf T and Enver T.
"Forcing cells to change lineages."
Nature, 462, 587- 594 (2009).

DIFFERENTIATION AND CANCER

Group: Epigenetics Events in Cancer

Luciano Di Croce has an ICREA Group Leader position.

Group structure:

Group Leader: Luciano Di Croce

Postdoctoral: Marcus Buschbeck (until October 2009), Luciana Rocha Viegas (until June 2009), Holger Richly, Celia Jeronimo, Elisabeth Simboeck, Martin Lange

PhD Students: Santiago Demajo, Sophia Teichmann, Iris Uribesalgo Micás, Joana Ribeiro, Paola Pisano

Technician: Arantxa Gutierrez

SUMMARY

Understanding the genetic basis of cancers has been a topic of intense research, and hundreds of gene mutations have been identified that can cause carcinogenesis. However, in the past few years, increasing evidence has suggested that mutations are not the only genetic changes that lead to cancer. Indeed, perturbations of chromatin structure and of other epigenetic mechanisms can cause inappropriate gene expression and genomic instability, resulting in cellular transformation and malignant outgrowth.

The focus of our laboratory is to understand the basic mechanism of gene regulation and the impact of epigenetic marks on chromatin metabolism, using normal cells, cancer cells, and mES cells as model systems. We will also address some of these questions using mouse models.

RESEARCH PROJECTS

1. Role of histone demethylases in leukemia

PcG and TrxG proteins were initially described in ***Drosophila*** as repressors and activators of Hox genes, respectively. More generally speaking, PcG and TrxG proteins play an important role in regulating lineage choices during development and differentiation. Additionally they are implicated in cell proliferation, stem cell identity and cancer, cellular senescence, genomic imprinting, X-inactivation and hematopoiesis.

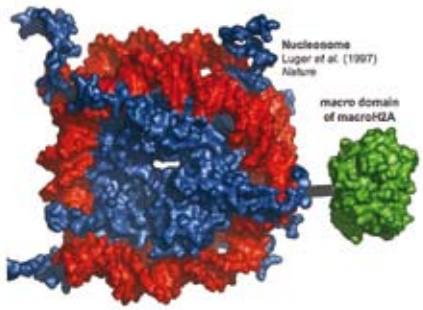
In collaboration with Dr. R. Shiekhattar (Wistar, USA), we identified UTX as the enzyme responsible for H3K27 demethylation (Lee et al, 2007). More recently, UTX have been found mutated in several tumour (van Haaften et al, 2009), thus corroborating the important role of epigenetic deregulation in human cancers. Interestingly, UTX is a component of the MLL complex, the mammalian orthologous of Trithorax. Our previous date indicates that after retinoic acid administration, Utx specifically demethylases H3K27 at several Hox genes.

We will investigate:
(i) how Utx is recruited to promoters,
(ii) which are the target genes in the human genome,
(iii) which role has Utx in tumorigenesis.

The oncoprotein PML-RARα is one of the most well-studied leukemogenic transcription factors. The PML-RARα fusion protein, responsible for 99% of acute promyelocitic leukemia (APL) cases, arises from a t(15;17) balanced reciprocal chromosomal translocation. It involves the PML gene and the retinoic acid receptor α (RARα) gene. PML-RARα represses target genes through recruitment of DNMTs and Polycomb complex (Villa et al, 2007). It is likely that Utx is important for the re-establishment of the proper gene transcription program necessary for cell differentiation. Thus the role of Utx in leukemia is also being investigated.

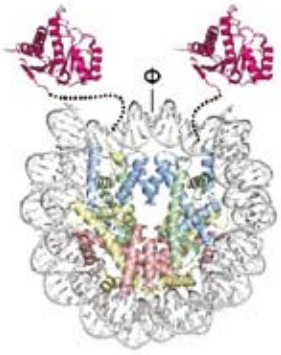
2. macroH2A and gene silencing

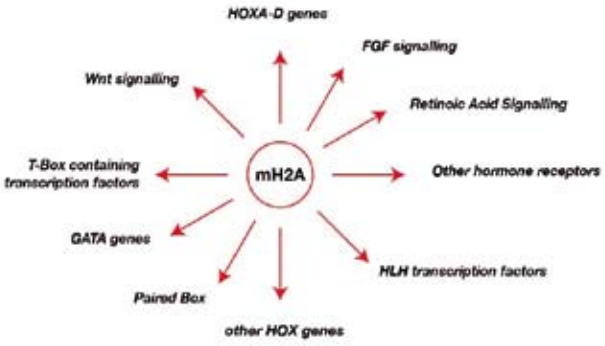
The most extensive histone modification is the complete exchange of canonical histones for variant ones. Among all known histone variants, the so-called macroH2A is the one that is the most divergent from its canonical histone and is the least understood in its function. In addition to a homologous histone domain, macroH2A possesses a large C-terminal domain of unknown function – the macro domain.
The H2A domain of macroH2A histones is ~65% identical to that of the conventional H2A.



The early observation that macroH2A is enriched on the inactive X chromosome and centrosomes suggested an involvement in gene repression and heterochromatinization. Biochemical studies using synthetic templates have indicated that macroH2A-containing nucleosomes are structurally different in the vicinity of the dyad axis, and this correlates with the inability of transcription factors to bind to DNA sites inserted nearby. These observations suggest that incorporation of macroH2A into nucleosomes could confer an epigenetic mark for gene repression. However, there are no known gene targets for macroH2A-dependent transcriptional repression, and no evidence for the mechanisms by which macroH2A could be recruited to specific genes and repress transcription in vivo.

Our microarray-based analysis in human male pluripotent cells uncovered occupancy of macroH2A at a large number of genes coding for key regulators of development and cell fate decisions. On these genes, presence of macroH2A is a repressive mark that locally and functionally overlaps with Polycomb repressive complex 2. We demonstrate that macroH2A contribute to the fine-tuning of temporal activation of HOXA cluster genes during neuronal differentiation. Furthermore, elimination of macroH2A function in zebrafish embryos produced severe specific phenotypes. Taken together, our data demonstrate that macroH2A constitutes an important epigenetic mark involved in the concerted regulation of gene expression programs during cellular differentiation as well as vertebrate development.



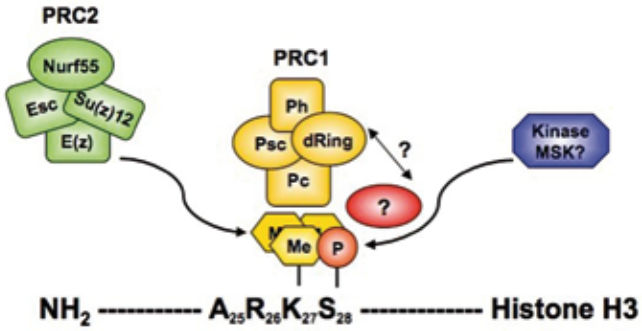


Some key questions remain still unsolved: which are the molecular mechanism involved in macroH2A deposition? Does macroH2A occupancy also control regulative regions in the genome other than gene promoters? A combination is biochemistry and genome-wide analysis are necessary to answer to those questions.

3. Cross-talk of epigenetic marks in gene regulation

Binding of chromatin-associated proteins can be influenced by adjacent additional modification. For instance, binding of HP1 to H3K9me is inhibited by an adjacent phosphorylation on serine 10 (H3S10ph). H3S10ph is associated with actively transcribed chromatin and is recognized by 14-3-3 proteins, whose binding is even favoured by additional lysine 14 acetylation (H3K14ac).

Therefore, we wonder whether additional posttranslational modifications in the proximity of H3K27me, in particular phosphorylation of Serine 28 (H3S28ph), would have any effect on Pc binding and consequently on target gene expression.



Similar to H3S10ph, H3S28ph is an abundant and transient mark during mitosis, while during interphase is also associated with actively transcribed chromatin. However it was shown that H3S28ph targets a different chromatin population than H3S10ph. Proteins, which may specifically bind to H3S28ph, are still unknown.

We have identified several H3S28ph "binder". We are now characterizing their function with respect to Polycomb binding, cell faith decision, and their role in promoter regulation and chromatin structure.

PUBLICATIONS

Buschbeck M and Di Croce L.
"Approaching the molecular and physiological function of macroH2A variants."
Epigenetics, in press.

Cardamonea MD, Bardella C, Gutierrez A, Di Croce L, Rosenfeld MG, Di Renzo MF, and De Bortoli M.
"ERa as ligand-independent activator of CDH-1 regulates determination and maintenance of epithelial morphology in breast cancer cells."
PNAS 106, 7420-25 (2009).

Buschbeck M, Uribesalgo I, Wibowo I, Rué P, Martin D, Gutierrez A, Morey L, Guigó R, López-Schier H, and Di Croce L.
"The histone variant macroH2A is an epigenetic regulator of key developmental genes."
Nature Struct Mol Biol 16,1074-79 (2009).

PATENT

Priority Date: 3rd July 2009
Application Nr.: ES P200901534
Application Date: 3rd July 2009
Office where the patent has been submitted: OEPM
Title: *"Complejos proteicos e inhibidores de los mismos para el tratamiento del cáncer."*
Applicants: Centre de Regulació Genòmica (CRG), ICREA
Inventors: Luciano Di Croce (CRG), Holger Richly (CRG)

DIFFERENTIATION AND CANCER

Group: Epithelial Homeostasis and Cancer

Salvador Aznar Benitah has a Junior ICREA Group Leader position.

Group structure:

Group Leader: Salvador Aznar Benitah

Postdoctoral: Cristina Hidalgo

PhD Students: Peggy Janich, Nuno Luis, Stefania Mejetta

Technicians: Bernd Kuebler, Gloria Pascual

Masters Student: Lara de Llobet

SUMMARY

The epidermis, and other stratified epithelia, needs to renew constantly in adults to maintain its function. This process is called homeostasis and relies on a population of epidermal stem cells (epSCs) that self renew and can undergo terminal differentiation. EpSCs adhere strongly to specialized niches where they remain relatively quiescent and unspecified. Upon a requirement of tissue replenishment, they become active, proliferate, and egress the niche to contribute to the differentiated compartment. The process is assymmetric, ensuring that the percentage of stem cells is maintained more or less constant after each cycle of activation. The transition between each state (quiescence vs proliferation; adherence vs egression; unspecified vs differentiated) is tightly regulated by the microenvironment and the intrinsic genetic program of the epSCs. Failures in this strict regulation can lead to premature ageing or to the development of tumours. The aim of our work is to understand the molecular mechanisms that control the behaviour of normal adult stem cells during tissue homeostasis and how their deregulation contributes to carcinogenesis.

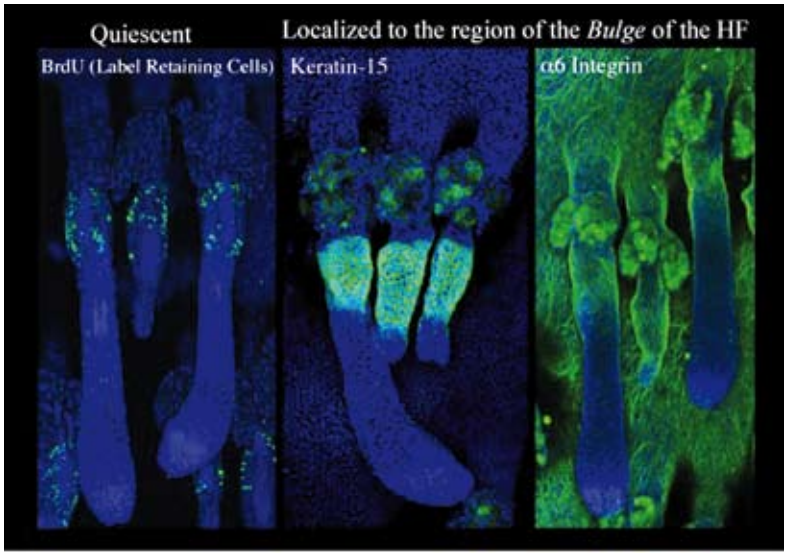


Figure 1
Epidermal stem cells are located at a region of the hair follicle known as the bulge. They are relatively quiescent and strictly positioned at their niche (Immunostaining and pictures by Peggy Janich).

RESEARCH PROJECTS

Little is still known about the spatiotemporal distribution, and the hierarchy, of the molecular pathways relevant to the transition between the inactive and active states of epidermal stem cells within the stem cell niche. Intriguingly, inactive epSCs express high levels of ¨molecular breaks¨, which make them refractory to activating stimuli. Why then do they respond to such stimuli? In addition, upon stimulation, a very small proportion of epSCs become active, whereas the bulk remains unresponsive. Why don´t all epSCs respond? What is the nature of this stem cell heterogeneity? The consequences of unbalancing this equilibrium must be underscored, since tilting it towards excessive or reduced activation may predispose the tissue to premature aging due to excessive stem cell depletion, carcinogenesis when combined with accumulation of DNA damage, or lack of regenerative potential due to the inability of the stem cells to become active upon tissue damage.

We can summarize our interests in three questions:

a) What is the molecular nature of the heterogeneity of epSCs within their niche? A small percentage of epSCs respond to activating stimuli; why do not all cells respond? How are these restricting mechanisms lost during carcinoma formation?

b) Are activating stimuli instructive or permissive for epSC activation? Inactive epSCs highly express inhibitors of activating stimuli: why do they respond to activating stimuli then? Is response predisposed by an intrinsic genetic program of the epSC?

c) What distinguishes the different choice of the two epSCs daughter cells? i.e. remain at or exit the stem cell niche. How and why is this mechanism lost in carcinomas?

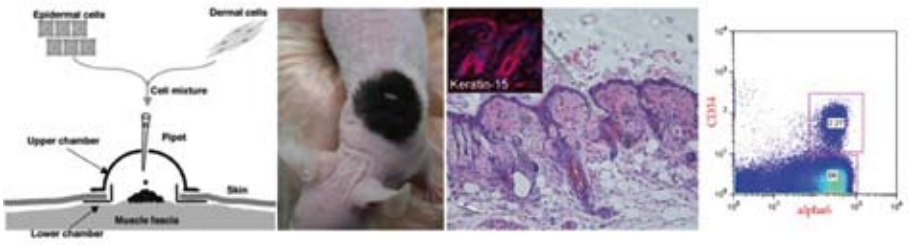


Figure 2

Epidermal stem cells are the only population that can regenerate the entire epidermal compartment. Epidermal stem cells transplanted onto nude mice can regenerate a homeostatic functional epidermis, hair follicles and sebaceous glands. A niche of epidermal stem cells is re-established once transplanted, which exhibits expression of stem cell markers such as keratin-15, integrin alpha6, and CD34 (Experiment done by Gloria Pascual).

1. Molecular and genetic mechanisms involved in epidermal self-renewal and differentiation.

We have previously identified a signalling axis important for establishing equilibrium between inactive and active epSCs. The small GTPase Rac1 promotes EpSCs quiescence and strong adhesion to the stem cell niche (Watt 2008, Benitah 2005). Conditional epidermal deletion of Rac1 causes an initial burst of proliferation and loss of the quiescent epSC pool, coupled to a massive exit of activated epSCs from their niche. Continuous epSC activation, upon Rac1 deletion, significantly depletes their number, leading to loss of epidermal maintenance and integrity.

Mechanistically, Rac1 modifies the activity of the transcription factor Myc to control the function of epSCs (Benitah 2005, Watt 2008). Myc is a general regulator of cell proliferation, metabolism and apoptosis, and together with p53, is one of the most common molecular lesions in human cancers (Watt 2008). Low, steady-state levels of Myc drive proliferation and repress the expression of matrix-adhering proteins, thereby allowing epSC niche exit (Watt 2008, Gebhardt 2006). However, increased Myc activity results in hyperproliferation, genome instability, and a predisposition to develop squamous carcinomas (Watt 2008). In inactvite epidermal stem cells, Rac1 stimulates the ser/thr kinase activity of PAK2, which then phosphorylates Myc at its C-terminus (Watt 2008, Benitah 2005). Phosphorylated Myc can no longer bind to specific DNA elements within the promoters of its target genes, nor interact with its transcriptional partner Max (Watt 2008, Benitah 2005). Consequently, EpSC proliferation and exit from the niche are impaired, preventing their activation.

Modulation of this pathway has enabled us to induce two states: inactivation (Rac1high/phospho-Myc) and activation (Rac1low/unphospho-Myc) of EpSCs. Based on this, the global comparative transcriptome (Affymetrix) of human EpSCs in their active versus inactive state has been obtained. In addition, we have undertaken a large scale proteomic approach to identify proteins that interact with members of the Rac1/Myc signalling cascade in epidermal stem cells. We are combining these two approaches, microarray data and proteomics, to identify new players relevant for different aspects of epidermal stem cell behaviour.

Selected targets are being validated using cellular and molecular biology tools with primary cultures, as well as with our *in vivo* mouse models.

2. Studying novel pathways relevant to epidermal and squamous tumour onset and progression.

Adult stem cells are potentially the few long term tissue residents that in time may accumulate enough somatic oncogenic mutations which result in the development of neoplasias. Moreover, the behaviour and molecular signature of a small percentage of cancer cells, known as cancer stem cells, recapitulate those of adult stem cells in the normal tissue. Cancer stem cell self-renewal, high potential of invasion and homing into a specific niche, with direct consequences over tumour maintenance and metastasis, are most probably characteristics inherited from normal adult stem cells. However, very little is known about the signalling events and the molecular signature that contribute to the behaviour of cancer stem cells in tumours of epithelial origin.

In collaboration with the Hospital del Mar we are obtaining samples of SCCs (fresh live tissue, and blocks for immunohistological analysis) at different stages of tumour progression. Squamous cell carcinomas are the most diagnosed types of tumours in western countries with poor prognosis when developed in the oral cavitiy. We are analyzing the status of the different pathways studied in the lab with respect to their possible role in squamous cancer stem cells and validating the results using our mouse models.

DIFFERENTIATION AND CANCER

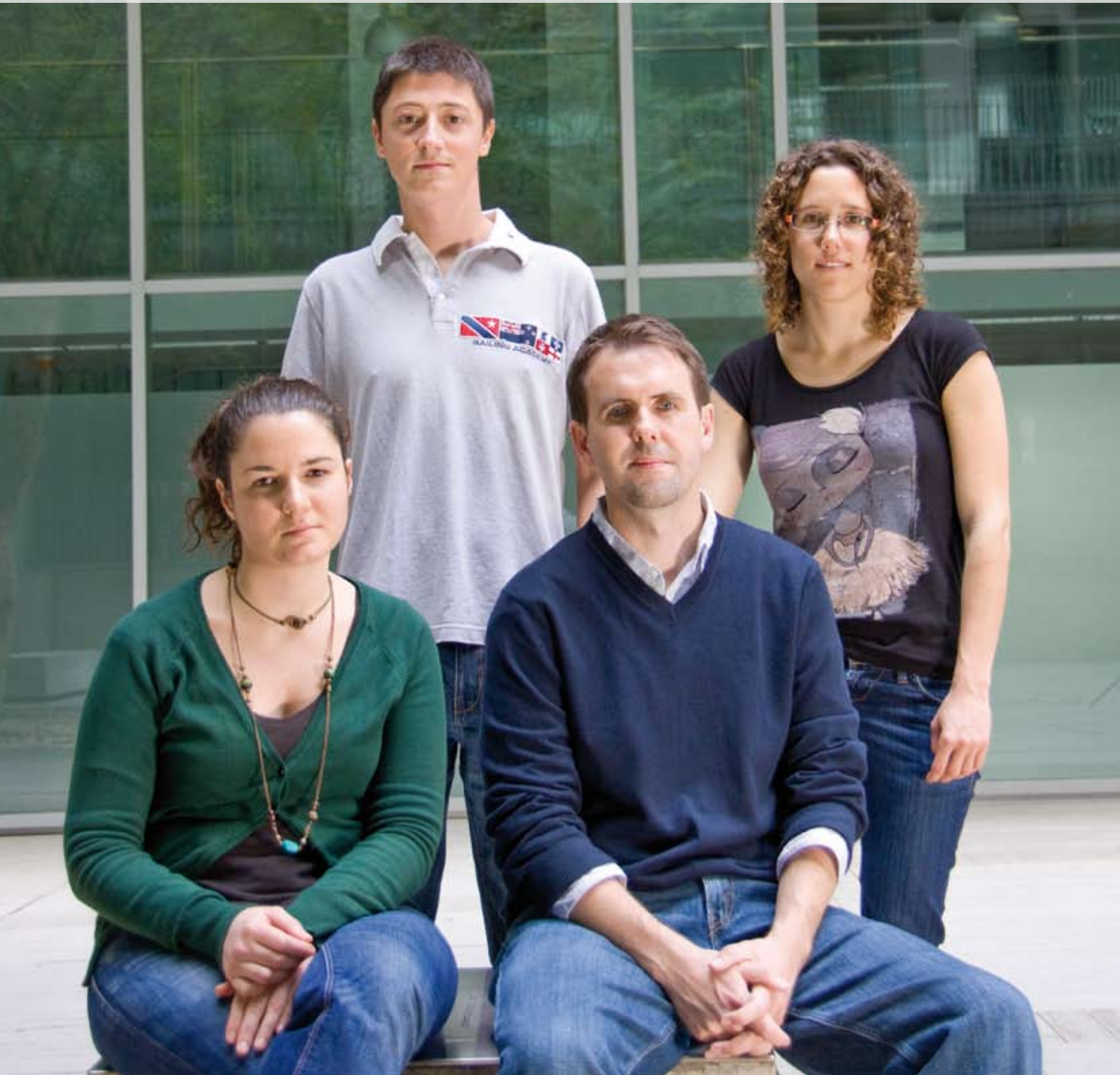
Group: Mechanisms of Cancer and Aging

Group structure:

Group Leader: Bill Keyes

PhD Students: Valeria Di Giacomo, Matteo Pecoraro

Technicians: Alba Mas



SUMMARY

Cancer and aging are two biological process that may seem unrelated. However, recent evidence demonstrates that they share many similarities, with the process of cellular senescence acting as a direct link between the two. Cellular senescence is an irreversible cell cycle arrest, that is induced in response to a variety of cellular stresses. The best characterised of these stimuli is the activation of oncogenic signaling. In response to such signals, cells enter into a senescent state as a primary defense mechanism to curtail aberrant proliferation. As such, senescence is a potent tumor suppressive mechanism that prevents the early stages of tumor growth. However, senescence is also induced when cells reach the end of their lifespan and are unable to proliferate further, a process induced by telomere shortening and DNA-damage. Increasing evidence suggests that such replicative senescence is a major cause of in vivo aging in animals and humans. At the molecular level, the main mediators of the senescence pathway include tumor suppressors and oncogenes such as p53, p16, Ras and Myc, all of which are increasingly being linked to the aging process. As such, any factor that regulates senescence can have a major impact on both the process of cancer and aging.

We have previously identified the p53-family member p63 as one such critical mediator of senescence, having direct effects on both cancer and aging. Using mouse models and primary cell cultures, we found that a deficiency of p63 induces a program of premature senescence and accelerated aging in vivo, providing further evidence linking these two processes. However, we have recently found that p63 is also a critical mediator of oncogene-induced senescence, having striking effects on tumor development. Together, these studies highlight the importance of p63 in regulating the balance between cancer and aging, most likely through the regulation of cellular senescence.

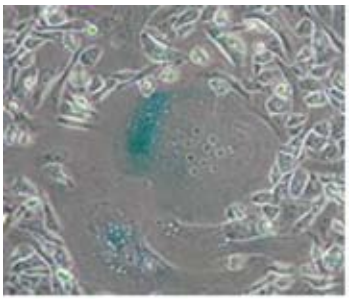


Figure 1
Senescence-associate β-galactosidase staining identifies senescent cells (blue stain). Note the characteristic morphology, including enlarged flat shape and multiple nuclei. Multiple smaller transformed cells surround.

RESEARCH PROJECTS

1. Investigating the role of p63 in the pathogenesis of Squamous Cell Carcinoma

Squamous cell carcinoma (SCC) is one of the most frequent solid tumors worldwide, representing the second highest cause of skin cancer, one third of lung cancers, and in the case of head and neck SCC, the sixth most common solid tumor type. Recently we identified ΔNp63α as an oncogene that is capable of inducing the development of SCC. Although this isoform of p63 is frequently overexpressed in human SCC, a causative role in tumor development had not been shown. In studies using primary mouse keratinocytes and nude mouse models, we found that ΔNp63α promotes SCC by inhibiting the process of Ras-induced senescence. Surprisingly however, we found that the tumorigenic cells that could bypass senescence and form SCC in response to overexpression of ΔNp63α possess many features of stem-like cells, including an ability to form self-renewing spheres in 3D-tissue culture, a capacity to differentiate and a resistance to DNA-damaging drugs.

Within many human tumors, including SCC, populations of cells exhibiting properties of stem cells have been identified. These cancer stem cells posess intrinsic growth properties that favor tumor development, resistance to treatment and tumor recurrence after treatment. Such properties include a capacity for self-renewal, an ability to differentiate and an inherent resistance to DNA-damage. However, the origin of these tumorigenic therapy-resistant cells and their mechanisms of proliferation are unknown in many tumor types. Understanding the processes that favour the proliferation of these cells is necessary to design more effective therapies for many cancers. In this project we propose to investigate the processes by which ΔNp63α is oncogenic, focusing on identifying the mechanisms involved and the processes by which stem-like properties may favour tumor development.

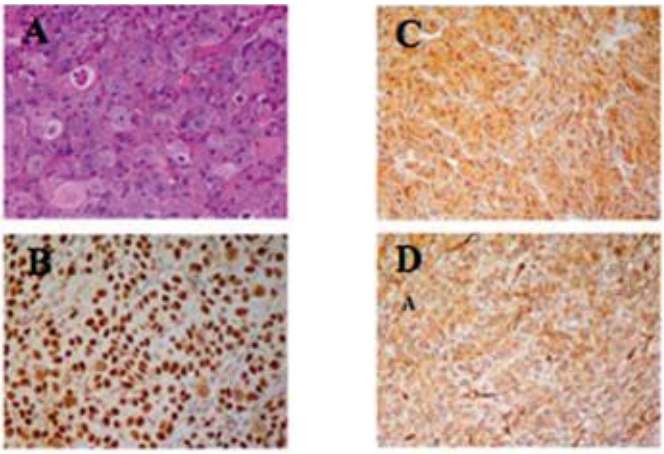


Figure 2
Squamous Cell Carcinoma induced by ΔNp63α overexpression with the oncogene Ras, stained for H&E (a), p63 (b), Keratin 14 (c) and β-catenin (d).

2. Determining the function of p63 in prostate stem cells and prostate tumor development

In prostate cancer, the role of p63 is unknown and controversial. Unlike the overexpression of p63 that is seen in SCC, during the development of adenocarcinoma, the most common prostate tumor type, p63 expression is actually lost from the cells that are undergoing malignant conversion. Indeed it has been suggested that it is the p63-positive stem cells that undergo malignant transformation during prostate tumor initiation. However, is not known if this loss of p63 facilitates or is required for prostate tumor development, or whether there is a shift in the ratio of expression of p63 isoforms during transformation. To further complicate the situation, loss of expression of p63 has been correlated with the upregulation of genes that are involved in epithelial-to-mesenchymal transformation and metastasis. In prostate tumors, the development of recurrent treatment-resistant tumors that undergo metastasis is the main cause of death from these tumors. The goal of this project is to investigate the function of p63 in the prostate gland, with particular emphasis on its roles in prostate stem cells and prostate tumor development. By taking a multidisciplinary approach, using functional genetics, ***in vivo*** animal models and high-throughput genomic screens, it is hoped that this work will identify key genes and mechanisms during prostate tumor development that can be targeted for therapy in future studies.

3. Novel pathways linking cancer and aging

In our ongoing investigations into the mechanism of action of p63 during tumor development, we have identified the Lsh/Hells gene as a novel p63-target that is required for senescence bypass and tumor initiation. Lsh is a member of the SNF2-family of chromatin remodelers that is involved in promoting DNA methylation and transcriptional silencing, through recruitment of DNA-methyltransferases or direct interaction with members of the Polycomb-repressive complex-1 (PRC1). Interestingly, mouse models deficient for Lsh exhibit premature aging and enhanced cellular senescence. In our studies, we find that Lsh is overexpressed in p63-expressing tumors, while knockdown of Lsh prevents senescence bypass and aberrant tumor-initiating cellular proliferation. We are currently investigating the function of Lsh in stem-cell homeostasis, tumor initiation and aging.

PUBLICATIONS

Guo X, Keyes WM, Papazoglu C, Zuber J, Li W, Lowe SW, Vogel H and Mills AA.
"TAp63 induces senescence and suppresses tumorigenesis in vivo."
Nat Cell Biol 11(12):1451-7 (2009).

GENES AND DISEASE

Coordinator: Xavier Estivill

The mission of the Genes and Disease Programme is to perform research on the molecular basis of human disease, from the discovery of genes involved in common human disorders to the development of preventive and therapeutic strategies. The Programme combines large-scale experimental approaches with advanced genetic strategies to elucidate biological determinants of common human disease.

Efforts of the Genes and Disease Programme researchers are focused on analysing sequence and genomic variants of the human genome that could participate in the predisposition to disease. Specific collective work within the Programme is focused on understanding the function of genes with potential implications in mental retardation and neurodegeneration, by using cellular and animal models. Programme researchers use murine models to develop therapeutic approaches that could correct features involved in mental retardation, anxiety disorders and cancer.

Translational research is a primary concern of many Genes and Disease members. To this aim, the Programme Group Leaders are part of the Rare Disorders and Epidemiology and Public Health Biomedical Research Networks, of the Spanish Medical Research Council, and other biomedical research networks at the regional, national or international levels. The programme isorganised in four groups and will expand with new groups in the field of medical systems biology in the next years. The genotyping core facility (CeGen) is associated to programme and has fully been supported by external funds (Genoma España).

The programme currently consists of four research groups:

> Gene Function and Murine Models of Disease (Mariona Arbonés/Susana de la Luna)
> Neurobehavioral Analysis (Mara Dierssen)
> Genetic Causes of Disease (Xavier Estivill)
> Gene Therapy (Cristina Fillat)

GENES AND DISEASE

Group: Genetic Causes of Disease

Group Structure:

Group Leader: Xavier Estivill

Staff Scientist: Eulàlia Martí

Scientific Officer: Veronique Blanc

Postdoctoral Fellows: Mariona Bustamante, Mónica Bañez-Coronel, Mario Cáceres (Ramón y Cajal), Mònica Gratacòs (CIBERESP), Raquel Rabionet (Ramón y Cajal), Eva Riveira (CIBERESP)

PhD Students: Johanna Aigner, Laia Bassaganyas, Elisa Docampo, Susana Iraola, Alexander Martínez , Elisabet Mateu, Elena Miñones, Margarita Muiños, Lorena Pantano, Raquel Rubio, Ester Saus Daniel Trujillano (since September 2009)

Technicians: Anna Carreras, Georgia Escaramís (CIBERESP), Birgit Kagerbauer, Marta Morell, Silvia Porta (CIBERESP), Sergi Villatoro, Anna Brunet (until September 2009)

SUMMARY

We are interested on how the variability of the human genome contributes to phenotype and common human disease. The availability of sequence information of the genomes of several individuals has provided the view that genetic variability is much larger than initial estimations. Genomic variability has different forms, single nucleotide polymorphisms (SNPs), structural variations (mainly copy number variants or CNVs, and insertion/deletion (indel) variants. Our work aims to understand how these variants influence common human disorders. We are also interested in the study of epigenetic modifications and small non-coding RNA pathways, as susceptibility factors in complex diseases (neuropsychiatric and neurodegenerative disorders). We are approaching these questions through genomic platforms that explore SNPs and CNVs and use deep sequencing for the characterisation and analysis of small non-coding RNAs. Functional studies are being used to address the consequences of the genetic and epigenetic changes in cell physiology and pathology. The study of disorders in which the interaction between genes and environment seems to play an important role in the development of the disease status is also a major focus of our research.

RESEARCH PROJECTS

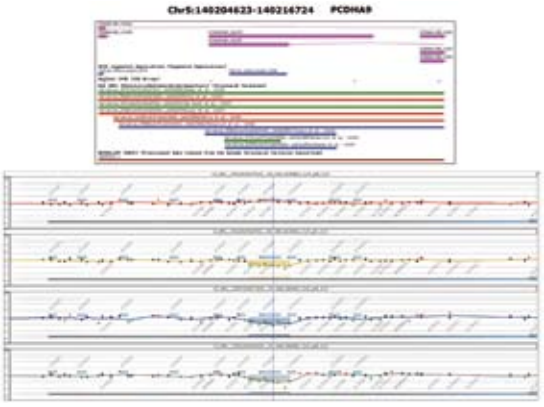
1. Structural variations and human disease

Large-scale segmental duplications (SDs) have played an important role in hominoid evolution and can be hotspots for non-allelic homologous recombination leading to deletion, duplication, inversion or translocation. Many of these SDs coincide with structural variations or copy number variants (CNVs) of the human genome. We are studying the genomic organisation of regions with structural variation in samples from different ethnic groups. We are also evaluating the nature of sequences contained in SDs and CNVs, including repetitive elements, transposable sequences and non-coding RNAs. Ongoing efforts will involve the evaluation of piRNAs in structural variations and their role in protecting the DNA sequence during important phases of the cell life.

We are systematically performing comparative genomic hybridization (CGH) experiments and we are using multiple technologies to further analyse SDs and CNVs at the genome scale for several common disorders. CNVs with specific profiles for several common disorders are being investigated. Genes located within sequences containing structural variants are object of major attention by the group. Disorders that are studied include schizophrenia, stroke, psoriasis, multiple sclerosis, asthma Parkinson's disease and fibromyalgia. The goal is to identify genomic variants that have functional consequences for these disorders.

We have made specific progress in psoriasis, a chronic disorder of the skin affecting most ethnic groups, but with the highest prevalence (3%) in northern Europeans. We have identified a common CNV that involves the deletion of two genes (LCE3B and LCE3C) in a significant fraction of patients with psoriasis. The lack of these two genes, highly expressed in the injured epidemis, brings skin barrier alteration as a new player in psoriasis susceptibility. This association has been replicated in several populations and constitutes one of the main loci associated with psoriasis.

The group is participating in the International Cancer Genomics Consortium (ICGC) with the aim to fully characterize the genome of chronic lymphocytic leukemia (CLL). The role of our group is to characterize the spectrum of non-coding RNAs and structural variants in normal and leukemic cells from patients with CLL. The Spanish contribution to the ICGC is a collaborative effort between several centers in Spain (Clinic Hospital, University of Oviedo and CNIO, among others).

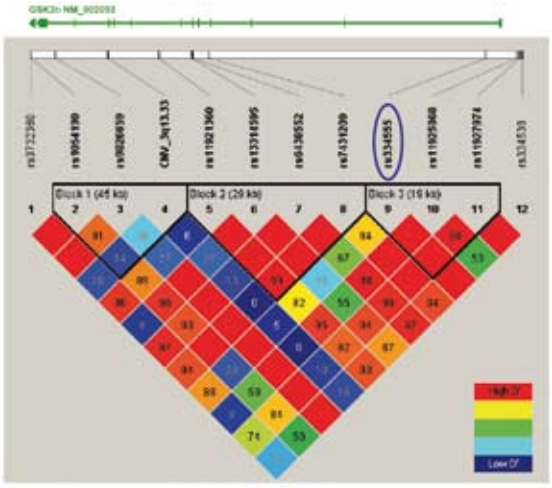


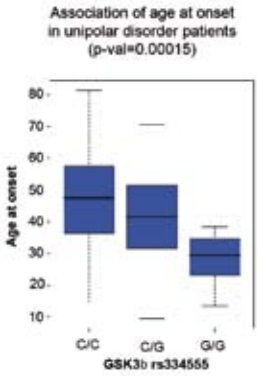
2. Genetic variants associated to psychiatric disorders

We are evaluating several psychiatric disorders (substance-abuse, anxiety, eating, psychotic and mood diseases) for potential functional SNPs in genes involved in neurotransmission and/or neurodevelopment. We found that multiple SNPs are strongly associated with several disorders. Specific progress has been achieved for bipolar disorder, obsessive-compulsive disorder and major depression. We have found that CRY1 and NPAS2 are associated with unipolar major depression and CLOCK and VIP with bipolar disorder. We have also identified that a brain-derived neurotrophic factor haplotype is associated with therapeutic response in obsessive-compulsive disorder. We have systematically characterized miRNA polymorphism in several psychiatric phenotypes and carried out functional assays. This has lead to detect allele variants in functional MicroRNA target sites of the neurotrophin-3 receptor gene (NTRK3) as susceptibility factors for anxiety disorders. Other studies for other psychiatric phenotypes are ongoing.

We are exploring the potential role of variants in miRNA targets in candidate genes for neuropsychiatric disorders. Reporter-gene assays based in dual luciferase and miRNA mimics transfection in HeLa cells have shown 11 miRNAs as post-transcriptional regulators of different NTRK3 isoforms and 7 miRNAs as regulators of BDNF. We have also studied the putative targets of two clustered miRNAs found to be downregulated in Parkinson disease samples. To this aim, a dopaminergic neuronal cell line was transfected with the miRNA mimics or inhibitors and the mRNA expression profiles were subsequently studied using Illunmina arrays. The mRNAs differently expressed in the cells presenting the miRNA downregulated should provide valuable information about the putative targets of this miRNA.

We have studied the genetic variability of miRNAs and their possible implication in susceptibility to disease. Case-control studies have been performed in more than 1300 patients with different psychiatric disorders and in control subjects. We have detected several associations for anxiety disorders, eating disorders and schizophrenia. Functional studies based in luciferase gene-reporter assays have already been performed and show that four of these miRNAs regulate candidate genes for panic disorder, such as HTR1A and HTR2C. Transcriptome analysis has revealed several neurological related pathways to be altered when two of these miRNA are overexpressed in a neuroblastoma cell line. One of these miRNAs is a brain-enriched miRNA that has been shown to regulate three candidate genes for schizophrenia. For obsessive-compulsive disorder and affective disorders several associations have also been found that are now under investigation. Replication studies in several cohorts of patients with psychiatric disorders are ongoing.



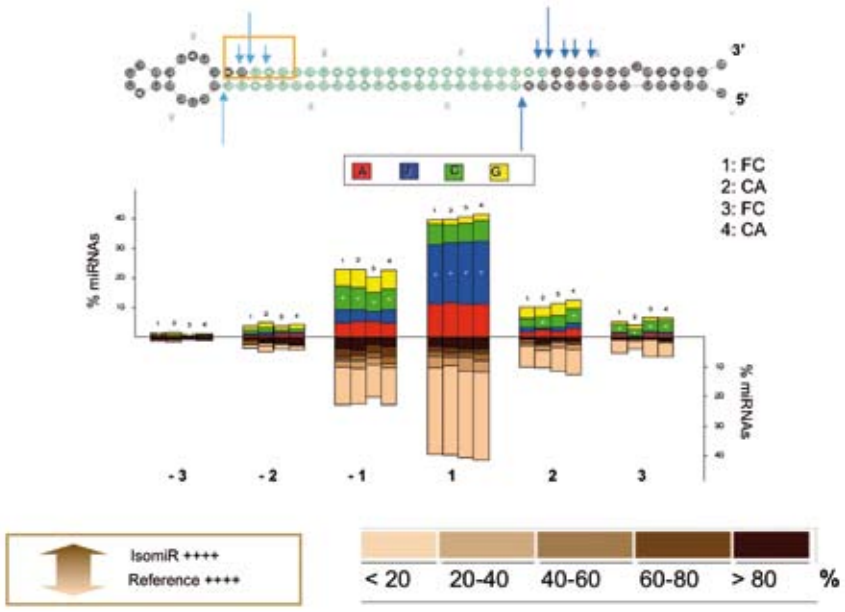


3. Functional genomics of neurological disorders

Small non-coding RNAs are involved in the guidance of diverse types of gene regulation, typically resulting in reduced expression of target genes. MicroRNA (miRNA) are a major class of non-coding RNAs. They are ubiquitously expressed and are believed to regulate most biological processes in a tissue- and temporal-specific manner. They are key players in the development of the central nervous system. The characterization of small RNA expression profiles in neurodegenerative disorders should provide important clues about the fine regulation of a plethora of genes that are involved in cell-specific survival of key neurones of brain regions that lead to disease.

We have used Illumina/Solexa deep sequencing to analyse and characterize small non-coding RNAs in different affected and non-affected brain areas of individuals with neurodegenerative disorders. miRNA variability has recently been described using several large scale sequencing strategies. These multiple mature miRNA variants have been referred as isomiRs. It has been proposed that much of the IsomiR variability can be explained by variability in either Dicer1 or Drosha cleavage positions at the 5'-end or 3'-end of the miRNA-precursor, during miRNA biogenesis (5'- or 3'-trimming isomiRs). The 5'-trimming variants are of particular interest since they affect the seed region of the miRNA (positions 2-8 of the miRNA), which is key for the recognition and silencing of mRNA targets. We have generated a web-based bioinformatic tool to analyze miRNA variability at different levels including length and frequency distributions and differential expression between samples. With this tool we have identified miRNA and IsomiRs with putative functions in physiological and pathological processes.

We are also studying the role of small non-coding RNAs in neurodegeneration associated to triplet repeat expansion diseases such as Huntington disease (HD) and fragile X-associated tremor ataxia syndrome (FXTAS). In triplet repeat expansion diseases, it has been recently shown that short CAG repeats generated from mutated long repeat hairpins, act as siRNAs and use the RNA interference machinery to trigger downstream effects.



PUBLICATIONS

Articles

Costas J, Gratacòs M, Escaramís G, Martín-Santos R, de Diego Y, Baca-García E, Canellas F, Estivil X, Guillamat R, Guitart M, Gutiérrez-Zotes A, García-Esteve L, Mayoral F, Moltó MD, Phillips C, Roca M, Carracedo A, Vilella E, Sanjuán .
"Association study of 44 candidate genes with depressive and anxiety symptoms in post-partum women."
J Psychiatr Res, in press.

Tortajada S, García-Gómez JM, Vicente J, Sanjuán J, e Frutos R, Martín-Santos R, García-Esteve L, Gornemann I, Gutiérrez-Zotes A, Canellas F, Carracedo A, Gratacos M, Guillamat R, Baca-García E, Robles M
"Prediction of postpartum depression using multilayer perceptrons and pruning."
Method Inform Med 48(3):291-8 (2009).

Brunet A, Armengol L, Heine D, Rosell J, García-Aragonés M, Gabau E, Estivill X, Guitart M.
"BAC array CGH in patients with Velocardiofacial syndrome-like features reveals genomic aberrations on chromosome region 1q21.1."
BMC Med Genet10:144. (2009).

Hüffmeier U, Bergboer JG, Becker T, Armour JA, Traupe H, Estivill X, Riveira-Munoz E, Mössner R, Reich K, Kurrat W, Wienker TF, Schalkwijk J, Zeeuwen PL, Reis A.
"Replication of LCE3C-LCE3B CNV as a Risk Factor for Psoriasis and Analysis of Interaction with Other Genetic Risk Factors."
J Invest Dermatol. 2009 Dec 17 [Epub ahead of print]

Pedraza V, Gomez-Capilla JA, Escaramis G, Gomez C, Torné P, Rivera JM, Gil A, Araque P, Olea N, Estivill X, Fárez-Vidal ME.
"Gene expression signatures in breast cancer distinguish phenotype characteristics, histologic subtypes, and tumor invasiveness."
Cancer 116(2):486-96. Published online: 22 Dec 2009.

Baca-Garcia E, Vaquero-Lorenzo C, Perez-Rodriguez MM, Gratacòs M, Bayés M, Santiago-Mozos R, Leiva-Murillo JM, de Prado-Cumplido M, Artes-Rodriguez A, Ceverino A, Diaz-Sastre C, Fernandez-Navarro P, Costas J, Fernandez-Piqueras J, Diaz-Hernandez M, de Leon J, Baca-Baldomero E, Saiz-Ruiz J, Mann JJ, Parsey RV, Carracedo A, Estivill X, Oquendo MA.
"Nucleotide variation in central nervous system genes among male suicide attempters."
Am J Med Genet B Neuropsychiatr Genet, 153B(1):208-13 (2009).

Pantano L, Estivill X, Martí E.
"SeqBuster, a bioinformatic tool for the processing and analysis of small RNAs datasets, reveals ubiquitous miRNA modifications in human embryonic cells."
Nucleic Acids Res. [Epub ahead of print]

Castro-Giner F, de Cid R, Gonzalez JR, Jarvis D, Heinrich J, Janson C, Omenaas ER, Matheson MC, Pin I, Anto JM, Wjst M, Estivill X, Kogevinas M.
"Positionally cloned genes and age specific effects in asthma and atopy: an international population-based cohort study (ECRHS)."
Thorax. [Epub ahead of print]

Castro-Giner F, Bustamante M, Ramon González J, Kogevinas M, Jarvis D, Heinrich J, Antó JM, Wjst M, Estivill X, de Cid R.
"A pooling-based genome-wide analysis identifies new potential candidate genes for atopy in the European Community Respiratory Health Survey (ECRHS)."
BMC Med Genet 10:128 (2009).

Castro-Giner F, Künzli N, Jacquemin B, Forsberg B, de Cid R, Sunyer J, Jarvis D, Briggs D, Vienneau D, Norback D, González JR, Guerra S, Janson C, Antó JM, Wjst M, Heinrich J, Estivill X, Kogevinas M.
"Traffic-related air pollution, oxidative stress genes, and asthma (ECHRS)."
Environ Health Perspect 117(12):1919-24 (2009).

Ribasés M, Bosch R, Hervás A, Ramos-Quiroga JA, Sánchez-Mora C, Bielsa A, Gastaminza X, Guijarro-Domingo S, Nogueira M, Gómez-Barros N, Kreiker S, Gross-Lesch S, Jacob CP, Lesch KP, Reif A, Johansson S, Plessen KJ, Knappskog PM, Haavik J, Estivill X, Casas M, Bayés M, Cormand B.
"Case-control study of six genes asymmetrically expressed in the two cerebral hemispheres: association of BAIAP2 with attention-deficit/hyperactivity disorder."
Biol Psychiatry 66(10):926-34 (2009).

Cukier HN,Rabionet R, Konidari I, Rayner-Evans MY, Baltos ML, Wright HH, Abramson RK, Martin ER, Cuccaro ML, Pericak-Vance MA, Gilbert JR.
"Novel variants identified in methyl-CpG-binding domain genes in autistic individuals."
Neurogenetics. [Epub ahead of print]

Albacar G, Sans T, Martín-Santos R, García-Esteve L, Guillamat R, Sanjuan J, Cañellas F, Carot JM, Gratacòs M, Bosch J, Gaviria A, Labad A, Zotes AG, Vilella E.
"Thyroid function 48h after delivery as a marker for subsequent postpartum depression."
Psychoneuroendocrinology. [Epub ahead of print]

Ogorelkova M, Navarro A, Vivarelli F, Ramirez-Soriano A, Estivill X.
"Positive selection and gene conversion drive the evolution of a brain-expressed snoRNAs cluster."
Mol Biol Evol 26(11):2563-71 (2009).

Navarro F, Gutman D, Meire E, Cáceres M, Rigoutsos I, Bentwich Z, Lieberman J.
"miR-34a contributes to megakaryocytic differentiation of K562 cells independently of p53."
Blood 114(10):2181-92 (2009).

de Cid R, Ramos MD, Aparisi L, García C, Mora J, Estivill X, Farré A, Casals T.
"Independent Contribution of Common CFTR Variants to Chronic Pancreatitis."
Pancreas. [Epub ahead of print]

Muiños-Gimeno M, Montfort M, Bayés M, Estivill X, Espinosa-Parrilla Y.
"Design and evaluation of a panel of single-nucleotide polymorphisms in microRNA genomic regions for association studies in human disease."
Eur J Hum Genet. Epub 2009 Oct 7. (2009).

Real E, Gratacòs M, Soria V, Escaramís G, Alonso P, Segalàs C, Bayés M, de Cid R, Menchón JM, Estivill X.
"A brain-derived neurotrophic factor haplotype is associated with therapeutic response in obsessive-compulsive disorder."
Biol Psychiatry 66(7):674-80 (2009).

Hammer C, Kapeller J, Endele M, Fischer C, Hebebrand J, Hinney A, Friedel S, Gratacòs M, Estivill X, Fichter M, Fernández-Aranda F, Ehrlich S, Rappold G, Niesler B.
"Functional variants of the serotonin receptor type 3A and B gene are associated with eating disorders."
Pharmacogenet Genomics 19(10):790-9 (2009).

Armengol L, Villatoro S, González JR, Pantano L, García-Aragonés M, Rabionet R, Cáceres M, Estivill X.
"Identification of copy number variants defining genomic differences among major human groups."
PLoS One 4(9):e7230 (2009).

Azkona G, Amador-Arjona A, Obradors-Tarragó C, Varea E, Arqué G, Pinacho R, Fillat C, de la Luna S, Estivill X, Dierssen M.
"Characterization of a mouse model overexpressing beta-site APP-cleaving enzyme 2 reveals a new role for BACE2."
Genes Brain Behav. [Epub ahead of print]

Gratacòs M, Costas J, de Cid R, Bayés M, González JR, Baca-García E, de Diego Y, Fernández-Aranda F, Fernández-Piqueras J, Guitart M, Martín-Santos R, Martorell L, Menchón JM, Roca M, Sáiz-Ruiz J, Sanjuán J, Torrens M, Urretavizcaya M, Valero J, Vilella E, Estivill X, Carracedo A; Psychiatric Genetics Network Group.
"Identification of new putative susceptibility genes for several psychiatric disorders by association analysis of regulatory and non-synonymous SNPs of 306 genes involved in neurotransmission and neurodevelopment."
Am J Med Genet B Neuropsychiatr Genet 150B(6):808-16 (2009).

Marti A, Santos JL, Gratacos M, Moreno-Aliaga MJ, Maiz A, Martinez JA, Estivill X.
"Association between leptin receptor (LEPR) and brain-derived neurotrophic factor (BDNF) gene variants and obesity: a case-control study."
Nutr Neurosci 12(4):183-8 (2009).

Muiños-Gimeno M, Guidi M, Kagerbauer B, Martín-Santos R, Navinés R, Alonso P, Menchón JM, Gratacòs M, Estivill X, Espinosa-Parrilla Y.
"Allele variants in functional MicroRNA target sites of the neurotrophin-3 receptor gene (NTRK3) as susceptibility factors for anxiety disorders."
Hum Mutat 30(7):1062-71 (2009).

Ma DQ, Rabionet R, Konidari I, Jaworski J, Cukier HN, Wright HH, Abramson RK, Gilbert JR, Cuccaro ML, Pericak-Vance MA, Martin ER.
"Association and gene-gene interaction of SLC6A4 and ITGB3 in autism."
Am J Med Genet B Neuropsychiatr Genet [Epub ahead of print]

Ponsa I, Ramos-Quiroga JA, Ribasés M, Bosch R, Bielsa A, Ordeig MT, Morell M, Miró R, de Cid R, Estivill X, Casas M, Bayés M, Cormand B, Hervás A.
"Absence of cytogenetic effects in children and adults with attention-deficit/hyperactivity disorder treated with methylphenidate."
Mutat Res 666(1-2):44-9 (2009).

Rodríguez-Santiago B, Brunet A, Sobrino B, Serra-Juhé C, Flores R, Armengol L, Vilella E, Gabau E, Guitart M, Guillamat R, Martorell L, Valero J, Gutiérrez-Zotes A, Labad A, Carracedo A, Estivill X, Pérez-Jurado LA.
"Association of common copy number variants at the glutathione S-transferase genes and rare novel genomic changes with schizophrenia."
Mol Psychiatry [Epub ahead of print]

González JR, Subirana I, Escaramís G, Peraza S, Cáceres A, Estivill X, Armengol L.
"Accounting for uncertainty when assessing association between copy number and disease: a latent class model."
BMC Bioinformatics10:172 (2009).

Morales E, Sunyer J, Julvez J, Castro-Giner F, Estivill X, Torrent M, De Cid R.
"GSTM1 polymorphisms modify the effect of maternal smoking during pregnancy on cognitive functioning in preschoolers."
Int J Epidemiol 38(3):690-7 (2009).

Hüffmeier U, Estivill X, Riveira-Munoz E, Traupe H, Wendler J, Lohmann J, Böhm B, Burkhardt H, Reis A.
"Deletion of LCE3C and LCE3B genes at PSORS4 does not contribute to Susceptibility to Psoriatic Arthritis in German patients."
Ann Rheum Dis. [Epub ahead of print].

Castro-Giner F, Kogevinas M, Imboden M, de Cid R, Jarvis D, Mächler M, Berger W, Burney P, Franklin KA, Gonzalez JR, Heinrich J, Janson C, Omenaas E, Pin I, Rochat T, Sunyer J, Wjst M, Antó JM, Estivill X, Probst-Hensch NM.
"Joint effect of obesity and TNFA variability on asthma: two international cohort studies."
Eur Respir J 33(5):1003-9 (2009).

Beà S, Salaverria I, Armengol L, Pinyol M, Fernández V, Hartmann EM, Jares P, Amador V, Hernández L, Navarro A, Ott G, Rosenwald A, Estivill X, Campo E.
"Uniparental disomies, homozygous deletions, amplifications, and target genes in mantle cell lymphoma revealed by integrative high-resolution whole-genome profiling."
Blood 113(13):3059-69 (2009).

Kaput J, Cotton RG, Hardman L, Watson M, Al Aqeel AI, Al-Aama JY, Al-Mulla F, Alonso S, Aretz S, Auerbach AD, Bapat B, Bernstein IT, Bhak J, Bleoo SL, Blöcker H, Brenner SE, Burn J, Bustamante M, Calzone R, Cambon-Thomsen A, Cargill M, Carrera P, Cavedon L, Cho YS, Chung YJ, Claustres M, Cutting G, Dalgleish R, den Dunnen JT, Díaz C, Dobrowolski S, dos Santos MR, Ekong R, Flanagan SB, Flicek P, Furukawa Y, Genuardi M, Ghang H, Golubenko MV, Greenblatt MS, Hamosh A, Hancock JM, Hardison R, Harrison TM, Hoffmann R, Horaitis R, Howard HJ, Barash CI, Izagirre N, Jung J, Kojima T, Laradi S, Lee YS, Lee JY, Gil-da-Silva-Lopes VL, Macrae FA, Maglott D, Marafie MJ, Marsh SG, Matsubara Y, Messiaen LM, Möslein G, Netea MG, Norton ML, Oefner PJ, Oetting WS, O'Leary JC, de Ramirez AM, Paalman MH, Parboosingh J, Patrinos GP, Perozzi G, Phillips IR, Povey S, Prasad S, Qi M, Quin DJ, Ramesar RS, Richards CS, Savige J, Scheible DG, Scott RJ, Seminara D, Shephard EA, Sijmons RH, Smith TD, Sobrido MJ, Tanaka T, Tavtigian SV, Taylor GR, Teague J, Töpel T, Ullman-Cullere M, Utsunomiya J, van Kranen HJ, Vihinen M, Webb E, Weber TK, Yeager M, Yeom YI, Yim SH, Yoo HS; Contributors to the Human Variome Project Planning Meeting.
"Planning the human variome project: the Spain report."
Hum Mutat 30(4):496-510 (2009).

Forcano L, Fernández-Aranda F, Alvarez-Moya E, Bulik C, Granero R, Gratacòs M, Jiménez-Murcia S, Krug I, Mercader JM, Riesco N, Saus E, Santamaría JJ, Estivill X.
"Suicide attempts in bulimia nervosa: personality and psychopathological correlates."
Eur Psychiatry 24(2):91-7 (2009).

Toll-Riera M, Bosch N, Bellora N, Castelo R, Armengol L, Estivill X, Albà MM.
"Origin of primate orphan genes: a comparative genomics approach."
Mol Biol Evol 26(3):603-12 (2009).

de Cid R, Riveira-Munoz E, Zeeuwen PL, Robarge J, Liao W, Dannhauser EN, Giardina E, Stuart PE, Nair R, Helms C, Escaramís G, Ballana E, Martín-Ezquerra G, den Heijer M, Kamsteeg M, Joosten I, Eichler EE, Lázaro C, Pujol RM, Armengol L, Abecasis G, Elder JT, Novelli G, Armour JA, Kwok PY, Bowcock A, Schalkwijk J, Estivill X.
"Deletion of the late cornified envelope LCE3B and LCE3C genes as a susceptibility factor for psoriasis."
Nat Genet 41(2):211-5 (2009).

Zadro C, Alemanno MS, Bellacchio E, Ficarella R, Donaudy F, Melchionda S, Zelante L, Rabionet R, Hilgert N, Estivill X, Van Camp G, Gasparini P, Carella M.
"Are MYO1C and MYO1F associated with hearing loss?"
Biochim Biophys Acta 1792(1):27-32 (2009).

Ribasés M, Ramos-Quiroga JA, Hervás A, Bosch R, Bielsa A, Gastaminza X, Artigas J, Rodriguez-Ben S, Estivill X, Casas M, Cormand B, Bayés M.
"Exploration of 19 serotoninergic candidate genes in adults and children with attention-deficit/hyperactivity disorder identifies association for 5HT2A, DDC and MAOB."
Mol Psychiatry 14(1):71-85 (2009).

Soria V, Martínez-Amorós E, Escaramís G, Valero J, Pérez-Egea R, García C, Gutiérrez-Zotes A, Puigdemont D, Bayés M, Crespo JM, Martorell L, Vilella E, Labad A, Vallejo J, Pérez V, Menchón JM, Estivill X, Gratacòs M, Urretavizcaya M.
"Differential Association of Circadian Genes with Mood Disorders: CRY1 and NPAS2 are Associated with Unipolar Major Depression and CLOCK and VIP with Bipolar Disorder."
Neuropsychopharmacology, in press.

Bosch N, Morell M, Ponsa I, Mercader JM, Armengol L, Estivill X.
"Nucleotide, cytogenetic and expression impact of the human chromosome 8p23.1 inversion polymorphism."
PLoS One 4(12):e8269 (2009).

Morales E, Julvez J, Torrent M, de Cid R, Guxens M, Bustamante M, Künzli N, Sunyer J.
"Association of early-life exposure to household gas appliances and indoor nitrogen dioxide with cognition and attention behavior in preschoolers."
Am J Epidemiol 169(11):1327-36 (2009).

Nelis M, Esko T, Mägi R, Zimprich F, Zimprich A, Toncheva D, Karachanak S, Piskácková T, Balascák I, Peltonen L, Jakkula E, Rehnström K, Lathrop M, Heath S, Galan P, Schreiber S, Meitinger T, Pfeufer A, Wichmann HE, Melegh B, Polgár N, Toniolo D, Gasparini P, D'Adamo P, Klovins J, Nikitina-Zake L, Kucinskas V, Kasnauskiene J, Lubinski J, Debniak T, Limborska S, Khrunin A, Estivill X, Rabionet R, Marsal S, Julià A, Antonarakis SE, Deutsch S, Borel C, Attar H, Gagnebin M, Macek M, Krawczak M, Remm M, Metspalu A.
"Genetic structure of Europeans: a view from the North-East."
PLoS One 4(5):e5472 (2009).

Reviews

Dierssen M, Herault Y, Estivill X.
"Aneuploidy: from a physiological mechanism of variance to Down syndrome."
Physiol Rev 89(3):887-920 (2009).

Letters

O'Roak B J, Morgan TM, Fishman D, Saus E, Alonso P, Gratacòs M, Estivill X, Teltsh O, Kohn Y, Kidd KK, Cho J, Lifton RP, State MW.
"Additional Support for the Association of SLITRK1 var321 and Tourette syndrome."
Molecular Psychiatry, in press.

Book chapters

Estivill X
"Bases moleculares de la herencia. Genes y enfermedad."
In: Medicina Interna, Decimosexta Edición, Ed. Farreras Rozman. Sección 9 Genética Médica. (2009).

Sala N, Bayés M, Armengol LI, Estivill X.
"Genoma humano. Métodos de análisis."
In: Medicina Interna, Decimosexta Edición, Ed. Farreras Rozman. Sección 9 Genética Médica. (2009).

Estivill X.
"Contribuciones Españolas a la Genética Humana y la Medicina Molecular."
In: España Siglo XXI. Instituto de España, Ed. Biblioteca Nueva. Tomo IV. (2009).

GENES AND DISEASE

Group: Gene Therapy

Group structure:

Group Leader: Cristina Fillat

Postdoctoral Fellows: Xavier Altafaj, Laura García, Maria Victoria Maliandi, Luciano Sobrevals

PhD Students: Daniel Abate Daga, Anabel Jose, Xavier Bofill, Ana Mato

Technicians: Núria Andreu

SUMMARY

Our laboratory focuses on the development of gene therapy strategies to understand and treat human complex genetic diseases. The advances on the basic biological understanding of molecular and cellular events underlying specific disease pathophysiology, facilitates the development of gene therapy in a broad spectrum of human diseases. In the coming decade, it is foreseen that molecular therapy in general and gene therapy in particular will be used to treat patients on an individual basis. Gene therapy has also a great potential in complex diseases as a powerful genetic tool to contribute to the identification of the functional role of specific genes and/or small non-coding elements. Our research is interested in genetic engineering viral vectors that can selectively and efficiently target specific cell types and to study their impact in living animals. Pharmacokinetic and pharmacodynamic studies are being conducted to evaluate therapeutic response or phenotype rescue in preclinical mouse models.

RESEARCH PROJECTS

1. Pancreatic cancer

Pancreatic cancer is a neoplasia with a very bad prognosis mainly due to the late diagnosis and inefficient current therapies. We take advantage of the current understanding of the mechanistic pathogenesis and of the comprehensive available data regarding inefficient therapeutic response to design alternative treatments. In the past few years we have been involved in exploring the feasibility of suicide gene-virotherapy. This system is based on the combination of two different principles, based on the attempt to achieve tumor destruction by allowing lytic adenovirus to replicate in tumor cells, multiply and disseminate, and by arming these viruses with the Herpes Simplex Thymidine Kinase gene (TK) that when combined with GCV induces apotosis mediated cell killing in cycling cells. We have proved to boost the potency of adenoviruses when TK was expressed at the late phase of viral replication. We are actively working on providing tumor selectivity to those viruses. This year we have shown that the transcriptional control of key viral genes using particular promoters confers selective replication. With the analysis of the keratin 7 promoter (K7p), a keratin expressed in simple epithelia, we have shown that by driving transcription with K7p it is feasible to achieve a restricted expression of the controlled genes, while mantains a selective expression in ductal cells of the pancreas. The studies on the uPAR promoter have shown the potential of this regulatory sequence to slectively and actively express genes in pancreatic tumor cells, and to trigger adenoviral replication efficiently to stop tumor progression and selectively eradicate liver metastasis in mouse models of tumorogenesis. We are currently working on the selective tumor control of adenoviral replication by means of several post-transcriptional strategies together with viral capsid modifications.

The group is also interested in the understanding of the basic mechanisms of tumor cell killing induced by viral lysis and the TK/GCV systems and on the definition of its activity in the heterogeneous population of cells that constitute a tumour.

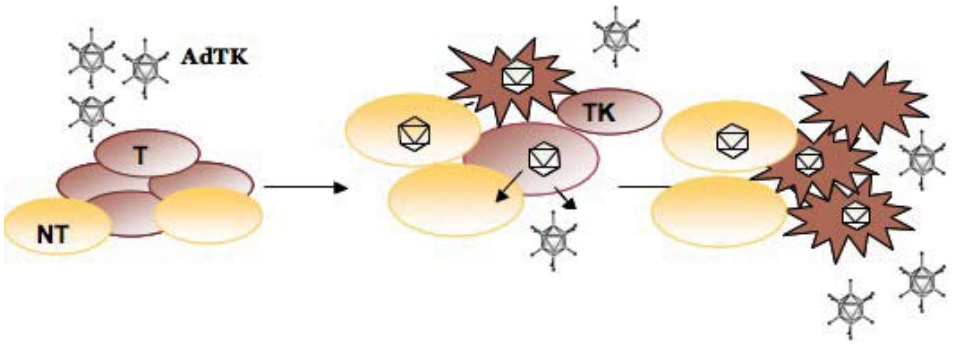


Figure 1.
Tumour Selective Gene-Virotherapy

2. Down Syndrome

Down syndrome (DS) is a complex multi-system disorder, resulting from an extra copy of the human chromosome 21 gene HSA21. Although HSA21 contains more than three hundred genes, it has been hypothesized that the excess of specific target genes (dosage-sensitive) will have a major impact on the DS phenotype. To better understand the contribution of these specific genes, we have engineer a novel approach that takes advantage of the current knowledge from the existing mouse models. This approach is based on the rescue of a phenotype in a mouse model by the viral delivery of shRNA sequences targeting a candidate gene. A proof of principle approach was developed and shown that normalizing the expression of the Dyrk1A dosage sensitive gene in the striatum of adult TgDyrk1A by AAVshDyrk1A delivery rescued motor impairment. This approach highlighted Dyrk1A phenotypic dependence in motor alterations. Mental retardation is the more diabling trait in DS. We are currently studying the impact that, normalizing Dyrk1A in the hippocampus can triggger on the cognitive function in DS models. Studies on the molecular mechanisms by which the normalization of Dyrk1A can reverse such phenotypes are also being explored. Despite the contribution triggered by the dosage-sensitive genes to the phenotype the over-representation of functional non-protein-coding elements might be involved in some of the abnormal phenotypes. On going efforts involve the development of viral gene-transfer approaches that help to evaluate the contribution of these non-coding elements to the cognitive phenotype of DS murine models.

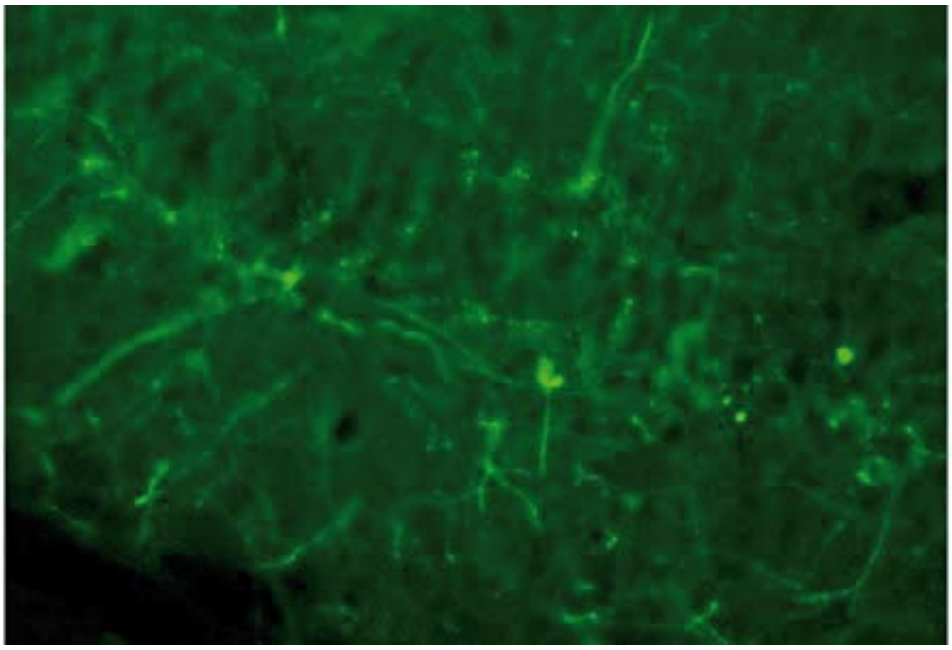


Figure 2.
Hippocampal neurons in vivo transduced with AAV2/1-GFP

PUBLICATIONS

Pujal J*, Huch M*, José A, Abasolo I, Rodolosse A, Duch A, Sánchez-Palazón L, Smith FJ, McLean WH, Fillat C*, Real FX*.
"Keratin 7 promoter selectively targets transgene expression to normal and neoplastic pancreatic ductal cells in vitro and in vivo."
FASEB J 23:1366-75 (2009). *(shared credit)

Martínez-Quintanilla J, Cascalló M, Fillat C, Alemany R.
"Antitumor therapy based on cellular competition."
Hum Gene Ther 20:728-738 (2009).

Huch M, Gros A, José A, González JR, Alemany R, Fillat C.
"uPAR Transcriptionally Controlled Adenoviruses Eradicate Pancreatic Tumors and Liver Metastasis in Mouse Models."
Neoplasia 11:518-528 (2009).

Martínez-Quintanilla J, Cascalló M, Gros A, Fillat C, Alemany R.
"Positive selection of gene-modified cells to increase the efficacy of suicide gene therapy."
Mol Cancer Ther 8: 3098-3107 (2009).

Azkona G, Amador A, Obradors C, Varea E, Arqué G, Pinacho R, Fillat C, de la Luna S, Estivill X, Dierssen M.
"Characterization of a mouse model overexpressing beta-site APP-cleaving enzyme 2 reveals a new role for BACE2."
Genes Brain Behav [Epub ahead of print]

Fillat C, José A, Bofill-De Ros X, Mato-Berciano A, Maliandi MV, Abate-Daga D.
"Controlling adenoviral replication to induce oncolytic efficacy."
The Open Gene Ther J, in press.

Fillat C, Dierssen M, Martínez de Lagrán M, Altafaj X.
"Insights from mouse models to understand neurodegeneration in Down Syndrome."
CNS Neurol Disor Drug Targets Review, in press.

Albert MH, Bittner TC, Nonoyama S, Notarangelo L, Burns S, Imai K, Español T, Fasth A, Pellier I, Strauss G, Morio T, Gathman B, Noordzij JG, Fillat C, Hoenig M, Nathrath M, Meindl A, Pagel P, Wintergerst U, Fisher A, Trasher AJ, Belohradsky BH, Ochs HD.
"X-linked thrombocytopenia (XLT) due to WAS mutations: clinical characteristics, long-term outcome and treatment options."
Blood, in press.

Book Chapters

Fillat C, Abate-Daga D.
"Terapia Génica. Biotecnología y Biofármacos."
Ed. Colegio Oficial de Farmacéuticos, in press.

Fillat C, González-Foix A Ma, Arán JM.
"Tratamiento de enfermedades genéticas y modelos de enfermedad."
In: Medicina Interna, Decimosexta Edición, Ed. Farreras Rozman. Sección 9 Genética Médica (2009).

Other Publications

Dierssen M, Azcona G, Fillat C, Toiber D, Soreq H.
"Bases moleculares de la neuropatología del síndrome de Down: implicación de DYRK1A."
Publicación del VI Premio «Jaime Blanco» de investigación en el síndrome de Down. Editorial Nueva Imprenta, Madrid (2009).

PATENTS

Inventors: Cristina Fillat, Meritxell Huch
Title: *"Adenovirus de replicación condicional efectivo en el tratamiento de tumores"*
Application form No: ES P200900597
Priority Country: Nacional
Priority date: 25/02/2009
Companies that are licensing it: not being licensed

GENES AND DISEASE

Group: Gene Function and Murine Models of Disease
Susana de la Luna has an ICREA Group Leader position.

Group structure:

Group Leaders: Mariona Arbonés / Susana de la Luna

Postdoctoral: María José Barallobre (CIBERER researcher, since Sept 2009), Esteban Rozen (since July 2009)

PhD Students: Krisztina Arató, Elisa Balducci, Chiara di Vona, Eulàlia Salichs (until July 2009)

Master Students: Sonia Najas, Andrea Senna (CIBERER researcher, since March 2009)

Technicians: Erika Ramirez, Alicia Raya

SUMMARY

The ability of a human being to perform complex tasks relies on the correct formation of neuronal circuits. For this to happen, the many cell types of a mature brain have to be generated at the correct numbers and then differentiate properly during development, a complex process that is exquisitely regulated in time and space. Mutations disturbing the division mode of neuronal precursor cells, cell fate acquisition, differentiation, or natural cell death are likely to cause alterations in brain connectivity resulting in neurological diseases and mental retardation.

Trisomy 21 is the main genetic cause of mental retardation, resulting from neurodevelopmental alterations and also changes in brain homeostasis in the adult. Research in the past few years has revealed that a number of human chromosome 21 (HSA21) genes are overexpressed in Down syndrome by at least 50% due to gene dosage. Because of the complexity of the Down syndrome phenotype, it is very likely that the increased expression leads to perturbations in a great variety of biological pathways. Furthermore, it is predictable that many HSA21 genes can interact functionally with each other within particular signalling pathways. Understanding the functional roles of the overexpressed genes will not only help to delineate the specific biological or biochemical processes affected but also to identify pathways that are particularly sensitive to dosage variations in any of their components. In this regard, strong emphasis has been made during the last years in trying to elucidate the role of a HSA21 kinase, DYRK1A, in central nervous system development.

The group studies the functional roles of several HSA21 genes using a combination of *in vitro* and *in vivo* analysis that involved the use of mouse models with different dosage of the genes of interest.

RESEARCH PROJECTS

1. Assigning functional roles to HSA21 genes.

Two HSA21 genes are special focus of the group interest, the protein kinase DYRK1A and the calcineurin regulator RCAN1/DSCR1. The products of the two genes participate in controlling the phosphorylation status of their target proteins, either directly in the case of the kinase or indirectly through the inhibition of the phosphatase calcineurin, in the case of RCAN1. The two proteins share targets such as the family of transcription factors NFAT or the endocytic protein dynamin, and their action on them leads to synergistically induce a hyperphosphorylated state.

One of the approaches we have chosen to elucidate the participation of DYRK1A in different cellular processes and signal transduction pathways is to inferred it from the activity of its interactors and substrates. With this approach, we have succeeded in establishing a functional link with receptor tyrosine kinase (RTK) signal pathways through the functional characterization of the RTK antagonist Sprouty 2 as a partner and substrate of DYRK1A. Likewise, the identification of caspase-9 as a substrate of DYRK1A has allowed us to associate pro-survival activities to DYRK1A. Other putative DYRK1A interacting-proteins are currently being explored.

Our work on RCAN1 has allowed us to propose a molecular mechanism for its inhibitory activity on calcineurin. The inhibitory action of RCANs on calcineurin-NFAT signaling results not only from the inhibition of phosphatase activity but also from competition between NFAT and RCAN for binding to the same docking site in calcineurin. The two acitivities of RCANs, binding to and inhibition of calcineurin, are structurally linked but do not overlap. Notably, this situation is reminiscent of some inhibitors of the phosphatase PP1. Thus, we have proposed that this mechanism, a combination of a docking site with a cis-regulatory element, could be a feature of the phosphatase superfamily. Our results have also highlighted that competition by substrates and modulators for a common docking site appears to be an essential mechanism in the regulation of Ca^{2+}-calcineurin signaling.

2. DYRK1A roles during nervous system development

DYRK1A plays an important role in mammalian brain development because mice and children with a loss-of function mutation in ***DYRK1A locus*** present microcephaly. The underlying pathological mechanisms are not well understood.

Some brain structures in Dyrk1a$^{+/-}$ heterozygous mutant mice present an important neuron loss. Likewise, the retina, which is another central nervous system structure that is significantly reduced in size in Dyrk1a$^{+/-}$ mice, present a severe hypocellularity. Conversely, mice with an additional copy of Dyrk1a gene have bigger retinas with increased number of neurons. To get some insights into the role of DYRK1A in nervous system development we have chosen the retina as a model system. We demonstrated that differences in retina cellularity in Dyrk1a mutants are due to an altered physiological apoptosis. In addition, we showed that caspase-9 is phosphorylated at a major inhibitory residue, Thr125, by DYRK1A and that the levels of phosphorylated caspase-9 in the developing retina correlate with the magnitude of apoptosis, which is decreased in the loss-of function mutants and increased in the gain-of function mutants. These results showed that phosphorylation of caspase-9 by DYRK1A is the mechanism whereby DYRK1A regulates cell number in the retina.

Examination of the formation of particular type of brain neurons, i.e. midbrain dopamine neurons, in mice with 1 or 3 functional copies of Dyrk1a have revealed an altered developmental apoptosis in the brain of these mutants. Like in the retina, apoptosis in the brain is increased in Dyrk1a loss-of function mutants and decreased in the gain-of function mutants. These observation indicate that deregulation of caspase-9 mediated apoptosis is likely contributing to the altered brain size and cellularity in Dyrk1a mutant mice and, probably, in people with aneuploidies involving DYRK1A gene. The putative role of DYRK1A in controlling the homeostasis of neuronal precursor cells in the brain is now under investigation.

3. Functional impact of homopolymeric histidine repeats

Eulália Salichs

Based on our findings that the histidine (His)-rich segment present within the DYRK1A carboxy-end is a signal for targeting to the subnuclear compartment known as nuclear speckles or splicing factor compartment, we would like to find out the functional impact of homopolymeric histidine repeats. This research line is carried out in collaboration with Mar Albá (ICREA-IMIM, Barcelona).

Single amino acid repeats, also known as homopolymeric tracts, are very common in eukaryotes and between 18-20% of proteins in the human genome contain such repetitive sequences. Although most of them are thought to be functionally neutral, recent evidence suggests they may play important functional or structural roles. Indeed, several studies have demonstrated that there is a strong correlation between genes associated with human diseases and genes encoding proteins with amino acid runs. Among homopolymeric tracts, His-repeats are relatively rare in eukaryotes. The physicochemical properties of His make it a versatile amino acid that can fulfill different roles, influencing protein conformation and enzymatic activity. Our work has shown that His-repeats can work as a general targeting signal to nuclear speckles, and further suggested that these repeats are a way of generating evolutionary diversification in gene duplicates.

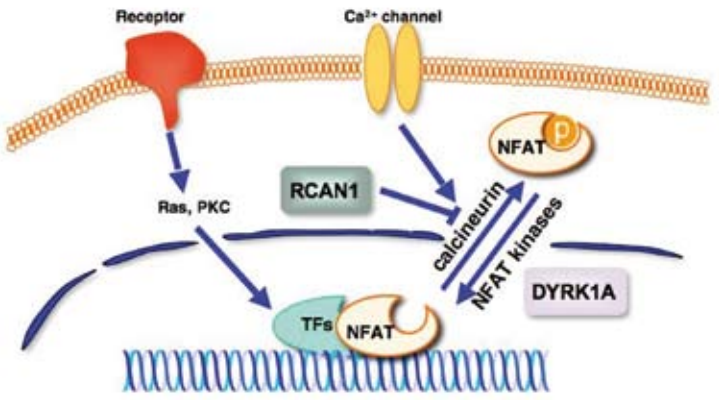


Figure 1.
Synergistic effect of DYRK1A and RCAN1 on the calcineruin-NFAT signalling pathway.

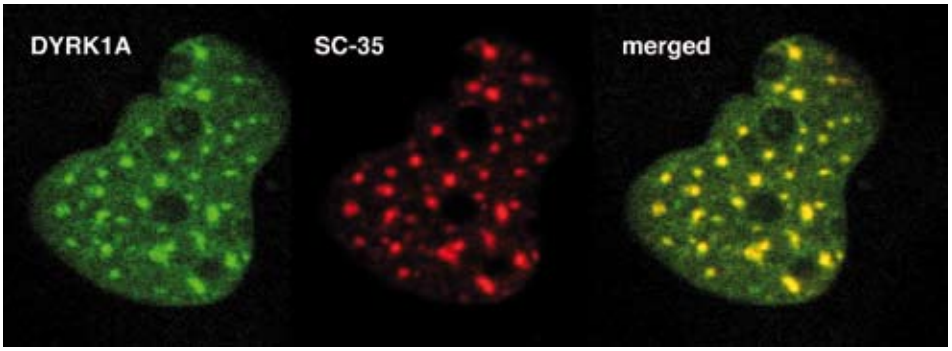


Figure 2.
DYRK1A accumulates in nuclear speckles (or splicing factor compartment) via its homopolymeric histidine repeat.

PUBLICATIONS

Arqué G, de Lagrán MM, Arbonés ML, Dierssen M.
"Age-associated motor and visuo-spatial learning phenotype in Dyrk1A heterozygous mutant mice."
Neurobiol Dis 36: 312-319 (2009).

Guedj F, Sébrié C, Rivals I, Ledru A, Paly E, Bizot JC, Smith D, Rubin E, Gillet B, Arbones M, Delabar JM.
"Green tea polyphenols rescue of brain defects induced by overexpression of DYRK1A."
PLoS One 4(2): e4606 (2009).

Martínez-Martínez S*, Genescà L*, Rodríguez A, Raya A, Salichs E, Were F, López-Maderuelo MD, Redondo JM, de la Luna S.
"The RCAN carboxyl end mediates calcineurin docking-dependent inhibition via a site that dictates binding to substrates and regulators."
Proc Natl Acad Sci USA 106: 6117-6122 (2009). (*equal contribution)

Nikolaienko O, Skrypkina I, Tsyba L, Fedyshyn Y, Morderer D, Buchman V, de la Luna S, Drobot L, Rynditch A.
"Intersectin 1 forms a complex with adaptor protein Ruk/CIN85 in vivo independently of epidermal growth factor stimulation."
Cell Signal 21: 753-759 (2009).

Salichs E, Ledda A, Mularoni L, Albà MM, de la Luna S.
"Genome-wide analysis of histidine repeats reveals their role in the localization of human proteins to the nuclear speckles compartment."
PLoS Genet 5(3): e1000397 (2009).

Azkona G, Amador-Arjona A, Obradors-Tarragó C, Varea E, Arqué G, Pinacho R, Fillat C, de la Luna S, Estivill X, Dierssen M.
"Characterization of a mouse model overexpressing beta-site APP-cleaving enzyme 2 reveals a new role for BACE2."
Genes Brain Behav. 2009 Sep 22. [Epub ahead of print]

GENES AND DISEASE

Group: Neurobehavioral Phenotyping of Mouse Models of Disease

Group structure:

Group Leader:	Mara Dierssen
Predoctoral Students:	Carla Obradors, Susanna Molas, Davide D'Amico, Meritxell Pons
Postdoctoral fellows:	Xavier Gallego Moreno, Monica Joana Pinto Do Santos
Technicians:	María Martínez de Lagrán Cabredo, Nuno Vasconcelos
Mouse phenotyping:	Ignasi Sahún Abizanda, Jerome MacDonald
Visiting Professor:	John Crabbe
Master/Graduate Students:	Isabel Fernández Vargas, Ornella Spadoni, Eva Aparicio, Marc Murall, Patricia Carrasco, Mariona Font, Cristina Vilella, Aida Regi

SUMMARY

The overall goal of our research is to understand the role of putative candidate genes for human complex genetic diseases that affect cognitive systems, using genetically modified mouse models as our main experimental tool. The characterization of these models allows obtain better knowledge of the genetic substrates regulating the expression of complex behavioral traits and the pathogenesis of neuropsychiatric and neurological disorders. Understanding the genetic and neural circuits disturbed in mental retardation and neuropsychiatric disorders is one of the significant challenges in ultimately treating it. Answers may emerge from systems neuroscience approaches that combine cognitive, imaging, and genetic analyses with the results from animal and cellular models.

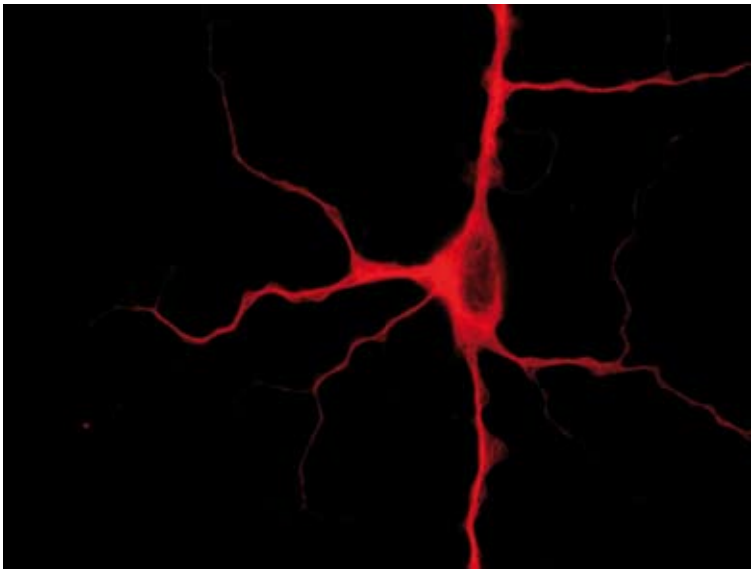


Figure 1.

RESEARCH PROJECTS

1. Mental retardation

The first research line of our group is aimed at dissecting the genetic architecture of mental retardation. During the last year we have contributed to characterize the functional and structural organization of the cognitive networks in Down syndrome (DS) models and the mechanistic aspects underlying the neuropathology. There are a number of mouse models of DS that replicate some of the cognitive and behavioral alterations of DS humans (for a review see Dierssen et al 2009). However, the link between the genetic and neuronal basis of the alteration and the corresponding behavioral and cognitive phenotype is still missing.

We have addressed different problems during 2009:

a) In order to explore the complex genetic interaction networks, we have performed a large-scale analysis of gene interactions involved in cognitive phenotype using genetically engineered aneuploid mice for overlapping mouse chromosomes homologous to HSA21 (Lopes-Pereira et al 2009). These tools will help to dissect the associated to specific trisomies/monosomies and their combinatory interactions. We have also continued the characterization of single gene transgenic models focusing on the fine dissection of different forms of learning and memory and other cognitive domains using innovative experimental paradigms and tracking tools (Arqué et al 2009, Azkona et al 2009, Azkona et al submitted) and at the molecular levels (Azkona, PhD thesis; Azkona et al, in press; J. Blanco Award 2009).

b) Dissection of the molecular mechanisms affecting neuronal network formation in primary cortical cultures, analyzing genetically-driven alterations in dendritic structure (Martínez de Lagrán et al in preparation) and the possible involvement of Rho signaling and actin cytoskeleton (EU project granted; M Dierssen, coordinator).

c) Epigenetic contributions to the functional and structural anomalies in DS. We have analysis histone deacetylation patterns (coll. L diCroce), and DNA methylation patterns (coll. M Esteller) and to the phenotypes of cortical neurons, by generating mice with chimeric cortices in which neuronal precursor cells of trisomic or transgenic (DS) genotypes are under the influence of euploid host neurons (G Arqué, S Martínez, M Dierssen, in preparation).

d) We explore if environmental complexity ameliorates motor, cognitive and emotional functions with stable consequences on the neuronal organization of perception and cognition and impact on the organization of the perceptual and cognitive systems to assess whether synthetic enrichment and augmented feedback is effective in promoting cognition in DS (M Pons PhD thesis).

e) Several physiological mechanisms of cortical neurons and synapses play a key role in constraining the extent and strength of network activation, among them are the balance between excitation and inhibition in the cortex (coll. MV Sánchez-Vives; IDIBAPS) and the cerebellum (Coll. P Verschure, D Konforty, UPF).

f) We are also collaborating with other research groups focusing on neurodegenerative disorders (Giralt et al 2009; Botella-López et al 2009).

Figure 2.

2. Neuropsychiatric disorders

A more recent research line is aimed at identifying genetic causative and vulnerability factors underlying neuropsychiatric disorders. We focus on:

a) Panic disorder
We address the pathogenetic mechanisms that underlie the inability of persons with anxiety disorders to correctly identify the fear-related information using genetically modified mouse models (D'Amico, PhD thesis). We also study the common neurobiological pathways responsible for co-morbid processes such as stress and drug abuse disorders (Amador-Arjona et al 2009; Gallego et al 2009), and we analyze the structure of fear memories: modulation of persistence of amygdala- and hippocampus-dependent fear memory, varying hippocampal involvement with contextual pre-exposure and contextual blocking, background and foreground context and trace conditioning protocols. In this regards it is important top study the modulation of memory persistence in training conditions with different salience, and the contribution of different candidate genes to the sensitivity to panicogenic/panicolitic agents (M Do Santos, in preparation).

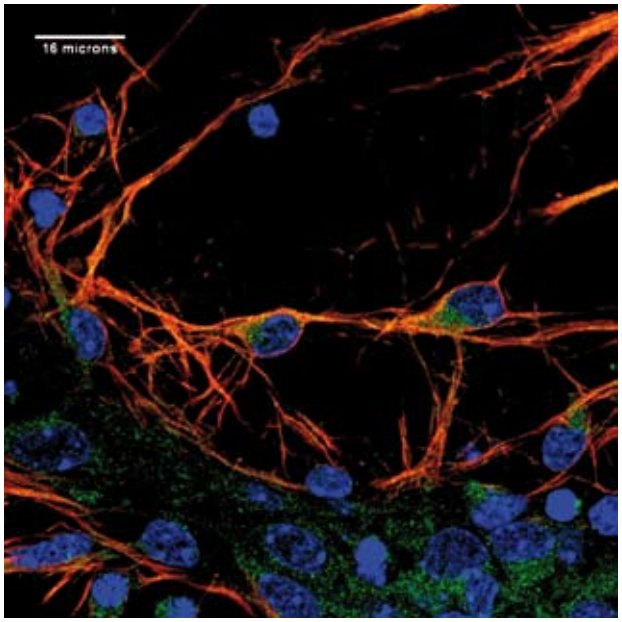


Figure 3.

b) Nicotine dependence
The second line is focused on the genetic-susceptibility to nicotine dependence (X. Gallego et al, presubmission enquiry accepted in Nat. Neurosci; S. Molas PhD project)

c) Obsessive-compulsive behavioral components of obesity
We focus on the development of a behavioral model of the obsessive-compulsive elements involved in eating disorders (J. MacDonald; Heyne et al 2009) and in the bioinformatic analysis of the behavioral structure.

PUBLICATIONS

Gallego X, Murtra P, Zamalloa T, Canals JM, Pineda J, Amador-Arjona A, Maldonado R, Dierssen M.
"Increased opioid dependence in a mouse model of panic disorder."
Frontiers in Behavioral Neurosci, in press.

Fillat C, Dierssen M, Martínez de Lagrán M, Altafaj X.
"Insights from mouse models to understand neurodegeneration in Down Syndrome."
CNS Neurol Disor Drug Targets, Review, in press.

Botella-López A, Cuchillo-Ibáñez I, Cotrufo T, Su San Mok, Qiao-Xin Li, Barquero M-S, Dierssen M, Soriano E Sáez-Valero J.
"Beta-amyloid controls altered Reelin expression and processing in Alzheimer's disease."
Neurobiol Disease [Epub 2009 Dec 16]

Heyne A, Kiesselbach C, Sahún I, McDonald J, Gaiffi M, Dierssen M, Wolffgramm J.
"An animal model of compulsive food taking behaviour."
Addiction Biology, 14, 373–383 (2009).

Azkona G, Amador-Arjona A, Obradors-Tarragó C, Varea E, Arqué G, Pinacho R, Fillat C, de la Luna S, Estivill X, Dierssen M.
"Characterization of a mouse model overexpressing BACE2 reveals a new role for BACE2."
Genes Brain Behav, [Epub ahead of print]

Lopes Pereira P, Magnol L, Sahún I, Brault V, Duchon A, Prandini P, Gruart A, Bizot J-C, Chadefaux-Vekemans B, Deutsch S, Trovero F, Delgado-García JM, Antonarakis SE, Dierssen M, Herault Y.
"A new mouse model for the trisomy of the abcg1-u2af1 region reveals the complexity of the combinatorial genetic code of Down syndrome."
Human Mol Genet 18:4756-69 (2009).

Amador-Arjona A, Delgado-Morales R, Belda X, Gagliano H, Gallego X, Keck ME, Armario A, Dierssen M.
"Susceptibility To Stress In Transgenic Mice Overexpressing TrkC, A Model Of Panic Disorder."
J Psychiat Res, [Epub 2009 Aug 20]

Arqué G, Martínez de Lagrán M, Arbonés ML, Dierssen M.
"Age-associated motor and visuo-spatial learning phenotype in Dyrk1A heterozygous mutant mice."
Neurobiology of Disease 36: 312-319 (2009).

Dierssen M, Herault Y, Estivill X.
"Aneuploidy: from a Physiological Mechanism of Variance to Down Syndrome."
Physiological Reviews, 89:887-920 (2009).

Giralt A, Rodrigo T, Martin ED, González JR, Ceña V, Dierssen M, Canals JM, Alberch J.
"Brain Derived Neurotrophic Factor modulates the severity of cognitive alterations induced by mutant huntingtin through changes in phospholipaseC activity and glutamate receptor expression."
Neuroscience 158:1234-50 (2009).

Santos M, Summavielle T, Silva-Fernandes A, Teixeira-Castro A, Dierssen M, Oliveira P, Sousa N, Maciel P.
"Age- and Region-Specific Analysis of Monoaminergic Systems in the Brain of Mecp2-Null Mice Implies the Prefrontal and Motor Corticesin the Earliest Stages of Disease."
Neurology 72: A347-A347 (2009).

Other published work

Dierssen M, Azcona G, Fillat C, Toiber D, Soreq H.
"Bases moleculares de la neuropatología del síndrome de Down: implicación de DYRK1A."
In: Publicación del VI Premio «Jaime Blanco» de investigación en el síndrome de Down. Editorial Nueva Imprenta, Madrid (2009).

Dierssen M.
"Estado anímico y vídeo/memoria en condiciones extremas. La Antártida."
Arts Santa Mónica, Universidad de Barcelona, Fundació Bosch i Gimpera, Barcelona. Editorial ACTAR, in press.

Dierssen M.
"El cerebro humano: un acontecimiento evolutivo especial."
Paradigma, Universidad de Málaga (ISSN: 1885-7604), in press.

Dierssen M.
"Una visión del arte desde la neurobiología de la discapacidad intelectual. Programas de educación especial de centros de arte contemporáneo."
Centre d'Art la Panera de Lleida, in press.

Dierssen M.
"La Música: ¿un impulso básico o un elemento clave en el cerebro humano?"
Rev Médica Int Síndrome de Down 13: 17-19 (2009).

M Dierssen, M Martinez, C Fillat.
"El envejecimiento en el síndrome de Down: Dyrk1A como gen candidato para el declive cognitivo."
SD, Revista Médica sobre el Síndrome de Down (2009).

Dierssen M.
"Música en tu cerebro. En Ressonancia KRTU."
Departament de Cultura i Mitjans de Comunicació de la Generalitat de Catalunya, in press.

PATENTS

Fenome-Biotech: phenotyping start up company (X. Gallego, J. Esteras)
License of exploitation: *"Mice and rat immobilization board"* CRG-UAB-Harvard Instruments, 2009
License of exploitation: *"Aversive radial arm maze"* CRG- Harvard Instruments, 2009

BIOINFORMATICS AND GENOMICS

Coordinator: Roderic Guigó

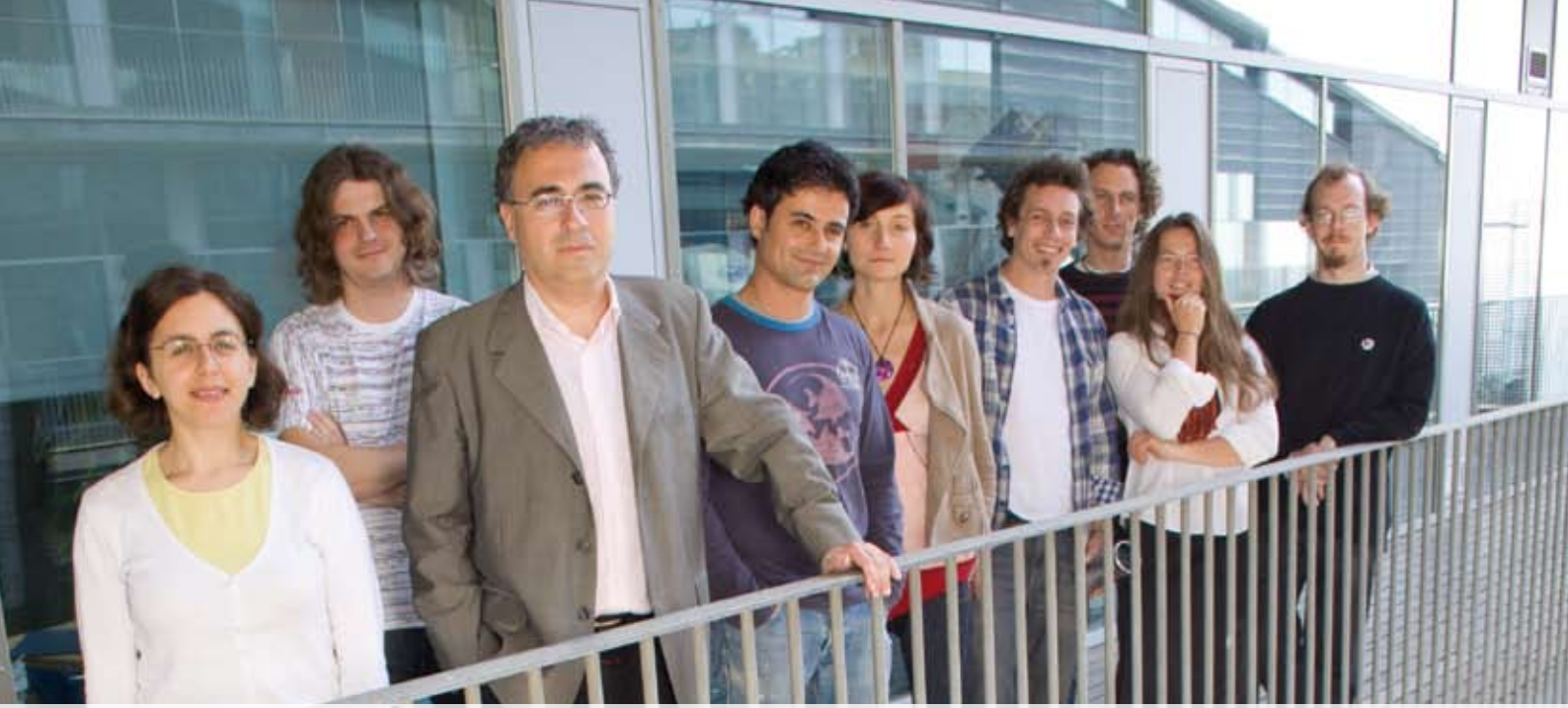
The Bioinformatics and Genomics programme includes a number of research groups, which use computational analysis in genome research. Currently, the programme includes four groups, those lead by Roderic Guigó, Cedric Notredame, Toni Gabaldón, and Fyodor Kondrashov. Dr. Gian Gaetano Tartaglia is expected to join in May 2010.

In addition to carry out their own research agenda, the groups have established a number of solid collaborations with several experimental groups from other CRG programmes. The program has also taken active leadership in the delineation of the newly established CRG genomics unit, which will provide service in ultrasequencing, microarrays, genotyping, and bioinformatics, as well as in the design and implementation of the CRG's computing scientific network.

Current structure of the programme:

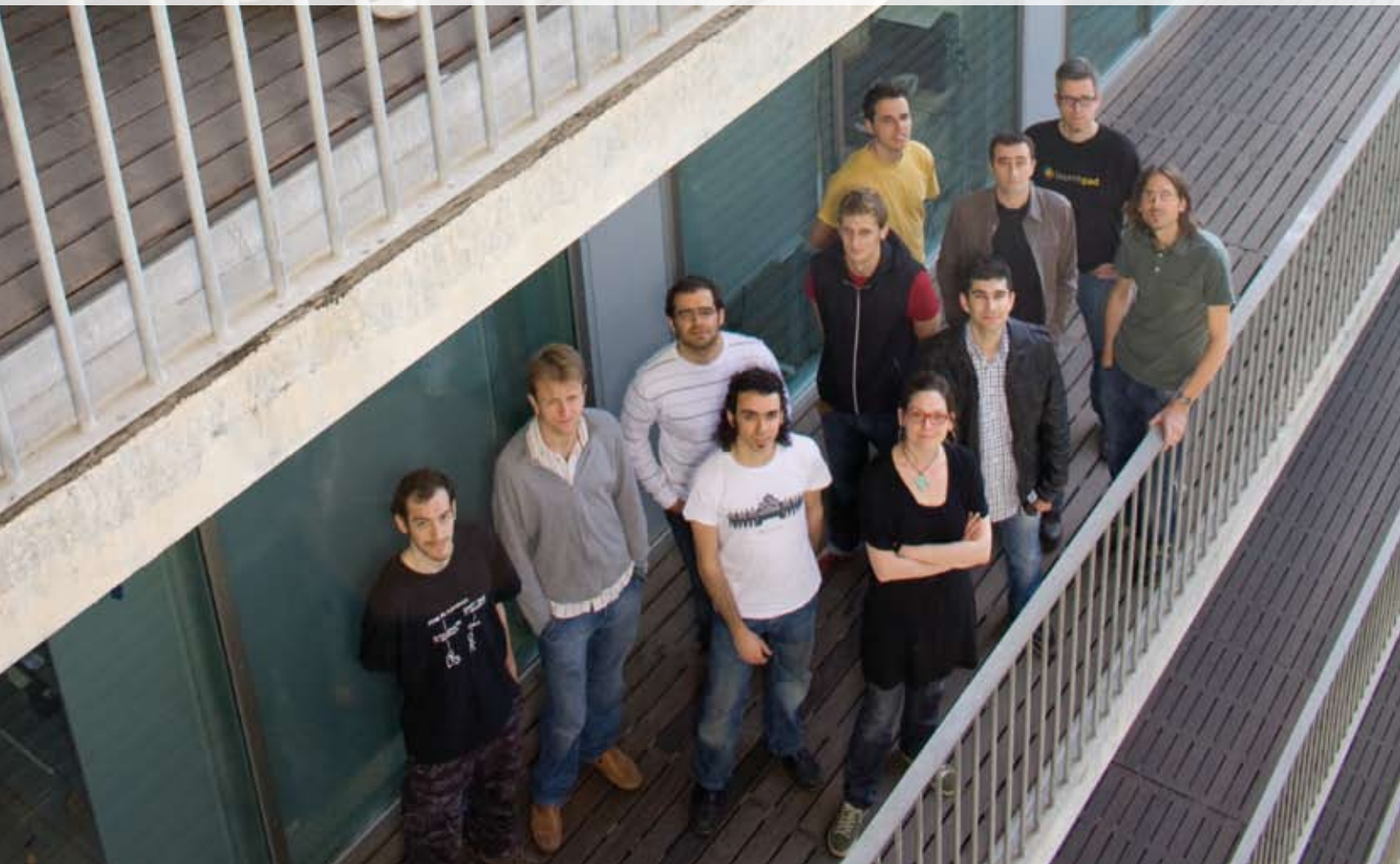
5 Research Groups:

> Genome Bioinformatics (Roderic Guigó)
> Comparative Bioinformatics (Cédric Notredame)
> Comparative Genomics (Toni Gabaldón)
> Evolutionary Genomics (Fyodor Kondrashov)
> Gene Function and Evolution (Gian Gaetano Tartaglia), starting on May 2010



BIOINFORMATICS AND GENOMICS

Group:	Bioinformatics and Genomics
Group structure:	
Group Leader:	Roderic Guigó
Postdoctoral Fellows:	Sarah Djebali, Tyler Alioto, David Martin, Christoforous Nikolau, Michael Sammeth, Pedro Gabriel, Dias Ferreira, Thomas Derrien, Paolo Ribeca, David González, Angelika Merkel
PhD Students:	Charles Chapple (IMIM), Hagen Tilgner, Anna Kedzierska, Marco Mariotti, Joao Curado
Master Students:	Alexis Grimaldi
Technicians:	Oscar González, Julien Lagarde, Francisco Câmara, Colin Kingswood, Judith Flo



SUMMARY

Research in the Genome Bioinformatics group focuses in the development and application of methods to identify functional domains in genomic sequences, with emphasis in protein coding genes and their splice variants. Our group is involved in both the development of software for gene prediction in genome sequences, and in the investigation of the signals involved in gene specification. Our group has actively participated in the analysis of many eukaryotic genomes and it in involved in NIH funded ENCODE project

RESEARCH PROJECTS

1. Characterization of long non coding RNAs

It has become recently evident that a major portion of the genome is being transcribed, and that protein-coding sequences only account for a minority of cellular transcriptional output[1,2,3]. Discovery of RNA interference (RNAi)[4] in ***Caenorhabditis elegans*** and a subsequent identification of a new class of small RNAs known as microRNAs[5,6] led to a greater appreciation of RNA's role in regulation of gene expression. Small RNAs, however, seem to represent only a minority of the non-coding transcriptome, and long noncoding RNAs are emerging as a major functional class. Cytosolic long noncoding RNAs (lnRNAs) originate both from genic and intergenic regions, have introns and poly(A)-tails, but do not contain a functional open reading frame (ORF). Until very recently, only a handful lnRNAs where known in mammals, such as those liying in the imprinting[7]and X-inactivation[8] loci.

However recent research using large scale cDNA sequencing[9] and specific chromatin signatures[10] have revealed the existence of hundreds of lnRNAs in the mouse genome. Comparative genome sequence analysis shows that lnRNAs tent to evolve slower than neighboring neutral sequences—and indication of functionality[11]. More recently, chromatin signatures have also helped to uncover about a few thousand intergenic lnRNAs (lincRNAs) in the human genome—many of which appear to affect gene expression by associating to chromatin modifying complexes[12]. The GENCODE annotation[13], produced within the framework of the ENCODE project[14] contains also a large collection of lnRNAs supported by cDNA evidence.

We have recently analyzed a set of 3,019 lnRNAs not overlapping protein-coding loci (equivalent to lincRNAs) from GENCODE. We have found that these lnRNAs also display evolutionary sequence conservation and a specific pattern of splicing distinct from that of protein-coding transcripts. A number of these RNAs cluster in well-defined gene families, and within some of which we have found a large number of compensating mutations, suggesting the presence of a common structural fold. Using RNAseq we have shown that lnRNAs display differential tissue expression, which is closely paralleled with their associated active or repressive chromatin signatures. Importantly, we found increased expression of a number of LnRNAs following differentiation of adult human stem cells. Functional analysis of these lnRNAs in multiple cell lines revealed their positive regulation of the neighboring protein-coding transcripts (Figure 1). A manuscript describing these findings is currently being considered by the journal Nature. Furher supporting the functionality of lnRNAs, we have found, after analysing RNAseq from 60 CEU individuals densily genotyped within the HapMap3 project[15], that the proportion of transcripts with eQTLs is similar in lnRNAs and protein-coding genes. In this work, we have applied the Flux Capacitor to quantify the expression of individual alternative transcript isoforms (see section 2). The paper describing these results has been accepted in the journal Nature.

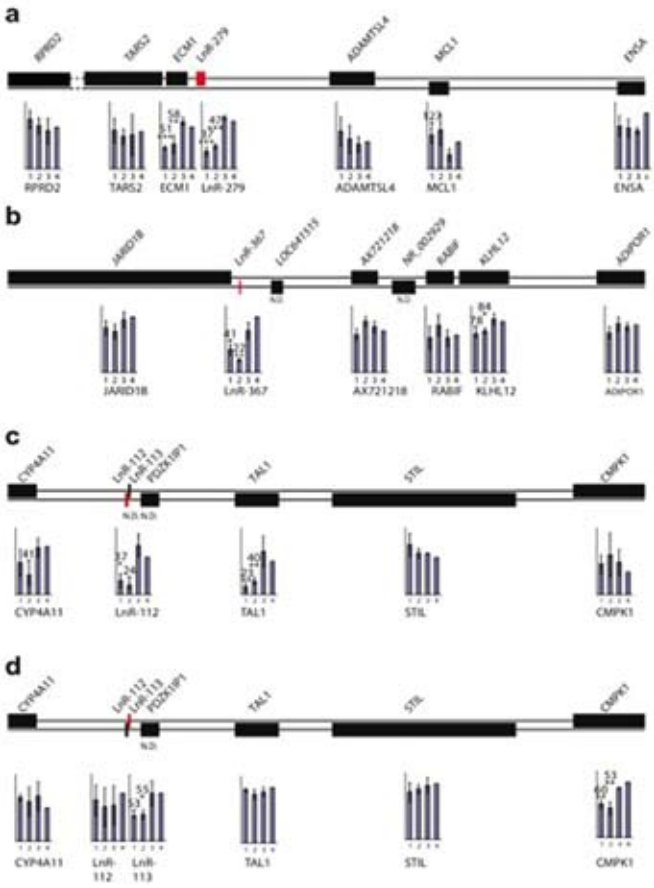


Figure 1.
*Knock-down analysis on LnRNA loci. The targeted LnRNAs are shown in red. In the bar-plots 1 and 2 represents specific siRNAs to the LnRNA in the depicted locus, 3 and 4 represents control siRNAs. a. Depletion of lnRNA LnR-279 locus in HEK293 cells results in a concomitant decrease in the expression of the neighboring EMC1 gene. Loss-of-function mutations of ECM1 results in lipoid proteinosis, a rare autosomal recessive genodermatosis24.b. Depletion of the LnR-367 locus in HeLa cells resulted in depletion of the opposite strand of the gene KLHL12 a gene known for its negative regulation of the Wnt-beta catenin pathway c. Depletion of the LnR-112 locus in MCF-7 cells resulted in a specific and potent reduction of TAL1 expression. d. Depletion of the LnR-113 locus in Jurkat cells resulted in a consistent and significant decrease in CMPK1 levels. LnR-112 and LnR-113 map within 300bp of the SCL locus, a basic helix-loop-helix protein which serves as the master regulator of hematopoiesis24. All values are relative to Gapdh expression and normalized to siRNA 4 set to 1. Error bars show standard deviation. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ by two-tailed Student's T-test.*

2. Transcriptome reconstruction from RNAseq experiments

Further enhancing our capacity to interrogate the transcriptional activity of the human genome is the recent development of sequence instruments of the so-called Next Generation (NGS). Indeed, these instruments provide the sufficient sequencing depth, in principle, to fully characterize the transcriptome of the cells, which is to identify all transcript species and to quantify their relative abundance. However, the most cost-effective such technologies typically produce very short sequence reads, which compounds transcript reconstruction and quantification. Indeed, we have recently shown [16] that in common scenarios neither single nor paired-end read sequences guarantee theoretically that all transcript variants and their abundances can be solved. We formulated the problems of transcript reconstruction as a system of linear equations in which the abundance of the sequence reads are simply the sum of the abundances of the transcript variants from which they originate. We demonstrated that even for very simple transcriptomes, and assuming infinite coverage and uniform read distribution, the system of linear equations is under-determined, and that it has, therefore, no unique solution. We have also shown, however, that in the practice paired-end reads can solve the majority of the transcript variants, and that single reads can be effectively used to efficiently discriminate between alternative splicing abundances, provided that an underlying transcript annotation exists. We have thus started to work in the development of computational tools to reconstruct and quantify transcript sequences from short sequence reads (see Research Design and Methods) and as mentioned above we are also currently collaborating with the ETD group in a population variation analysis of RNAseq. Methods like this, together with technological advances that result in increasing read length and throughput, promise that full characterization of the RNA complement of the cell will soon be possible. A number of groups worldwide are working on the development of methods to reconstruct transcript sequences and infer transcript abundances from massively parallel RNA sequencing (RNAseq) experiments, and an assessment project has been launched (*http://www.sanger.ac.uk/PostGenomics/encode/RGASP.html*) to evaluate current progress towards that end. Assessment of RNAseq reconstruction methods, however, is seriously compounded by the lack of an experimentally verified gold standard, that is, of a complex RNA population which can be easily replicated and in which

the abundance of all transcripts, including alternative spliceforms within the same gene, have been carefully calibrated. To address this shortcoming we have been working on the development of an RNAseq simulator, which we call the FluxSimulator, a program that given a transcriptome, that is, a set of RNA sequences, and attached abundances, simulates the different steps of an RNAseq experiment and produces a set of short reads originating from the transcript sequences. In this way, we can evaluate how different methods and strategies perform when reconstructing the original transcript population from the set of simulated reads.

The FluxSimulator contains two modules. The first module is a transcriptome simulator. It takes as input a reference annotation for the genome—that is a set of genes and transcripts—and assigns to each transcript a expression value y according to the formula:

where x is the expression rank of the $y = \left(\frac{x}{x_0}\right)^{k} \exp^{-\left(\frac{x-x_0}{x_0}\right)^{2}}$ transcript, x_0 the expression level of the highest abundant transcript, and k is a cell type specific parameter close to -1. Our analysis of published transcriptomic data shows that such a profile exhibits a linear log-log shape for highly expressed genes, which becomes exponential decay towards the spectrum corresponding to low abundant transcripts. This is represented by factor x^1, that sumarizes the turnover of rare RNA species, cell population size and sequencing depth of the conducted RNASeq experiment (Figure 4, top-left).

The actual RNAseq simulator module takes as an input a real or simulated transcriptome, that is, a set of RNA sequences with attached abundances, and simulates the different steps of the RNAseq experiment: the reverse transription (RT) of the RNA to generate a cDNA library, the fragmentation of the DNA, and the sequencing itself. Each of these steps can be carried out using different protocols leading to alternative strategies to perform the RNA experiments, each of which has its own inherent biases. Our goal is for the FluxSimulator to be able to simulate multiple protocols and strategies. A few of these protocols (poly-dt vs random priming for the reverse transcription; nebulization vs cleavage for fragmentation, paired- vs single- reads, stranded vs unstranded for sequencing) have already been implemented. In Figure 2, we display the results of the simulation of a number of real RNAseq experiments, and the comparison of our simulations with the experimental results. In general, there is a good agreement between experimental and simulated data—at least when comparing the bias in the read distribution along the transcript. Figure 3 shows a prototype for the interface of the Flux Simulator.

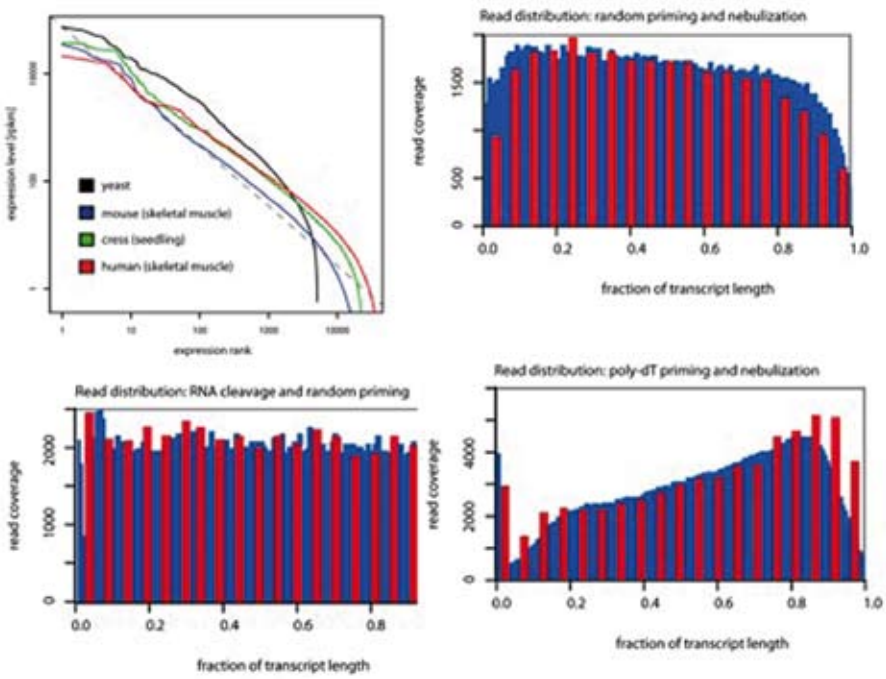


Figure 2.
Simulation of transcriptomes and RNAseq experiments.
Top Left: *Expression profiles obtained from RNASeq experiments in different species: Homo sapiens [17], Mus musculus [18], Arabidopsis thaliana [19] and Saccharomyces cerevisiae Reads have been mapped to reference annotations for each of these species: RefSeq [20] for human and mouse, TAIR [21] for A. thaliana, and SGD [22] for yeast. Normalized expression values (rpkm) have been ranked and they are given for every rank percentile.* ***Rest of the panels:*** *Distribution of real and simulated reads (in blue) along the transcripts in different RNAseq experiments. The X-axis is the normalizaed transcript length, and the Y-axis normalized read density. Transcriptomes have been simulated using as input the reference annotations mentioned above, and the RNAseq has been simulated according to the experimental protocol employed. The distribution of the corresponding experimental data is shown by superimposed red bars.* ***Top Right:*** *RT with Random Priming followed by Nebulization in human (unpublished data);* ***Bottom left:*** *Chemical cleavage followed by RT with random priming in mouse (Mortazavi et al, 2008);* ***Bottom Right:*** *Poly-dT Priming followed by Nebulization in A.thaliana [19].*

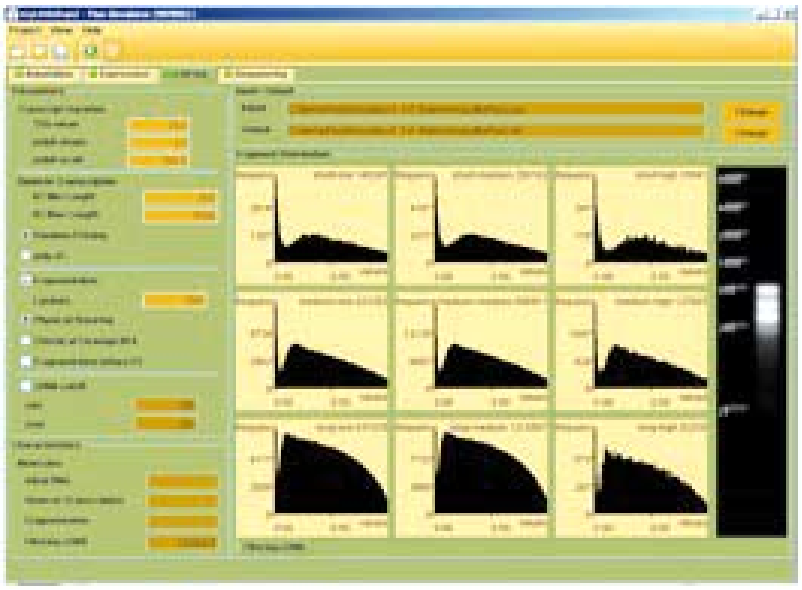


Figure 3. ***Graphical user interface of the FluxSimulator.***
The simulation pipeline comprises different steps. The histograms in the screenshot show prospective read distributions along transcripts, broken down by different transcript lengths and expression levels. A simulated northern blot summarizes the length distribution of obtained cDNA fragments.

2.1 Reconstruction and Quantification of RNAseq data

We are developing a software, the FluxCapacitor, to reconstruct transcript abundances from RNAseq data. The current version of the software assumes an underlying annotation, and therefore it is only able to attach abundances to previously known transcripts. We are working on a new version of the FluxCapacitor so that it is able 1) to assemble "de novo" transcripts from the RNAseq data, and attach abundances to all, annotated and novel, transcripts, and 2) attach allelic origin to the quantified transcripts.
Currently the FluxCapacitor takes as input an annotation on a reference genome and a set of sequence reads produced through an RNAseq experiment and mapped into the genome. In the first step the transcript annotation corresponding to each locus is converted into a special form of a splicing graph (Figure 6, top and middle), onto which sequence reads are mapped. Edges in the graph represent (parts of) exons and introns, and reads that map to an edge correspond to the reads originating from the different transcripts including that edge. In our approach to reconstruct transcript abundances, the number of reads mapping to each edge is the sum of the abundances of the transcripts including that edge plus some experimental noise d_i (equation in Figure 3, top). The collection of all those edge-constraints across the splicing graph results in a system of linear equations, which is subsequently solved using linear programming, with the objective function of minimizing the sum of all edge-noises d_i. Transcript expression levels are derived directly from the solution of the system. Because experimental biases, reads are not uniformely distributed across the transcript length (Figure 2). Therefore, the expected number of reads originating from a given transcript mapping to an edge does not depend only on the size of the edge, but on the relative position of the edge along the transcript, and on the length of the transcript (Figure 4, bottom). As a general approach, we build expression profiles directly from the input read data, which serve as templates for the estimation of expression fractions.

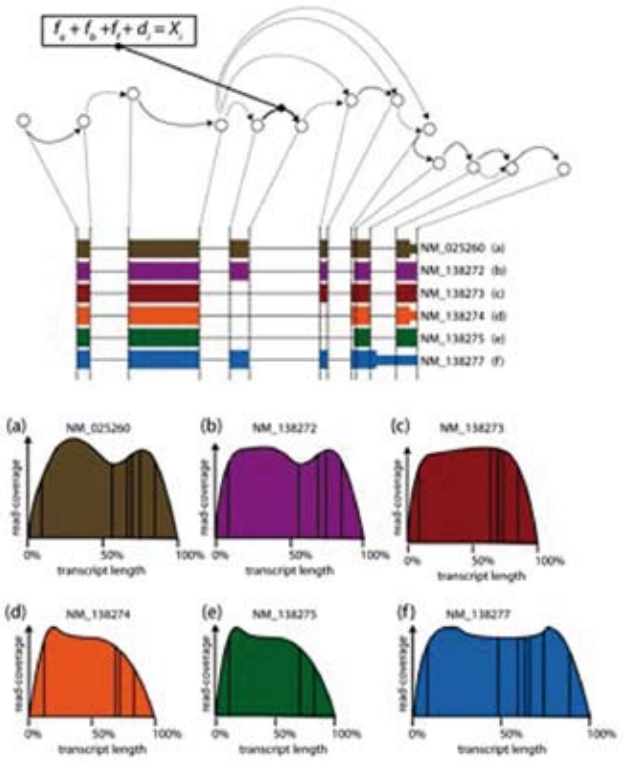


Figure 4. ***The FluxCapacitor.*** *Splicing graph representing the alternative splice forms corresponding to one locus. Edges correspond to unique exon segments (solid arrows) and introns (dashed arrows). For each such Xi the number of reads mapping it is the sum of of the expected contributions f_x (i.e., "flows") of the transcripts including the edge, that is the sum of the transcript's expression, the unknowns, plus some error or noise d_j. Because of technical biases, reads are not distributed uniformly across the genome (bottom of the figure). The expected contribution from each transcript to each edge depends on the length and the position of the edge along the transcript. The collection of linear constraints along all exonic edges of the labeled splice graph (i.e., flow network) forms a system that is solved by linear programming, minimizing the sum of all edge noise levels d_j.*

3. Nucleosome positioning

The search of sequence prerequisites for nucleosome positioning has been a long-standing problem at the intersection of chromatin structure and gene regulation. With the use of structural information directly related to the primary DNA sequence we propose a new property of natural nucleosome forming sequences, which is inherently related to their intrinsic curvature and its symmetry. A measure of this property has been introduced and a corresponding method validated against novel high quality datasets of human nucleosomes, obtained through close collaboration with CRG's Chromatin and Gene Expression Group (Miguel Beato). Based on this symmetrical curvature property, we have implemented a computational ab initio method for nucleosome positioning prediction and have shown this method to be more efficient than previously published ab initio methods aiming at the same goal. The aforementioned method is already being applied in a number of analyses related to nucleosomal patterns of regions of interest in various organisms. The results have contributed to the formulation of the "key nucleosome" concept.

4. Nucleosome positioning and splicing

We have found evidence that strongly suggests a direct involvement of chromatin structure in RNA splicing. We have analyzed high throughput sequencing data recently produced by Schones et al. (2008) on positioning of nucleosomes in resting and activated human CD4+ T cells. Our analyses show evidence for stable nucleosome occupancy within internal exons of human genes, a pattern absent in pseudoexons (non-repetitive intronic sequences flanked by strong splice sites), which in contrast exhibit a weak nucleosome depletion (Figure 3). Remarkably, this pattern depends on the strength of the splice sites. Indeed, the nucleosome peak is accentuated within exons with weak splice sites, while in the exons with strong splice sites, as well as in pseudoexons, a region of stable nucleosome occupancy is observed upstream of the acceptor sites. As a result, pseudoexons with strong splice sites—in which splicing does not occur despite the strength of the sites—show a pattern of nucleosome occupancy which is the mirror image of that observed on exons with weak sites—in which splicing occurs despite the weakness of the sites (Figure 5). These observations strongly suggest that positioning of nucleosomes influences RNA splicing.

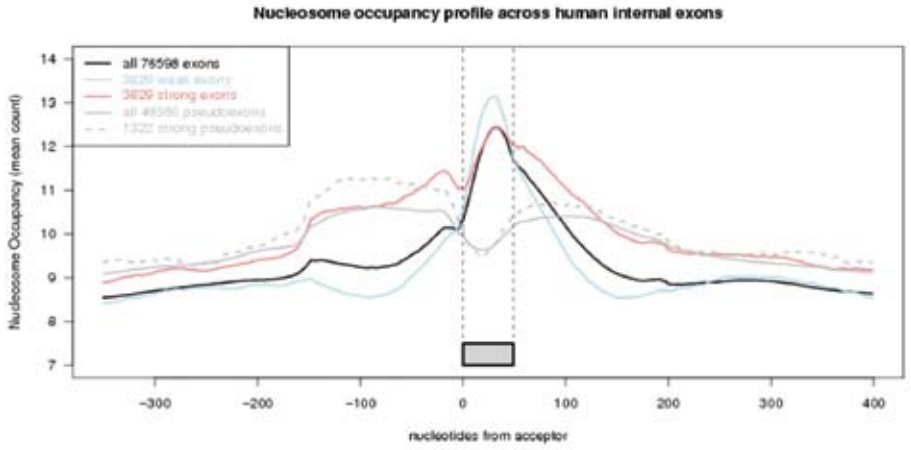


Figure 5. ***Nucleosome occupancy profile across internal human exons in resting CD4+ T cells.***
Upstream and Downstream from the idealized exon (3' acceptor site at position 0), we plot the average number of nucleosome reads per nucleotide position. Within the exon, we have binned nucleosome reads in 50 intervals, irrespective of the length of the exon, and we plot the average number of reads in a 6bp window centered in the midpoint of the interval. In black, the profile for all internal exons, in light read, for the subset of exons with combined donor and acceptor score among the top 5%, in light blue, for the subset of exons with combined score among the bottom 5%, in solid grey for a set of intronic pseudoexons, and in dashed grey for the subset of pseudoexons with a combined score higher than the 90% percentile of real exons.

5. ENCODE Project

We are part of the NHGRI funded ENCODE project. During the pilot phase of ENCODE, we lead one of the research groups. In the scaling phase of ENCODE we participate in the grants lead by Tom Gingeras (from Cold Spring Harbour Laboratory), where we are responsible for the bioinformatics, and by Tim Hubbard (from the Sanger Institute), where we contribute to build the reference annotation of the human genome. We are also leading the RNA analysis working group. A manuscript describing the results after the first phase of the scaling project will be ready by the end of this year.

Recently, we have started a research line on the development of methods to infer alternative transcript abundance from the short reads produced by the Next Generation Sequencing Instruments (Lacroix et al., 2008)

REFERENCES

1 Birney E et al.
"Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project."
Nature 447:799-816 (2007).

2 Cheng J et al.
"Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution."
Science 308:1149-1154 (2005).

3 Kapranov P et al.
"RNA maps reveal new RNA classes and a possible function for pervasive transcription."
Science 316:1484-1488 (2007).

4 Fire A et al.
"Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans."
Nature 391:806-811 (1998).

5 Lee RC, Feinbaum RL and Ambros V.
"The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14."
Cell 75:843-854 (1993).

6 Wightman B, Ha I and Ruvkun G.
"Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans."
Cell 75:855-862 (1993).

7 Yang PK and Kuroda MI.
"Noncoding RNAs and intranuclear positioning in monoallelic gene expression."
Cell 128:777-786 (2007).

8 Heard E and Disteche CM.
"Dosage compensation in mammals: fine-tuning the expression of the X chromosome."
Genes Dev 20:1848-1867 (2006).

9 Carninci P et al.
"The transcriptional landscape of the mammalian genome."
Science 309:1559-1563 (2005).

10 Guttman M et al.
"Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals."
Nature 458:223-227 (2009).

11 Marques AC and Ponting CP.
"Catalogues of mammalian long noncoding RNAs: modest conservation and incompleteness."
Genome Biol 10:R124 (2009).

12 Khalil AM et al.
"Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression."
Proc Natl Acad Sci U S A 106:11667-11672 (2009).

13 Harrow J et al.
"GENCODE: producing a reference annotation for ENCODE."
Genome Biol 7 Suppl 1:S4 1-9 (2006).

14 Birney E et al.
"Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project."
Nature 447:799-816 (2007).

15 Frazer KA et al.
"A second generation human haplotype map of over 3.1 million SNPs."
Nature 449:851-861 (2007).

16 Lacroix V, Sammeth M, Guigo R and Bergeron A.
"Exact transcriptome reconstruction from short sequence reads."
Proceedings of WABI'08, Lecture Notes in Bioinformatics 5251:50-63 (2008).

17 Wang ET et al.
"Alternative isoform regulation in human tissue transcriptomes."
Nature 456:470-476 (2008).

18 Mortazavi A, Williams BA, McCue K, Schaeffer L and Wold B. *"Mapping and quantifying mammalian transcriptomes by RNA-Seq."*
Nat Methods 5:621-628 (2008).

19 Weber AP, Weber KL, Carr K, Wilkerson C and Ohlrogge JB.
"Sampling the Arabidopsis transcriptome with massively parallel pyrosequencing."
Plant Physiol 144:32-42 (2007).

20 Clark AG et al.
"Evolution of genes and genomes on the Drosophila phylogeny."
Nature 450:203-218 (2007).

21 Swarbreck D et al.
"The Arabidopsis Information Resource (TAIR): gene structure and function annotation."
Nucleic Acids Res 36:D1009-1014 (2008).

22 Christie KR et al.
"Saccharomyces Genome Database (SGD) provides tools to identify and analyze sequences from Saccharomyces cerevisiae and related sequences from other organisms."
Nucleic Acids Res 32:D311-314 (2004).

PUBLICATIONS

Koscielny G, Texier VL, Gopalakrishnan C, Kumanduri V, Riethoven JJ, Nardone F, Stanley E, Fallsehr C, Hofmann O, Kull M, Harrington E, Boué S, Eyras E, Plass M, Lopez F, Ritchie W, Moucadel V, Ara T, Pospisil H, Herrmann A, Reich JG, Guigó R, Bork P, Doeberitz MV, Vilo J, Hide W, Apweiler R, Thanaraj TA and Gautheret D.
"ASTD: The Alternative Splicing and Transcript Diversity database."
Genomics Mar 93(3):213-20 (2009).

Chapple CE, Guigó R, Krol A.
"SECISaln, a web-based tool for the creation of structure-based alignments of eukaryotic SECIS elements."
Bioinformatics 25(5):674-5 (2009).

Boj SF, Servitja JM, Martin D, Rios M, Talianidis I, Guigo R, Ferrer J.
"Functional targets of the monogenic diabetes transcription factors HNF-1alpha and HNF-4alpha are highly conserved between mice and humans."
Diabetes 58(5):1245-53 (2009).

Takeuchi A, Schmitt D, Chapple C, Babaylova E, Karpova G, Guigo R, Krol A, Allmang C.
"A short motif in Drosophila SECIS Binding Protein 2 provides differential binding affinity to SECIS RNA hairpins."
Nucleic Acids Res 37(7):2126-41 (2009).

Harrow J, Nagy A, Reymond A, Alioto T, Patthy L, Antonarakis SE, Guigó R.
"Identifying protein-coding genes in genomic sequences."
Genome Biol 10(1):201. Review. (2009).

Servitja JM, Pignatelli M, Maestro MA, Cardalda C, Boj SF, Lozano J, Blanco E, Lafuente A, McCarthy MI, Sumoy L, Guigó R, Ferrer J.
"Hnf1alpha (MODY3) controls tissue-specific transcriptional programs and exerts opposed effects on cell growth in pancreatic islets and liver."
Mol Cell Biol 29(11):2945-59. (2009).

Pignatelli M, Serras F, Moya A, Guigó R, Corominas M.
"CROC: Finding Chromosomal Clusters in Eukaryotic Genomes."
Bioinformatics 25(12):1552-3. Epub 2009 Apr 23. (2009).

Bovine Genome Sequencing and Analysis Consortium (including Guigó R).
"The genome sequence of taurine cattle: a window to ruminant biology and evolution."
Science 324(5926):522-8. (2009).

Arman M, Aguilera-Montilla N, Mas V, Puig-Kröger A, Pignatelli M, Guigó R, Corbí AL, Lozano F.
"The human CD6 gene is transcriptionally regulated by RUNX and Ets transcription factors in T cells."
Mol Immunol 46(11-12):2226-35. Epub 2009 May 14. (2009).

Castellano S, Andrés AM, Bosch E, Bayes M, Guigó R, Clark AG.
"Low exchangeability of selenocysteine, the 21st amino acid, in vertebrate proteins."
Mol Biol Evol 26(9):2031-40. Epub 2009 Jun 1. (2009).

Sammeth M.
"Complete alternative splicing events are bubbles in splicing graphs."
J Comput Biol 16(8):1117-40. (2009).

Makrythanasis P, Kapranov P, Bartoloni L, Reymond A, Deutsch S, Guigó R, Denoeud F, Drenkow J, Rossier C, Ariani F, Capra V, Excoffier L, Renieri A, Gingeras TR, Antonarakis SE.
"Variation in novel exons (RACEfrags) of the MECP2 gene in Rett syndrome patients and controls."
Hum Mutat 30(9):E866-79 (2009).

Tilgner H, Nikolaou C, Althammer S, Sammeth M, Beato M, Valcárcel J, Guigó R.
"Nucleosome positioning as a determinant of exon recognition."
Nat Struct Mol Biol 16(9):996-1001 (2009).

Buschbeck M, Uribesalgo I, Wibowo I, Rué P, Martin D, Gutierrez A, Morey L, Guigó R, López-Schier H, Di Croce L.
"The histone variant macroH2A is an epigenetic regulator of key developmental genes."
Nat Struct Mol Biol 16(10):1074-9. Epub 2009 Sep 6 (2009).

Horner DS, Pavesi G, Castrignanò T, De Meo PD, Liuni S, Sammeth M, Picardi E, Pesole G.
"Bioinformatics approaches for genomics and post genomics applications of next-generation sequencing".
Brief Bioinform. 2009 Oct 27. [Epub ahead of print]

BIOINFORMATICS AND GENOMICS

Group: Comparative Bioinformatics

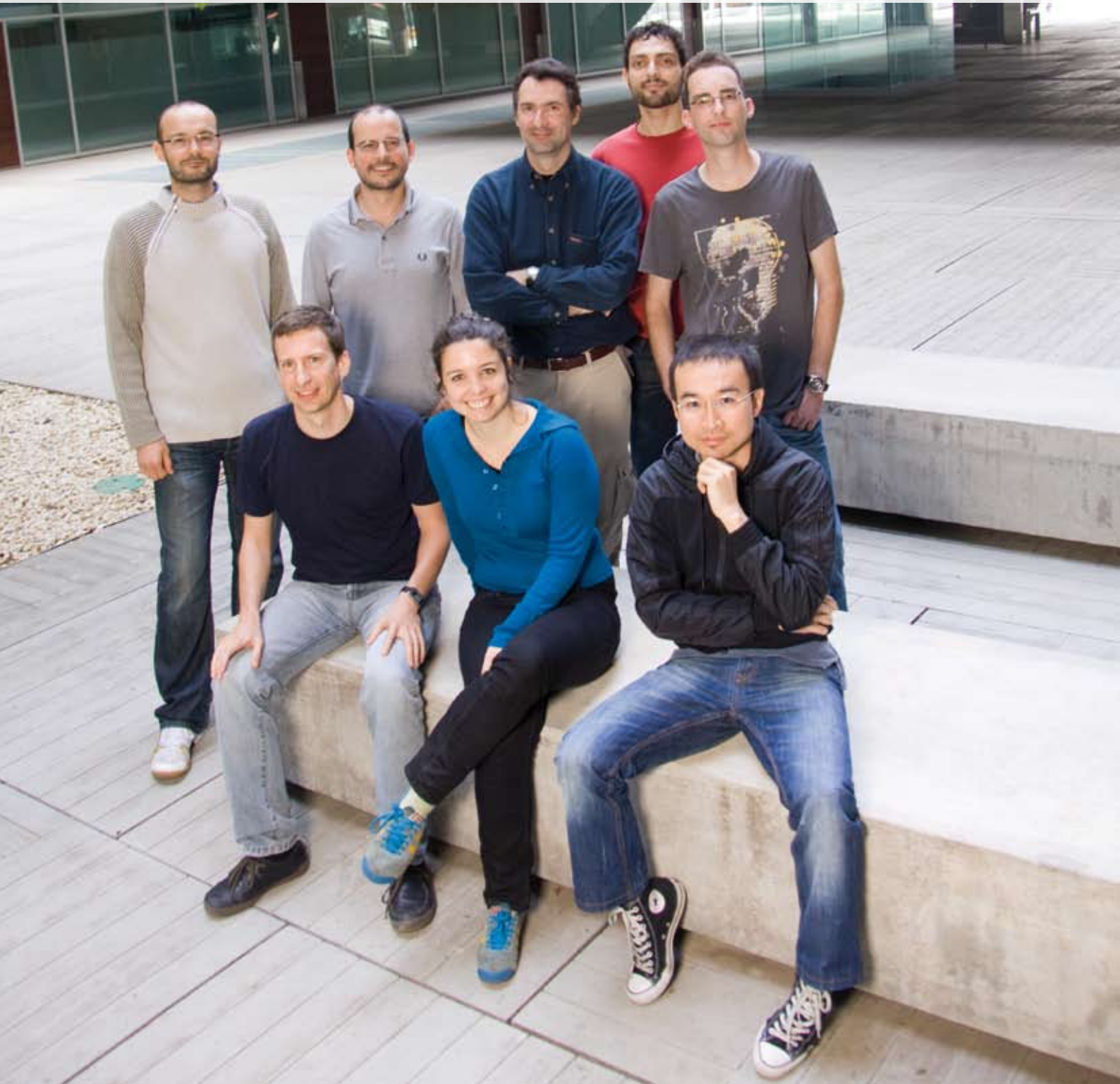
Group structure:

Group Leader: Cédric Notredame

Postdoctoral Fellows: Ionas Erb, Emmanuele Raineri, Jean-François Taly, Thomas Lingner

Students: Jia-Ming Chang, Giovanni Bussoti, Carten Kemena, Meritxell Oliva, Isabel Fernandez

Technician: Paolo di Tomasso



SUMMARY

The main focus of the group is the development of novel algorithms for the comparison of multiple biological sequences. Multiple comparisons have the advantage of precisely revealing evolutionary traces, thus allowing the identification of functional constraints imposed on the evolution of biological entities. Most comparisons are currently carried out on the basis of sequence similarity. Our goal is to extend this scope by allowing comparisons based on any relevant biological signal such as sequence homology, structural similarity, genomic structure, functional similarity and more generally any signal that may be identified within biological sequences. Using such heterogeneous signals serves two complementary purposes: (i) producing better models that take advantage of the evolutionary resilience, (ii) improving our understanding of the evolutionary processes that leads to the diversification of biological features. We develop these novel methods in close collaboration with experimental groups and make them available through an international network of web severs: www.tcoffee.org. In addition to the CRG and the Catalan government, the group is supported by the Plan Nacional, La Caixa and by two international FP7 consortiums: Quantomics (dedicated to the survey of genomic variation in farm animals) and Leishdrug, a project dedicated to the development of a new class of drugs targeting kinases in Leishmania Major.

RESEARCH PROJECTS

1. Integration of Multiple Sequence and Multiple Genome Alignments

Our lab has pioneered the development of a new generation of multiple sequence alignment methods, based on consistency. We are working actively on the development of a new complementary concept: template based sequence alignments. The principle is relatively new and has marked an important shift in the field of multiple sequence alignments. While most methods tend to merely exploit the information contained in the provided sequences, template based methods couple the multiple sequence alignment with a either a database search or a structural prediction that enriches the information content of the sequence. After focusing on the combination of multiple protein structures and multiple sequences protein sequences, we have now started exploring the alignment of structured non coding RNAs. For this purpose we have developed a novel approach able to incorporate RNA secondary structural prediction within the computation of a multiple sequence alignment. The resulting algorithm (R-Coffee) is one of the most accurate RNA multiple sequence aligner currently available (as judged by benchmarking on BraliBase). R-Coffee is the result of a collaboration with the lab of Des Higgins and the Swiss Institute of Bioinformatics. Building up on this framework, we have now added two new types of templates: homology based templates, where templates are replaced with a profile automatically built and promoter alignments where promoter sequences are aligned by taking into account their overlap with known binding sites. The long term goal of this approach is to bridge the methodological gap that exists between genome alignments (based on nucleotides) and model based sequence alignments. This project is supported by the Plan Nacional (G-Coffee project) and is carried out in collaboration with the European Institute of Bioinformatics. We also have a very active collaboration with Knut Reinert (Berlin Free University) and with the computational department of Lleida University who are helping us improving the underlying algorithms and deploying our algorithms onto new parallel hardware.

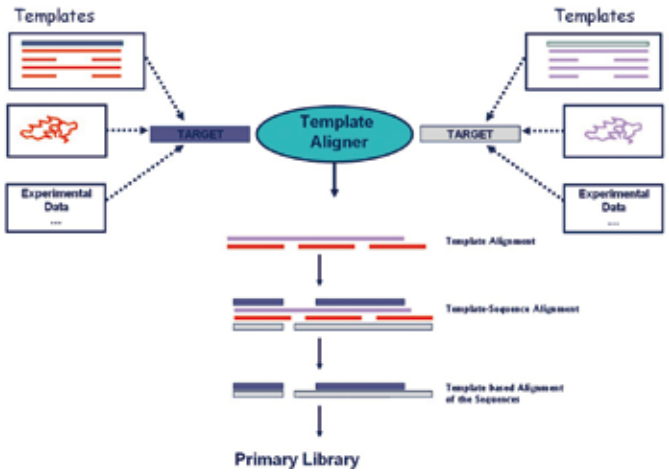


Figure 1.
Principal of the Template based sequence alignments in T-Coffee

2. Phenotypic and Genomic Comparisons

One of the goals of genomics is to survey variability within species and identify which element of this genomic variability can be related to variations measured at the phenotypic level. Achieving such an aim supposes that one is able to quantitatively compare genomic and phenotypic variability across individuals. Yet, reliable methods able to achieve these two goals still need to be developed. Through various projects and collaborations, our lab is addressing this problem at all levels. At the genomic level, we are developing novel methods and approaches in order to obtain better multiple genome alignments. One of our approaches involves exploring how multiple genome alignments can be evaluated with respect to their biological correctness. In collaboration with the European Institute of Bioinformatics, we are also developing a new generation of multiple genome aligners, designed to handle very large datasets. The multiple genome alignments thus obtained are mostly meant to be used for surveying variability across species and understand how this evolutionary variability relates to phenotypic variability within populations. We are applying this methodology within the context of the Quantomics project where our goal is to produce multiple genome alignments of farm species and project onto these alignments the information acquired through genotyping, while the inter species comparison is meant to provide functional clues and allow the identification of the genetic variation more likely to be associated with a phenotype. We are also using RNAseq data in order to understand how genomic variations may affect gene expression (that may be viewed as a low level phenotype). Phenotype analysis (i.e. recording and comparison of phenotypes) are rapidly becoming one of the bottleneck of genomics. In order to address this important question we collaborate with the lab of Mar Diersen (CRG), developing new methods, based on Hidden Markov Modelling for the analysis of longitudinal recording in mice. Longitudinal recording are methodologies that allow the recording of a given parameter over time. In this context our goal is to seek new ways of comparing these recordings. For that purpose we are using a methodology that was originally developed in collaboration with the department of sociology in Lausanne University and that involves Hidden Markov Modelling.

3. Reconstruction of large gene families

While it has dramatically empowered the scientific with new knowledge, high throuput sequencing also create many yet unsolved problems, the most obvious one being the reconstruction of large protein families across collection of low quality genomes. This type of analysis constitute a straightforward application of the various methodologies developed within our lab. We currently have two lines of research building on this expertise. In the context of the Leishdrug consortium and in collaboration with the lab of Gerald Spaeth (Pasteur Institute, Paris) we are currently analyzing the genome of various Trypanosomidae in order to establish the complete kinome of these species, the aim being the identification of new pharmacological targets among the protein kinases contained in these organisms. Using a similar methodology we are also analyzing the genomes of all sequenced drosophilaes in order to establish an exhaustive repertory of the genes involved in odour recognition (collaboration with the lab of Mathieu Louis, CRG). These analysis involve developing a full range of new tools making it possible to correct all possible issues associated with gene annotation, including genome mis-assembly, gene mis-prediction and the mis-identification of remote homologues.

4. Multiple Sequence Alignment Servers

Thanks to a partnership with the Swiss Institute of Bioinformatics, we are currently running a very powerful multiple alignment server mirrored by several international institutions (EBI, Cornell, CNRS, SIB). Altogether all the mirrors receive close to 20.000 hits a month originating from an average of 90 countries. In 2009 the various publications describing either the server of the original algorithms have received more than 450 citations. The latest addition to this popular service has been the R-Coffee alignment mode. We are now working on an upgrade of this server. An effort to upgrade this service and make it available via a computer cloud is now under way.

Figure 2.
The T-Coffee web server (http://www.tcoffee.org)

PUBLICATIONS

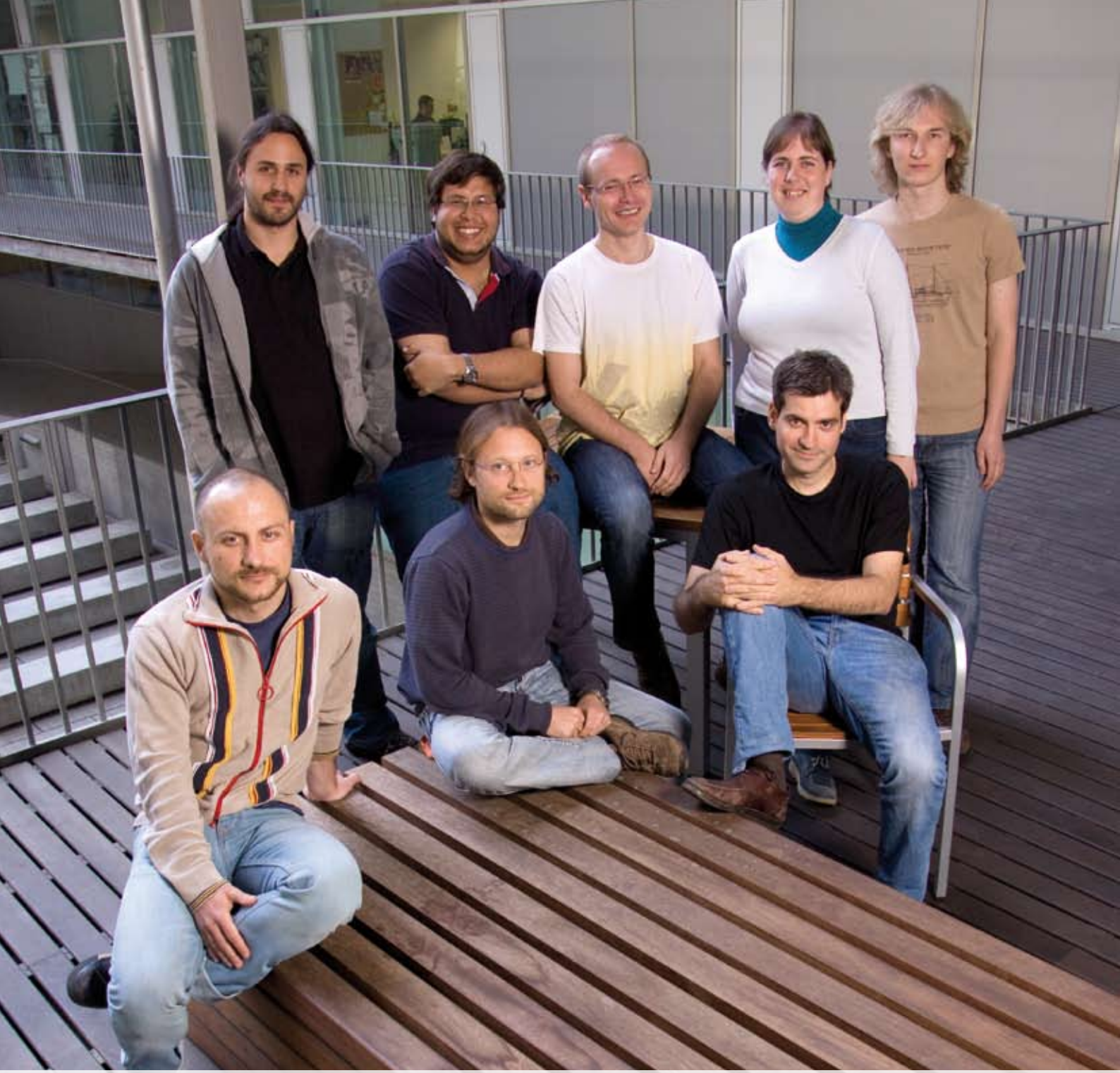
Varas F, Stadtfeld M, de Andres-Aguayo L, Maherali N, di Tullio A, Pantano L, Notredame C, Hochedlinger K, and Graf T.
"Fibroblast-derived induced pluripotent stem cells show no common retroviral vector insertions."
Stem Cells 27(2):300-6 (2009).

Orobitg M, Guirardo F, Notredame C and Cores F.
"Exploiting parallelism on progressive alignment methods."
Journal of Supercomputing 1(1):1-9 (2009).

Kemena C and Notredame C.
"Upcoming challenges for multiple sequence alignment methods in the high-throughput era."
Bioinformatics 25(19):2455-65 (2009).

Gauthier JA, Widmer ED, Bucher P and Notredame C.
"How much does it cost? Optimization of costs in sequence analysis of social science data."
Sociological Methods and Research 38(1):197-231 (2009).

Coll O, Villalba A, Bussotti G, Notredame C and Gebauer F.
"A novel, noncanonical mechanism of cytoplasmic polyadenylation operates in Drosophila embryogenesis."
Genes Dev, in press.



BIOINFORMATICS AND GENOMICS

Group: Comparative Genomics

Group structure:

Group Leader: Toni Gabaldón

Postdoctoral Fellows: Jaime Huerta-Cepas

Students: Salvador Capella-Gutiérrez, Leszek Pryscsz

Technicians: Marina Marcet-Houben, Diego Kormes, Javier Díez-Pérez

SUMMARY

Our research interests are focused around the use of comparative genomics and phylogenomics to study the origin, evolution and function of complex biological systems. This includes understanding how specific biochemical pathways, protein complexes or cellular organelles emerged and evolved as well as using this evolutionary information to gain insight into their function. Through collaborations with experimental groups we apply comparative genomics to discover new mechanisms and genes involved in interesting processes, especially those of clinical relevance (see lines of research). On the technical side, our work often involves the development of new bioinformatics tools and algorithms that we make available to the community. You can access more info at *http://www.crg.es/toni_gabaldon* and *http://gabaldonlab.crg.es*

RESEARCH PROJECTS

1. Discovery of new genes involved in mitochondrial disease

Mitochondria play a central role in the cellular metabolism and the impairment of many mitochondrial proteins leads to disease. The list of such diseases is continuously growing and includes Parkinson´s, Alzheimer´s and Huntington´s diseases. Despite recent advances, the molecular basis of many mitochondrial diseases is yet to be understood and there is growing need to identify disease-causing genes and to unravel their functions. Recently, proteomics analyses have identified a large set of proteins that function inside the mitochondrion, the so-called mitochondrial proteome. Most of these proteins are not functionally characterized and it is expected that many of them may be involved in mitochondrial diseases. Moreover, other mitochondrial proteomic sets are being characterized in different species facilitating an evolutionary analysis of the mitochondrial system. The aim of this research line is to integrate different types of data and automatically combine them in order to facilitate the identification and functional characterization of mitochondrial disease-related genes. This year we discovered that a mutation in a non-coding region of the BCS1L gene was responsible for Complex III deficiency.

2. Comparative genomics of fungal pathogens

Fungal infections constitute an ever-growing and significant medical problem. Diseases caused by such pathogens range from simple toe nail infections, to life-threatening systemic mycoses in patients with impaired immune systems. The molecular mechanisms driving invasion of mammalian hosts by fungal pathogens poses many scientifically challenging problems, which are as yet little understood. *Candida glabrata* is an emerging pathogen that represents the second most common cause of fungal infection. *C. glabrata* is evolutionarily closer to non-pathogenic Saccharomyces species than to other pathogen *Candida* species and uses different virulence mechanisms. Therefore, the problem of elucidating the mechanism for pathogenesis of *Candida glabrata*, as proposed here, can be approached with an evolutionary perspective by detecting specific adaptations in the *C. glabrata* genome that may explain why, in contrast to its closest relatives, is able to infect the human host. The present research line will be developed in the context of a international consortium (FunPath). developed in the context of a international consortium (FunPath).

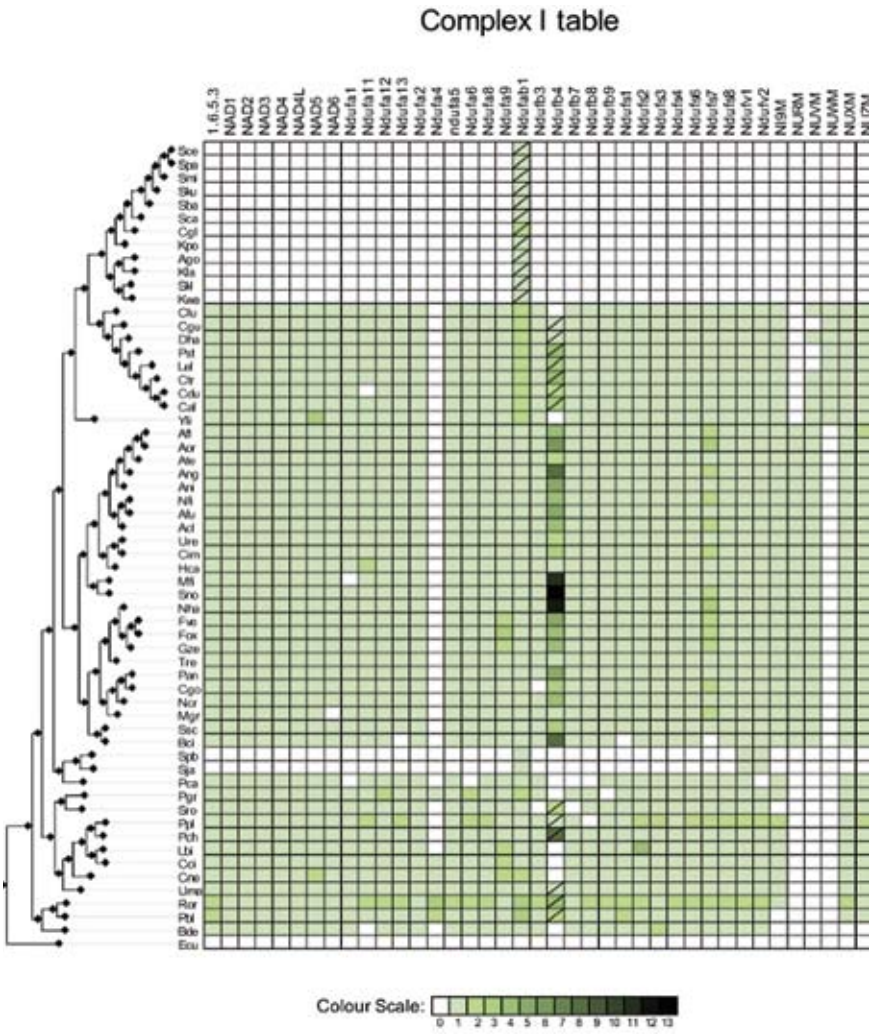


Figure 1.
Phylogenetic profile of subunits of mitochondrial Complex I across 60 completely-sequenced fungal genomes. Green boxes indicate presence of an ortholog of that subunit, where the intensity of the green corresponds to the number of members of that orthologous group in that species (Marcet-Houben and Gabaldón 2009)

3. Comparative genomics of apoptosis and other programmed cell deaths.

Programmed cell-death is a central biological process that in eukaryotes it evolved into different complex pathways. It acts as a control mechanism of homeostasis in cell number and is triggered by the onset of a coordinated biochemical cascade of events, a process called apoptosis in metazoan species. Although much progress has been made in the recent years, its origins and evolutions remain to be resolved. The aim of this project is to trace the evolution of the cell-death pathways in eukaryotes and to unravel the evolutionary relationship between caspases and mitochondria, which play a central role in the caspases-dependent apoptotic process. This aim includes tracing the origin of key components of the apoptotic pathways in order to ascertain their eukaryotic or endosymbiotic origin and to determine the evolutionary period in which they were recruited to function in cell-death pathways. Furthermore we will use genomic-context techniques to discover proteins that evolved coordinately with known apoptotic components. Such predictions will be tested experimentally. This project is performed in collaboration with the lab of Dr. Cristina Muñoz-Pinedo (IDIBELL, Barcelona) and is funded through as AICR grant. We have developed the DeathBase, a database for the structure, function and evolution of apoptotic proteins *(www.deathbase.org)*

4. Phylogenomics and genome evolution

In the genomic era it has been possible to move from the evolutionary analysis of single protein families (phylogenetics) to that of complete genomes and proteomes (phylogenomics). To achieve this transition new tools have been developed that allow the large-scale reconstruction of thousands of phylogenetic trees in an automatic way. This computerization of the whole process of tree construction often involves the use of standard parameters and conditions for all tree families, inevitably resulting in poor or incorrect phylogenies in many cases.

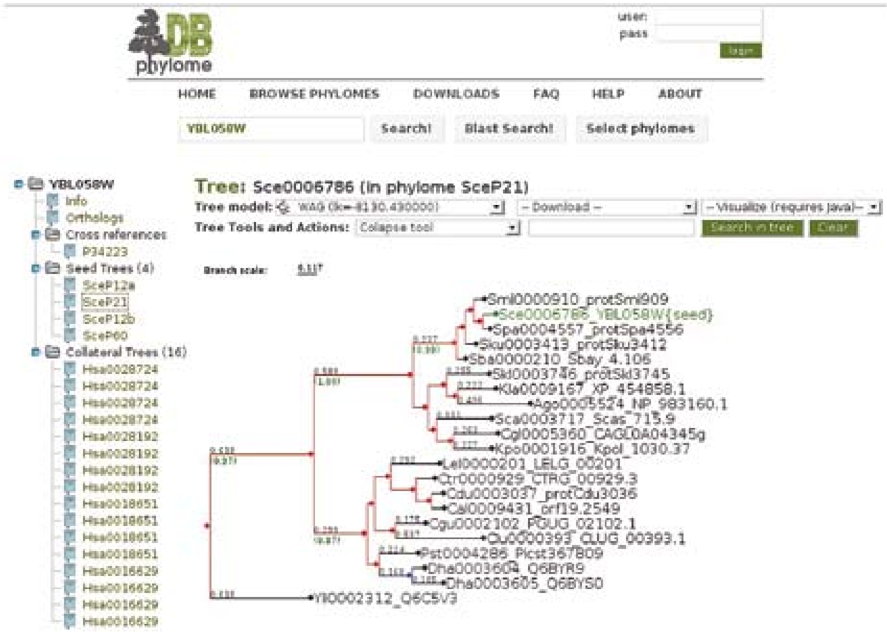


Figure 2.

Screenshot of a PhylomeDB entry (http://phylomedb.org) our database for complete phylomes. All visualizations are created by in-house programs (e.g http://ete.cgenomics.org)(Huerta-Cepas et al. 2008).

Moreover, interpreting such type of complex data poses many difficulties and does require the development of novel algorithms, tools, forms of representing the data and even new semantics and concepts. We combine the development of original algorithms to treat phylogenomic data with its application to gain knowledge on problems of biological relevance. In particular we are interested in developing post-processing methods to interpret sequence alignments and phylogenetic trees in a large-scale and to mine such data to find evidence for functional interactions between proteins.

5. Evolution of the eukaryotic cell

Every eukaryotic organism shows a high level of sub-cellular compartmentalization that is significantly more intricate than the most complex prokaryotic cell. How such degree of complexity came to be is still not fully understood. In this context, endo-symbiotic events with bacterial organisms have been proposed to be the source of a number of organelles including mitochondria, chloroplasts and peroxisomes. Only recently, it has been possible to contrast these hypotheses with the growing availability of completely sequenced genomes and organellar proteomic data. We use large-scale evolutionary analyses to investigate the origin and evolution two most widespread organelles for which an endosymbiotic origin has been proposed: mitochondria and peroxisomes.

PUBLICATIONS

Gabaldón T.
"Peroxisome diversity and evolution."
Philos Trans R Soc Lond B Biol Sci, in press.

Mihailovich M, Militti C, Gabaldón T, Gebauer F.
"Eukaryotic cold shock domain proteins: highly versatile regulators of gene expression."
Bioessays, in press.

Huerta-Cepas J, Dopazo J, Gabaldón T.
"ETE: a python Environment for Tree Exploration."
BMC Bioinformatics, in press.

Marcet-Houben M, Marceddu G, Gabaldón T.
"Phylogenomics of the oxidative phosphorylation in fungi reveals extensive gene duplication followed by functional divergence."
BMC Evol Biol 9:295 (2009).

Marcet-Houben M, Gabaldón T.
"Acquisition of prokaryotic genes by fungal genomes."
Trends Genet. 2009 Dec 5. [Epub ahead of print]

Schlüter A, Real-Chicharro A, Gabaldón T, Sánchez-Jiménez F, Pujol A.
"PeroxisomeDB 2.0: an integrative view of the global peroxisomal metabolome."
Nucleic Acids Res. Epub 2009 Nov 5.

Gabaldón T, Dessimoz C, Huxley-Jones J, Vilella AJ, Sonnhammer EL, Lewis S.
"Joining forces in the quest for orthologs."
Genome Biol 10(9):403 (2009).

Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T.
"trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses."
Bioinformatics 25(15):1972-3 (2009).

Gil-Borlado MC, González-Hoyuela M, Blázquez A, García-Silva MT, Gabaldón T, Manzanares J, Vara J, Martín MA, Seneca S, Arenas J, Ugalde C.
"Pathogenic mutations in the 5' untranslated region of BCS1L mRNA in mitochondrial complex III deficiency".
Mitochondrion 9(5):299-305 (2009).

Marcet-Houben M, Gabaldón T.
"The tree versus the forest: the fungal tree of life and the topological diversity within the yeast phylome."
PLoS One 4(2):e4357 (2009).

BIOINFORMATICS AND GENOMICS

Group Evolutionary Genomics

Group structure:

Group Leader: Fyodor Kondrashov

Research Assistants: Margarita Meer, Inna Povolotskaya, Michael Breen

SUMMARY

Evolutionary thought is pertinent to all aspects of biological or biomedical sciences. Similarly, discoveries in molecular and systems biology provide a deeper understanding of the mechanisms involved in the origin and continued evolution of living things. Our laboratory attempts to strengthen the connection between evolution and other areas of biology by working on the interface of several disciplines. In the past year we have started the transition from being an exclusively computational laboratory towards integrating some of our computational work with basic labwork. However, the majority of our research is still centred around the application of computational and bioinformatical tools to study any publicly-available genome-wide data.

We make no categorical distinction in our choice of the organism of study, nor do we exclude particular evolutionary mechanisms from our interests. Since our work relies almost exclusively on the available data many of our studies use the available data from humans and model organisms. However, for our experimental work we are focusing on some specific organisms. Over the past year we have worked on several different projects but we completed only some of them.

RESEARCH PROJECTS

1. Compensatory evolution

We are interested in describing and defining cases of epistasis, or cases where the fitness in one site depends on the state in other sites. Perhaps the simplest form of epistasis is found in compensatory interactions within the structure of a single molecule. For example, in an RNA stem structure, the interacting Watson-Crick complementary pairs depend on each-other to maintain the structure and high fitness. Likewise, the fitness of a substitution in one of the two complementary sites in a RNA stem structure depends on the nucleotide of the other site.

Last year we have made a comprehensive study of compensatory, Watson-Crick switches in the stem structures of mammalian mitochondrial tRNAs. Such switches, when one Watson-Crick pair in a secondary structure is substituted for another (for example an GC pair into an AU pair), are frequently observed in mammalian mt-tRNAs (Figure 1). Obviously, such switches require two substitutions, so a GC -> AU switch can proceed either through an GU or through the AC intermediate. Previous works assumed that such switches occurred through the stable GU intermediate. We tested this hypothesis using evolutionary and polymorphism data and showed that both the GU and AC intermediate variants are deleterious. Thus, compensatory evolution in mt-tRNAs must proceed through deleterious intermediates indicating that deleterious variants are not evolutionary dead ends but may rather be evolutionary stepping stones for further evolution.

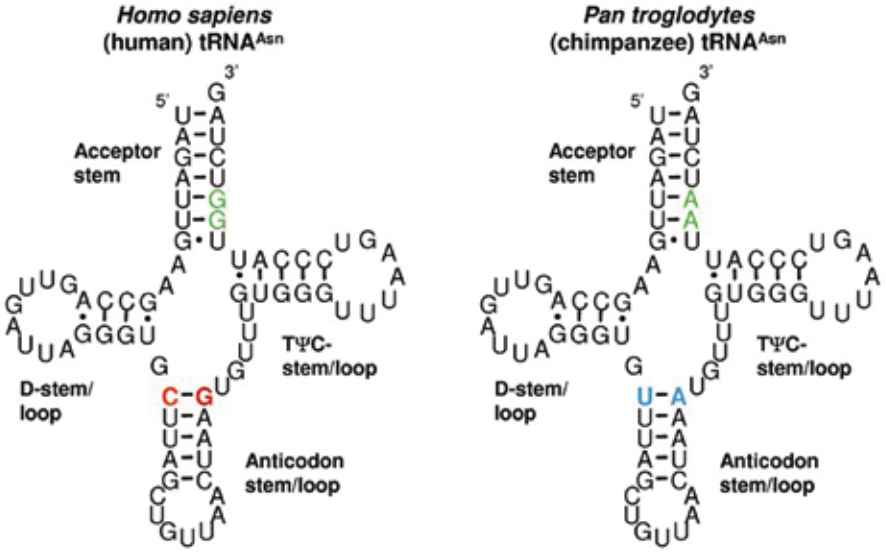


Figure 1.
An example of a Watson-Crick compensatory switch in mt-tRNA^{Asn}. Humans have a CG interacting pair (red) while chimps have an UA pair (blue) in the anticodon stem, which must have occurred through two substitutions and rapidly, as there are only two other differences in the entire tRNA (green).

2. The limit of protein evolution

In the context of another project we formulated a basic population genetics question: what is the limit of protein divergence under constant (non-epistatic) selection and what is the rate of reaching this limit? We were surprised, but apparently this fundamental question in molecular evolution has never been asked previously. The formulation of the theory for 2 alleles is simple. We assume that the rate of mutation is sufficiently low such that only one allele is segregating in the population at any given time. This assumption allowed us to use the Bulmer equations (Genetics 1991) that show the flux from one allele to another, which is a function of both selection and mutation. We model ed two diverging sequences starting from a mutation-selection equilibrium. Obviously, the limit of neutral divergence of two nucleotide sequences (4 alleles) will be 0.75, or for a sequence with two alleles 0.5. Also, it is evident, that constant selection will reduce this limit. If every substitution is strongly deleterious than the limit of divergence will be 0, that is to say that the two sequences will never diverge.

Formal theory for the case of two alleles is simple. The rate of change of divergence (dD/dt where D is the fraction of divergent sites in the two evolving sequences) is determined by the rate at which matching sites in the two sequences become mismatches and the rate at which mismatching sites become matching. The limit of the evolution is reached with the number of mismatches becoming matches is equal to the number of matches becoming mismatches. Formally, for the case of two alleles, this can be expressed in the following equation:

$$\frac{dD}{dt} = 2f_{1,2}x_{1,1} + 2f_{2,1}x_{2,2} - (f_{1,2} + f_{2,1})x_{1,2} \tag{1}$$

where $x_{1,1}$ is the frequency of matches of allele one in the two sequences, $x_{2,2}$ is the frequency of matches of allele 2 and $x_{1,2}$ is the frequency of all mismatches, $f_{1,2}$ is the flux of allele 1 into allele 2 and $f_{2,1}$ is the flux of allele 2 into allele 1. The fluxes and frequencies of matches depend both on the mutation rate and the strength of selection and are derived from Bulmer (Genetics 1991). Thus, the equation expands into

$$\frac{dD}{dt} = \frac{2\mu S}{1-e^{-S}}\left(\frac{e^{-S}}{1+e^{-S}} - \frac{1}{2}D\right) - \frac{2\mu S}{1-e^{S}}\left(\frac{e^{S}}{1+e^{S}} - \frac{1}{2}D\right) - D\mu S\left(\frac{1}{1-e^{-S}} - \frac{1}{1-e^{S}}\right) \tag{2}$$

with an analytical solution of

$$D_{2,any\_selection}(t) = \frac{2e^{S}}{\left(e^{S}+1\right)^{2}}\left(1 - \exp\left(-2\mu S\frac{e^{S}+1}{e^{S}-1}t\right)\right) \tag{3}$$

implying that the limit of sequence divergence for two alleles is

$$E = \frac{2e^{S}}{\left(e^{S}+1\right)^{2}} \tag{4}$$

and the asymptotic rate at which this limit is reached is

$$r(S) = 2\mu S\frac{e^{S}+1}{e^{S}-1} \tag{5}$$

It is simple to show that there is a simple maximum of r at s=0. Therefore, constant negative selection can only accelerate the rate at which the limit of sequence divergence is reached. This is slightly different and more complicated for the case of 4 alleles. However, qualitatively the result holds: constant negative selection cannot substantially slow down reaching the limit of sequence divergence. Since stable sequence divergence limits are practically never observed in real genome comparisons, our model implies that epistasis must be rampant across all types of functional sequences.

3. Unfinished business

There are several projects that we have started but not finished in the last year. Firstly, we began studying the evolution of Methanosarcina, which is a methane-metabolizing, perrolysine-utilizing archaea. We suspect that there may be interesting evolutionary patterns in the evolution of stop codons of this group of archeas. To be able to investigate this aspect of their genome evolution we required several genomes from closely related strains. Thus, with the help of our Ultrasequencing core facilities we have sequences and began to assemble the genomes of six Methanosarcina strains.

PUBLICATIONS

Alkalaeva E, Eliseev B, Ambrogelly A, Vlasov P, Kondrashov FA, Gundllapalli S, Frolova L, Söll D, Kisselev L.
"Translation termination in pyrrolysine-utilizing archaea."
FEBS Lett. 583:3455-3460 (2009).

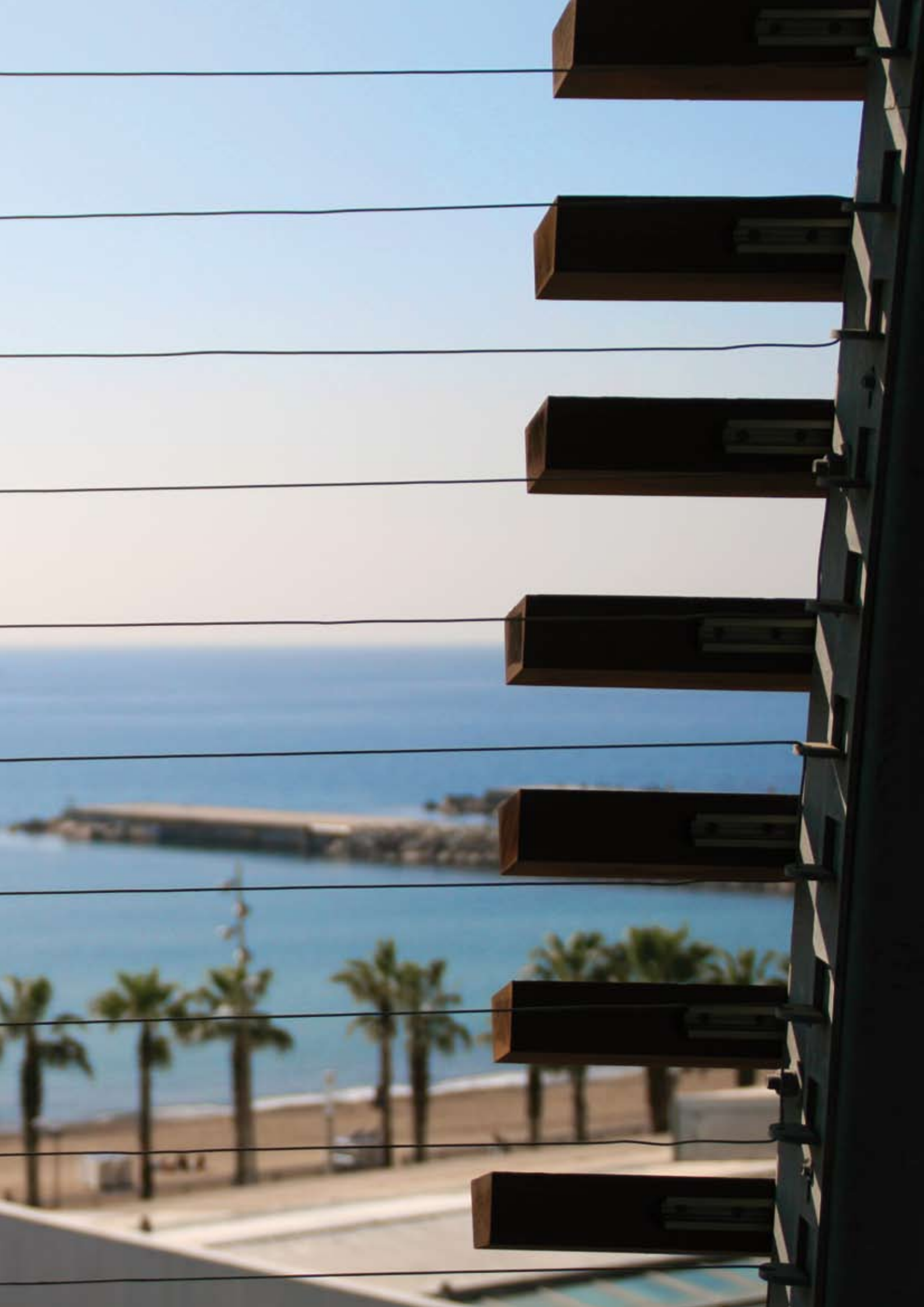
Innan H, Kondrashov FA.
"The evolution of gene duplications: classifying and distinguishing between models."
Nat Rev Genet, in press.

Meer MV, Kondrashov AS, Artzy-Randrup Y, Kondrashov FA.
"Compensatory evolution in mt-tRNAs navigates valleys of low fitness."
Nature, in press.

Kondrashov AS, Povolotskaya IS, Ivankov DN, Kondrashov FA.
"Rate of sequence divergence under constant selection."
Biology Direct, in press.

Book chapters

Kondrashov, FA.
"Evolution after Gene Duplication."
In: Gene Dosage and Duplicaton. Publisher: Willey, in press.



CELL AND DEVELOPMENTAL BIOLOGY

Coordinator: Vivek Malhotra

The programme of Cell and Developmental Biology has succeeded in recruiting a new group leader, Pedro Carvalho from Harvard University. Pedro will arrive at the CRG in July 2010. Pedro is interested in the mechanism of retro translocation at the endoplasmic reticulum and the process of lipid body biogenesis. The arrival of Pedro will strengthen the overall interest of our department in the process of protein targeting and organelle biogenesis. Vivek Malhotra is elected a member of EMBO.

Structure of the Programme:

Coordinator:	Vivek Malhotra
Senior Group:	Isabelle Vernos
Junior Group:	Hernan Lopez-Schier
Junior Group:	Manuel Mendoza
Junior Group:	Jérôme Solon

CELL AND DEVELOPMENTAL BIOLOGY

Group: Intracellular Compartmentation

Vivek Malhotra is a Senior ICREA Group Leader and the coordinator of the department.

Group Structure:

Group Leader: Vivek Malhotra

Postdoctoral Fellows: Kota Saito, Amy Curwin, Yuichi Wakana, Juan Duran, Julia von Blume, Sandra Mitrovic, Josse van Galen, Felix Campelo

PhD student: Caroline Bruns

Visiting student: Elena Forlanelli

Technician: Anne-Marie Alleaume

SUMMARY

We are interested in the mechanism of protein secretion and biogenesis of Golgi membranes.

RESEARCH PROJECTS

1. Membrane fission

Yuichi Kawana, Felix Campleo and Anne-Marie Alleume

The Golgi apparatus is composed of cisternae (flat membranes) that are stacked and kept near the centrioles in mammalian cells. Why such unique organization and spatial location? How is this organization regulated during protein transport? The secret here is membrane fission (cutting) that is necessary to generate transport carriers. Membrane fission must be regulated to generate transport carriers commensurate with cargo size, prevent formation of empty carriers or complete conversion of Golgi into small vesicles, during protein secretion? We asked a simple question whether there were chemicals that would vesiculate Golgi membranes by uncontrolled activation of membrane fission. Identification of such a compound would make the process of membrane fission amenable to molecular analysis

Fluorescence microscopy was used to screen 600 sponge metabolites and a compound called Ilimaquinone (IQ) was identified based on its property to specifically vesiculate Golgi membranes into small 60-90 nm size vesicles (Takizawa et al., Cell 1993). IQ mediated Golgi vesiculation was reconstituted in vitro, which revealed the involvement of trimeric G protein subunits ßg. This was the first clear documentation of the involvement of trimeric G proteins in membrane fission (Jamora et a., Cell. 1997). Soon thereafter we found that Gßg activated a serine/threonine protein kinase called PKD (Jamora et al., Cell. 1999). The next obvious question was to test whether PKD is required for formation of transport carriers from Golgi membranes during protein secretion. This was indeed the case and we found that inhibition of PKD blocked secretion of proteins, interestingly and specifically, from Golgi to the cell surface. Under these conditions, cargo destined to the cell surface was packed into transport carriers, but the carriers failed to undergo cutting (fission) and grew, consequently, into large tubules. This was the first key component, a protein kinase, identified to be essential for regulating events leading to the biogenesis of a cargo filled Golgi to cell surface transport carrier (Liljedahl et al., Cell. 2001). How is PKD recruited to the Golgi membranes? We found that diacylglycerol was required for the recruitment of PKD to the Golgi membranes (Baron and Malhotra, Science. 2001). In collaboration with Dr. Thomas Suefferlein, we have recently found that PKD also binds to ARF1 at the Golgi. ARF1 is known for its role in protein transport by recruiting coat proteins such as COPI and Clathrin in a process that leads to the formation of transport carriers. We suggest that AFR1 bound to PKD activates phospholipase D (PLD), which is then involved in production of phosphatidic acid (PA) at the TGN. Therefore, DAG and ARF, required for the recruitment of PKD generate PA, which is then [somehow] involved in events leading to the biogenesis of cargo filled transport carriers. There are three different isoforms of PKD in the mammalian cells and surprisingly, all three are required for Golgi to cell surface transport (Yeaman et al., Nature Cell Biology. 2004). Moreover, all three forms are specific for the trafficking of only those proteins that are destined to the basolateral cell surface. Apically targeted proteins do not require the activity of PKD. Our most recent findings suggest that all mammalian cells contain 2 of the 3 isoforms and these two isoforms form homo (1-1; 2-2, for example) and heterodimers (1-2, for example. Bossard et al., Journal of Cell Biology. 2007). There are numerous reports now on the involvement of PKD in protein secretion but the challenge is to understand the molecular mechanism of downstream events. PKD has been shown to activate a lipid kinase called PI4KIIIb, which converts phophatidylinositol (PI) into phophatidylinositol 4-phosphate (PI4P). PI4P is required for Golgi to cell surface transport. We suggest that local production of PI4P by PKD dependent activation of PI4KIIIb recruits specific effectors that are required for membrane fission (Klaus Pfizenmeir and colleagues. Germany). PKD also phosphorylates a ceramide transfer protein called CERT. Phosphorylation of CERT by PKD prevents the attachment of the former to the Golgi membranes (Klaus Pfzemmeir and colleagues. Germany). The role of ceramide at the TGN is not clear (Bard and Malhotra, Ann Rev Cell and Dev Biol.2006). Clearly, a number of components involved in PKD dependent TGN to cell surface transport have been identified and in order to gain mechanistic insights, we this process in permeabilized mammalian cells. The transport carrier containing membrane fraction has been analyzed by mass spectrometry and revealed a number of new proteins. This assay is being used to reveal the role of proteins and lipids that assemble at the TGN for the generation of TGN to cell surface transport carriers.

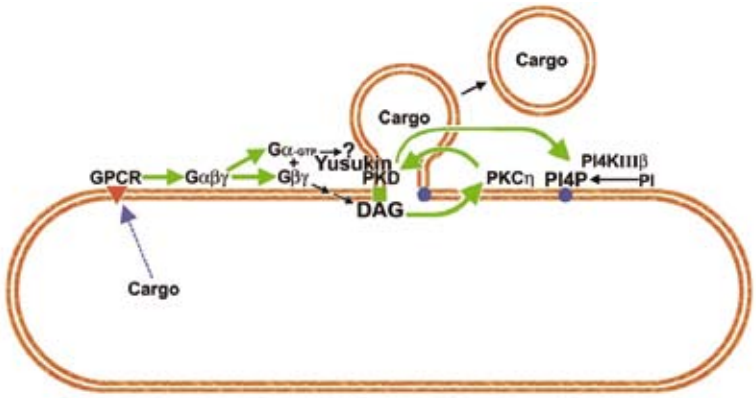


Figure 1.
Regulation of membrane fission by PKD. A trimeric G-protein is required for the recruitment and activation of PKD. PKD regulates the activity of a number of components to assemble a fission machine, which is used to cut cargo filled transport carriers from the TGN.

2. New protein transport components

Kota Saito and Julia von Blume

A genome wide screen was carried out to identify new components required for protein secretion. From the 22,000 genes tested, 110 gene products were identified to be essential for secretion. This contained 22 previously known components. The rest were cloned, tagged and expressed to identify the intracellular location of the cognate proteins. 20 genes products were further selected, based on their localization to compartments of the secretory pathway (Bard et al., 2006. Nature). The new genes of interest are called TANGO for Transport And Golgi Organization. TANGO1, 5 and a previously identified gene called twinstar are being further characterized for their roles in protein transport and Golgi organization.

TANGO1 is required for loading cargo into transport carriers at the Endoplasmic Reticulum (ER) exit sites (Kota et al., Cell. 2009). Twinstar is required for the sorting of secretory cargo at the TGN (Von Blume et al., JCB. 2009). Our aim is to understand the mechanism by which TANGO loads cargo without entering the transport carrier, and the process by which twinstar dependent actin remodeling helps in cargo sorting.

3. Unconventional secretion

Juan Duran, Josse van Galen, Amy Curwin and Caroline Bruns

Some of the interleukins, fibroblast growth factor-2 (FGF2), inhibitor of macrophage migration (MIF), Galectins etc., are secreted from cells without entering the ER-Golgi pathway. These secreted proteins are key players in the immune response, cell growth, angiogenesis, but the mechanism of their release from cells remains mysterious. We have found that the Golgi associated protein called GRASP in Dictyostelium is required for secretion of a protein called AcbA. AcbA, like the proteins mentioned above, lacks a signal sequence necessary for targeting to the ER, and is secreted unconventionally (Kinseth et al., Cell. 2007). Unconventional secretion of AcbA ortholog Acb1 in yeast requires GRASP, genes involved in autophagosome formation, transport to the early endosome, multivesicular body formation and the cell surface specific T-SNARE (Duran et al., JCB. 2010 in press). We are using yeast and mammalian cells to reveal the mechanism of unconventional secretion.

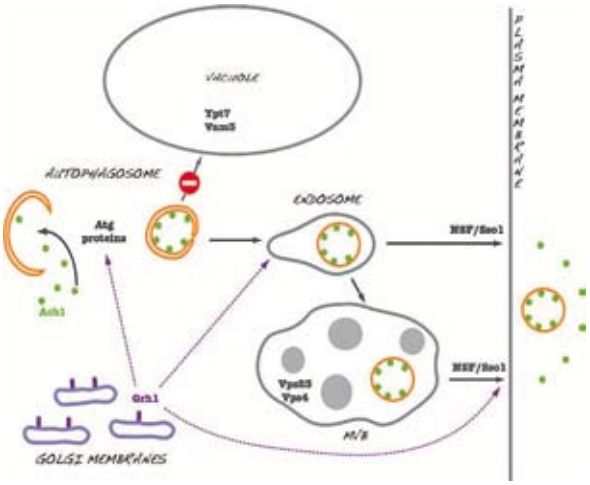


Figure 2.
The Acb1 secretion pathway. Cytosolic Acb1 is packaged into autophagosomes, which fuse with early endosomes. The early endosomes containing Acb1 either fuse directly with the cells surface, or more likely mature into a multivesicular body (MVBs). The MVB fuses with the cell surface to release exosomes containing Acb1. Grh1 is required for the secretion of Acb1 and has to be membrane associated for its role in this pathway. However, the exact site of action for Grh1 in unconventional secretion remains unknown.

4. Mechanism of Mucin secretion

Sandra Mitrovic

Mucins are heavily glycosylated secretory proteins, however, the mechanism by which they are sorted at the TGN, packed into secretory storage granules and then secreted in a regulated manner is not known. We have reconstituted mucin secretion in human goblet cells and screened the genome to identify components involved in mucin synthesis and secretion. A large number of gene products have been identified and our aim is to understand their mechanism of action in mucin secretion. We are specifically interested in cell surface proteins that can be targeted to inhibit mucin secretion. This has the potential of developing valuable therapeutics for Asthma and chronic obstructive pulmonary disorders (COPD).

5. A Golgi organization specific cell cycle checkpoint

Julia Von Blume

Inhibiting changes in Golgi organization prevents entry of cells into mitosis (Sutterlin et al., Cell. 2004). Thus a mechanism exists to monitor organizational changes in Golgi and if there is any defect, entry of cells into mitosis is blocked. What is the molecular mechanism of this Golgi specific checkpoint? How is this event coordinated with other check points that are activated upon DNA damage and defective spindle dynamics? A combination of in vitro approaches and system wide SiRNA is being used to identify the Golgi specific cell cycle checkpoint.

PUBLICATIONS

Saito K, Chen M, Bard F, Chen S, Zhou H, Woodley D, Polischuk R, Schekman R and Malhotra V.
"TANGO1: a protein required for collagen VII packing into COPII transport carriers."
Cell 136:891-902 (2009).

Malhotra V.
"Vivek Malhotra: Gaga for the Golgi."
[Interview by Liz Savage]. J Cell Biol 184:4-5 (2009).

Emr S et al.
"Journey through the Golgi-taking stock in a new era."
J Cell Biol 187:449-453. Review. (2009).

Von Blume J, Duran JM, Forlanelli E, Alleaume AM, Egorov M, Polischuk R, Molina H and Malhotra V.
"Actin remodeling by ADF/Cofilin is required for cargo sorting at the Trans Golgi Network."
J Cell Biol 187:1055-1069 (2009).

Duran JM, Anjard C, Stefan C, Loomis WF and Malhotra V.
"Unconventional secretion of Acb1 is mediated by autophagosome."
J Cell Biol, in press.

Pusapati GV, Krndija D, Armacki M, vonWickert G, vonBlume J, Malhotra V, Adler G and Seufferlein T.
"Role of the second cystein-rich domain and Pro 275 in PKD2 interaction with ARF1, TGN recruitment and protein transport."
Mol Biol Cell, in press.

Book chapters

Malhotra, V, Warren, G and Mellman, I.
"Protein trafficking between membranes."
In: Lewin's Cells, Second edition, in press.

CELL AND DEVELOPMENTAL BIOLOGY

Group: Microtubule Function and Cell Division

Isabelle Vernos is a Research professor from ICREA.

Group structure:

Group Leader: Isabelle Vernos

Postdoctoral Fellows: Teresa Sardon, Sylvain Meunier, Roser Pinyol

Students: Vanessa Dos Reis Ferreira, Isabel Peset, Martin Schütz, David Vanneste, Antonios Lioutas

Technicians: Luis Bejarano (until June 2009), Leonor Avila (since July 2009), María Sanz, Nuria Mallol

SUMMARY

Research in my lab is directed at understanding the role of the microtubule network in cell organization and function. To address this question we study various microtubule-associated proteins (molecular motors and MAPs) and their regulators (kinases, phosphatases and the small GTPase Ran during M-phase). One major goal is to unravel how the self-organization of cellular components results in the morphogenesis of dynamic molecular machines. In 2007, we have focused on the morphogenesis of the bipolar spindle in mitosis and meiosis.

Our favourite experimental system is the Xenopus egg extract system for studies on cell cycle progression and regulation, microtubule dynamics, spindle assembly and chromosome behaviour (Karsenti and Vernos, 2001). We combine it with the use of human tissue culture cells in which we validate some of the results obtained in egg extract.

RESEARCH PROJECTS

Cell division is characterized by the dramatic reorganization of the microtubule network into a spindle shaped apparatus that segregates the chromosomes into the two daughter cells. Spindle assembly and function rely on complex protein interaction networks that are finely regulated in time and in space. In addition to phosphorylation-dephosphorylation reactions, recent work has shown that the small GTPase Ran in its GTP bound form plays an important role in the spatial regulation of spindle assembly (Gruss and Vernos, 2004). To understand the molecular mechanism underlying cell division we study the process of microtubule nucleation and stabilization during M-phase and the role of molecular motors in bipolar spindle assembly and chromosome movements.

1. The centrosome: a cellular regulatory hub

The centrosome is the major site for microtubule nucleation in animal cells and its activity is finely regulated during the cell cycle. In late G2 and prophase, the pericentriolar material expands by recruiting additional components, such as the γ-tubulin-ring complex and as a result the MT nucleation activity of the centrosome increases. Several kinases from different families are recruited to the centrosome in G2/M phase and become specifically activated as the cell enters mitosis. We want to understand how these kinases participate in centrosome maturation and activity, promoting microtubule nucleation and stabilization during cell division.

The centrosomal Aurora A kinase has been implicated in several important processes including centrosome maturation during G2, mitotic entry, centrosome separation and bipolar spindle assembly. We have previously shown that Aurora A interacts with TPX2 in a RanGTP dependent manner after nuclear envelope breakdown resulting in kinase activation. Using the Xenopus egg extract system we found last year that Aurora A works though different mechanisms to regulate MT assembly during mitosis, ensuring bipolar spindle formation (Sardon et al, 2008). At the centrosome active Aurora A is required for efficient MT nucleation. In addition, it promotes MT growth by recruiting TACC3 that works in concert with XMAP215 to oppose the destabilizing activity of XKCM1 (Peset I. et al, 2005). To obtain a useful tool to further investigate the function of Aurora A during the cell cycle we have collaborated with the group of Prof Giannis (Leipzig, Germany) to look for a specific inhibitor for this kinase. We have found one compound that shows specificity for Aurora A inhibition in cells suggesting that it could be useful for basic and applied research (Sardon et al, 2009).

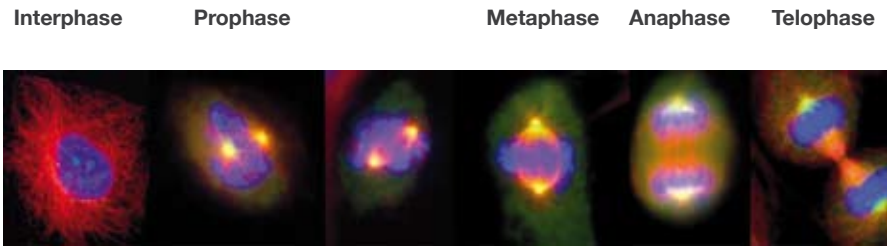


Figure 1.
Immunufluorescence analysis of HeLa cells shows the specific localization of the kinase Aurora A to the centrosome and the spindle poles during mitosis. MTs are labelled in red, DNA in blue and Aurora A in green.

2. The RanGTP pathway for MT nucleation and stabilization in M-phase

After nuclear envelope breakdown, a centrosome independent pathway relying on a RanGTP gradient triggers MT nucleation and promotes MT stabilization in the vicinity of the condensed chromatin. We want to understand the molecular mechanism that triggers MT nucleation in this acentrosomal pathway.

We have previously identified TPX2 as being essential for this pathway and for spindle assembly both in mitosis and meiosis (Gruss and Vernos, 2004; Gruss et al., 2002; Wittmann et al., 2000). Interestingly, TPX2 is highly conserved in plants that relie exclusively on an acentrosomal pathway for spindle assembly as they do not contain centrosomes like animal cells. Last year in collaboration with two groups working in plants we have shown that TPX2 has conserved functions in animals and plants (Vos et al, 2008; Evrard et al, 2009). We are continuing our efforts to unravel the molecular mechanism underlying the RanGTP dependent pathway for MT assembly in M-phase, focusing on TPX2 and its interaction partners.

3. Role of molecular motors in spindle assembly and chromosome movements

Spindle bipolarity is essential for correct chromosome segregation but the mechanism underlying its establishment is still not completely understood. Furthermore very little is known about how metaphase spindles maintain a stable bipolar configuration before anaphase. Although it has been established that a balance of forces, generated by plus and minus-end directed motors, mostly Eg5 and dynein, is required for bipolar spindle assembly, whether these forces still play a role in metaphase is unknown.

This year we have characterized another motor, Hklp2, the human homologue of Xklp2 (Boleti et al, 1996; Wittmann et al, 1998) and found that this motor plays a role in bipolar spindle assembly and stability.

Although Eg5 has a predominant and essential role in bipolar spindle establishment, Eg5 inhibition does not compromise the stability of the bipolar spindle at metaphase. We found that Hklp2 becomes essential under these conditions for spindle stability. Hklp2 localizes to the spindle microtubules and the chromosomes in metaphase. This steady state distribution is essential for Hklp2 role in promoting the switch from the monopolar to the bipolar configuration and in stabilizing spindle bipolarity in metaphase. Our data provide an additional understanding of the mechanism driving the initial establishment and subsequent maintenance of the bipolar spindle at metaphase, and reveals the existence of a novel specific mechanism that stabilizes bipolar spindles and may be essential to prevent mitotic defects (Vanneste et al, 2009).

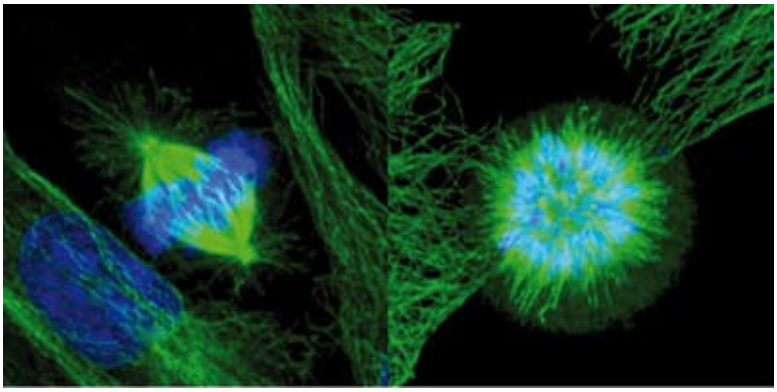


Figure 2.
Confocal images of mitotic HeLa cells. At metaphase, the bipolar spindle has assembled around the condensed chromosomes positioned at an equatorial position ready to be segregated by this molecular machine (left). When some motors like Eg5 and Hklp2 do not function properly the bipolar spindle is unstable and collapse into a monopolar configuration with the subsequent failure of cell division (right). MTs are shown in green, chromosomes in blue.

PUBLICATIONS

Sardon T, Cottin T, Xu J, Giannis A and Vernos I.
"Development and Biological Evaluation of a novel Aurora A Kinase inhibitor."
ChemBioChem 10(3):464-478 (2009).

Evrard JL, Pieuchot L, Vos JW, Vernos I and Schmit AC.
"Plant TPX2 and related proteins."
Plant Signaling & Behavior 4(1):69-72 (2009).

Vanneste D, Takagi M, Imamoto N and Vernos I.
"The role of Hklp2 in the stabilization and maintenance of spindle bipolarity."
Current Biology 19(20):1712-7 (2009).

PATENT

Application number: EP 09382002
Priority date: 15/01/2009
Applicant(s): Universität Leipzig [DE]; CRG[SP]; ICREA[SP]
Inventor(s): Athanassios Giannis (Univ. Leipzig) Thomas Cottin (Univ. Leipzig) Teresa Sardón (CRG) Isabelle Vernos (CRG)

CELL AND DEVELOPMENTAL BIOLOGY

Group: Sensory Cell Biology and Organogenesis

Group structure:

Group Leader: Hernán López-Schier (Ramón y Cajal fellow)

Postdoctoral Fellows: Marana Muzzopappa (EMBO postdoctoral fellow), Adele Faucherre (Marie Curie postdoctoral fellow), Jean Pierre Baudoin (FRM postdoctoral fellow)

Graduate Students: Indra Wibowo, Filipe Pinto Teixeira, Jesús Pujol Martí

Technician: Andrea Durán

Visiting professor: Giuseppe Montalbano (University of Mesina, Italy)

SUMMARY

Research in my laboratory focuses on understanding the fundamental principles that govern the development of tissues and organs in vertebrates. We attempt to define the cellular and molecular bases of the acquisition and maintenance of tissue architecture and neurral circuits, and their relationship to the function of sensory organs.

RESEARCH PROJECTS

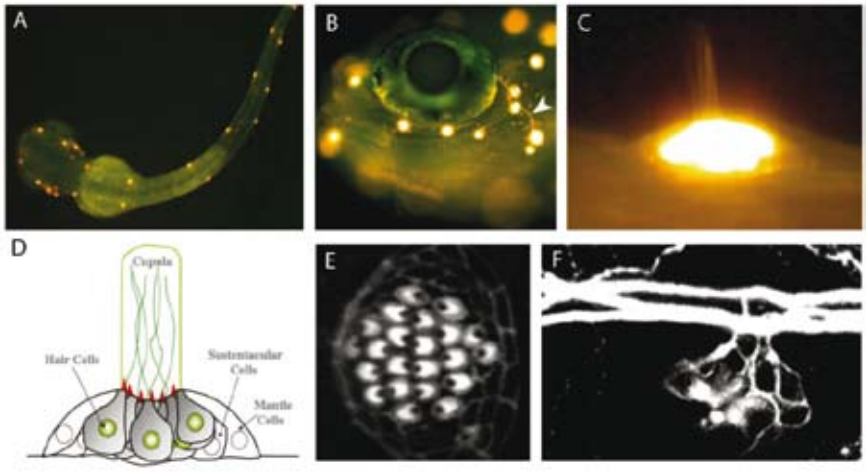
1. Cellular responses to polarity signals, in particular planar cell polarity

We use the mechanosensory lateral line of the zebrafish (*Danio rerio*) as a model system to study cell-fate specification, and the formation and remodelling of epithelial architecture during organ development and regeneration. For our studies, we employ cellular, genetic and molecular approaches and state-of-the-art optical imaging techniques, and are also developing methods to analyse *in vivo* the reinnervation of sensory cells during regeneration. In the long term, our studies should provide insight into how sensory organs develop and regenerate, and how their cellular organisation and function are maintained throughout life.

The coordinated orientation of polarised cells within the plane of an epithelium is termed planar cell polarity. The orientation of hair cells within the neuroepithelium of the inner ear represents a striking example of planar cell polarity in vertebrates. Directional deflections of apical mechanosensitive organelles (stereocilia), respectively open or close transduction channels to depolarise or hyperpolarise the hair cell's plasma membrane. The axis of morphological polarity of the stereocilia therefore corresponds to the direction of excitability of the hair cell, and bestows the organ with maximal sensitivity to mechanical stimuli. The senses of hearing and equilibrium thus rely on the exquisite precision with which hair cells are oriented across the sensory epithelium. In spite of its importance, we only have a very superficial knowledge of the mechanisms that control the planar polarisation of hair cells.

Some aquatic vertebrates sense directional water movements with the lateral-line system, a sensory organ closely related to the inner ear. This system comprises a stereotyped array of sensory clusters called neuromasts, each with a very simple organisation. A neuromast contains two types of peripheral supporting cells and a few centrally located hair cells innervated by afferent and efferent axons (Figure 2 left). Hair cells in neuromasts are polarised within the plane of the epithelium in a way comparable to that of the inner ear (Figure 2 right). The lateral-line organ of the zebrafish is thus ideally suited to investigate the mechanisms that control hair-cell planar polarisation.

Figure 1.
Low magnification view of a living zebrafish larva whose hair cells in the lateral line were labelled with the fluorescent vital dye DiAsp (bright orange). It shows the superficial and systematic distribution of neuromasts along the anterior (head) and posterior (trunk and tail) lateral-line systems. B: Higher magnification of head neuromasts revealing the transynaptic transport of DiAsp from hair cells to afferent nerves (white arrowhead). C: The highest magnification of a neuromast labeled with DiAsp shows a core of hair cells and their apical-projecting hair bundles. D: Scheme of a neuromast with its constituent cells. Neurons are not depicted. E: The orientation (planar polarity) of hair cells is evident in this neuromast whose actin-rich hair bundles were stained with phalloidin. The kinocilium (not stained) appears as a black hole giving the hair bundle its horseshoe appearance. F: Axonal arbor of two afferent neurons branching off the lateralis nerve below a neuromast.



We have now identified a series of mutations that disrupt the establishment and maintenance of planar cell polarity in neuromasts. The combination of these strains with several of our multicolour transgenic animals will permit us to generate three- and four-dimensional images of living wild type and mutant specimens with great precision, and to track protein localisation patterns within seconds, or cellular behaviours over days. The combination of the genetic approaches afforded by the zebrafish with live imaging shall allow us to understand sensory-organ development, regeneration, and function in whole animals and at the single-cell level.

2. Sensory organ innervation

Historically, planar cell polarity has been studied in invertebrates on tissues that undergo polarisation during a very brief period, to eventually become fixed with negligible or non-existent plasticity, including lack of cellular proliferation, tissue remodelling or cell migration. Such tissues, consequently, will not undergo repair or regeneration after cell death or mechanical damage. Genetic and molecular studies in *Drosophila* have shown that the establishment of planar polarity relies on the concerted activity of many proteins. The cellular responses to polarity cues, especially in remodelling tissues, are not understood. We are defining the cellular and molecular bases underlying the acquisition and maintenance of planar cell polarity, epithelial architecture and innervation in a vertebrate, and its relationship to organ function.

3. Sensory-organ growth and regeneration, with an emphasis on epithelial remodelling

Sensory perception is a complex process that allows organisms to sample the environment and to react appropriately. Sensory dysfunction can thus be a major handicap that dramatically decreases the quality of life of the affected individual. All sensory modalities are liable to deteriorate during one's lifetime. Hearing deficits, for example, afflict more than 10% of the population in industrialized countries, including 0.1% of newborn children and 50% of those aged 80 years or over. Some sensory organs have an impressive capacity to recover after environmental insult, while others can lose function permanently. The inner ear is among the later: hearing loss owing to the degeneration or denervation of the mechanosensory hair cells is irreversible.

Although the search for a hair-cell progenitor resident in sensory epithelia has been pursued for over twenty years, to date there are no reports demonstrating the identification, or even the existence of such

cell type. Our recent work has identified a hair-cell progenitor in neuromasts, which allows us the analysis of hair-cell development from its very outset. It also suggests the existence of a stem-cell population, and pinpoints its location within the neuromast. Within the context of this research, we are also attempting to devise methods to follow every cells and complex tissue movements to reconstruct a digital organ *in vivo* (Figure 2).

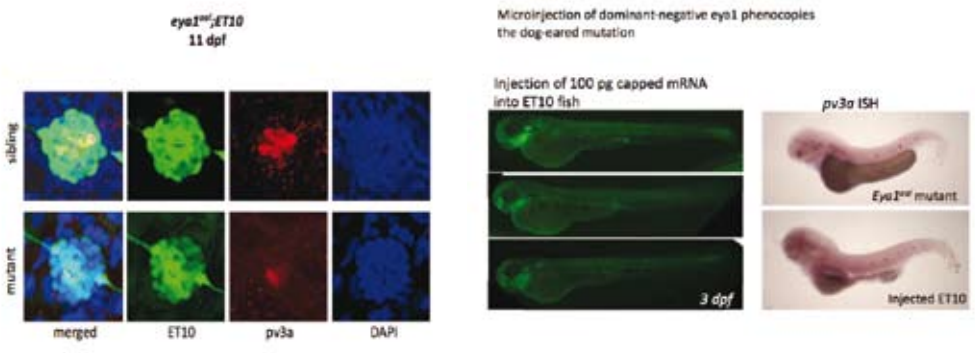


Figure 2.
High magnification view of wild type and eya1/dog eared zebrafish larvae whose hair cells in the lateral line were labelled with PV3a (red). Fish are transgenics for ET10, which expresses the GFP in supporting cells (green). Blue is DAPI staining. It shows that hair cells degenerate in eya1/dog eared. B: Low magnification of zebrafish of the ET10 line expressing Eya1DN. It phenocopies the eya1/dog eared mutation.

4. Tissue mechanics

We have recenlty begun to investigate the physical forces that shape and maintain the achitecture of epithelia in the zebrafish. We plan to use genetic, microscopic, and biophysical approaches to adreess this biological problem. We are also developing mathematical tools to analyse the quantitative data generated through the above-mentioned approaches, in the hope to model organ formation in silico.

These studies shall provide insight into how organs develop and function throughout life, and also how they regenerate and re-innervate to recover function after damage. This not only represents a very interesting biological problem, but also is relevant to the successful application of therapies aimed to restore sensory function in humans, for aberrant repair would prevent the organ from performing properly.

PUBLICATIONS

Gleason MR, Nagiel A, Jamet S, Vologodskaia M, López-Schier H, Hudspeth AJ.
"The transmembrane inner ear (Tmie) protein is essential for normal hearing and balance in the zebrafish".
Proc Natl Acad Sci U S A106(50):21347-52 (2009).

Faucherre A, Pujol-Martí J, Kawakami K, López-Schier H.
"Afferent neurons of the zebrafish lateral line are strict selectors of hair-cell orientation".
PLoS One 4(2):e4477 (2009).

Buschbeck M, Uribesalgo I, Wibowo I, Rué P, Martin D, Gutierrez A, Morey L, Guigó R, López-Schier H, Di Croce L.
"The histone variant macroH2A is an epigenetic regulator of key developmental genes".
Nat Struct Mol Biol 16(10):1074-9 (2009).

Froehlicher M, Liedtke A, Groh K, López-Schier H, Neuhauss SC, Segner H, Eggen RI.
"Estrogen receptor subtype beta2 is involved in neuromast development in zebrafish (Danio rerio) larvae."
Dev Biol 330(1):32-43 (2009).

CELL AND DEVELOPMENTAL BIOLOGY

Group: Coordination of Cytokinesis with Chromosome Segregation

Group structure:

Group leader:	Manuel Mendoza
Postdoctoral Fellow:	Alexandre Vendrell (from September 2009)
PhD students:	Gabriel Neurohr, Aina Masgrau (from September 2009), Iris Titos (from September 2009)
Technician:	Trinidad Sanmartin
Visiting student:	Eulalia Scheenaard (July-September 2009)

SUMMARY

Living cells have a fascinating ability to generate complex and dynamic internal structures. Nowhere is this property more evident than during mitosis and cytokinesis: in a very short time (often of the order of a few minutes) cells alter their shape, duplicate and partition their internal components, and divide into two apparently identical halves. These dramatic morphological changes need to be carefully coordinated with each other in space and time. To learn more about the principles underlying this coordination, we focus on the events at the end of the cell cycle: chromosome segregation and cytokinesis in the budding yeast, *Saccharomyces cerevisiae*.

RESEARCH PROJECTS

1. Mechanisms of chromosome segregation sensing by the NoCut checkpoint

Cell division is completed through partition of the cytoplasm by ingression of the cleavage furrow, and the subsequent cleavage of the cell membrane into two during abscission. At the same time, the replicated chromosomes are segregated to opposite ends of the cell. Whereas cleavage furrow ingression usually proceeds concomitantly with poleward movement of the chromosomes, abscission must take place exclusively after the last pair of sister chromatids have been pulled out of the cleavage plane. In budding yeast and human cells, a checkpoint known as NoCut delays completion of cytokinesis when chromosome segregation is impaired. Inactivation of NoCut leads to premature abscission, and late-segregating chromosomes are trapped and (at least in yeast) cut by the cytokinesis machinery.

The Chromosome Passenger Complex (CPC) component Aurora B kinase (Ipl1 in yeast) plays a central role in the NoCut checkpoint (Fig. 1). The CPC localizes to the spindle midzone during anaphase, and it is activated by the presence of acetylated chromatin around the midzone. Ipl1 activity mediates the translocation of two anillin-related proteins, called Boi1 and Boi2, to the site of cytokinesis during anaphase, where they inhibit cytokinesis until chromosomes are cleared from the cleavage plane.

NoCut relies on a complex network of factors to monitor chromatin segregation away from the spindle midzone. We are systematically characterizing the NoCut mechanism, taking advantage of genetic screens to identify new checkpoint components. We have already completed a genome-wide screen for non-essential NoCut genes. This has identified about 50 NoCut candidate genes, whose products fall in five main categories: (1) proteins of the spindle midzone; (2) histone modification and chromatin organization factors, (3) kinetochore proteins, (4) cleavage site proteins, and (5) 13 genes of unknown function. We predict that analysis of the identified mutants will identify factors that act upstream of the CPC, enabling its activation by chromatin (***sensors***); downstream of Boi1 and Boi2, inhibiting the abscission machinery at the site of cytokinesis (***effectors***); and mediating the communication between the CPC and Boi1/2, by shuttling between the nucleus and the cytoplasm or affecting the shuttling of other NoCut components (***transducers***). We are particularly encouraged by the identification of chromatin organization genes as NoCut candidates, since their characterization could shed light into the mechanism by which chromatin activates the CPC.

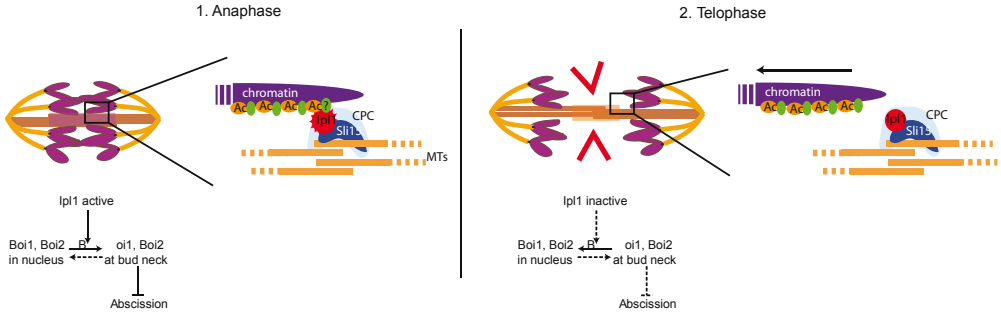


Fig. 1. *Model of the NoCut checkpoint. (A) In early anaphase, chromosomes (depicted in purple) are in contact with the CPC, which binds to spindle midzone microtubules (in orange). Inset: the CPC subunits Ipl1 and Sli15 are depicted in red and blue, respectively. Midzone-bound Ipl1 is kept active, probably through interaction with chromatin-associated factors (in green) which require histone acetylation (orange circles) to activate the CPC. As a result, Boi1 and Boi2 localize to the bud cortex, where they inhibit abscission. (B) When chromosomes are segregated away from the midzone, the CPC is no longer activated by chromatin and the NoCut signal is turned off; Boi1 and Boi2 leave the bud neck, and abscission (represented by red triangles) ensues.*

2. Coordination of chromosome length with spindle elongation

We are also interested in other mechanisms ensuring robust chromosome segregation. The function of the mitotic spindle is to separate sister chromatids away from each other. When the length of the anaphase spindle reaches twice the length of the longest chromosome arm, chromosome segregation is complete. Are there mechanisms that allow the cell to adjust the size of the anaphase spindle and/or anaphase dynamics, to the length of chromosome arms? We are addressing this question through manipulation of chromosome arm length in yeast cells, and analysis of mitotic and cytokinetic events by live cell microscopy in normal and mutant strains (Fig. 2).

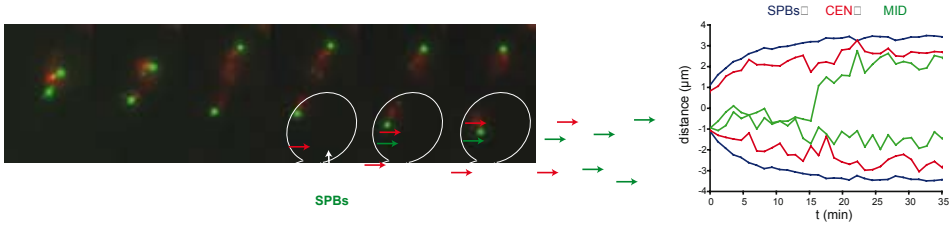


Fig. 2. *Analysis of chromosome segregation in living yeast cells. Spindle pole bodies are labeled with GFP (bright green dots) to track spindle elongation during anaphase. mRFP and GFP protein fusions allow tracking of the centromere (CEN) and middle of the right arm (MID) of chromosome IV, seen as dimmer red and green dots, respectively. The dynamics of the marked loci, relative to the mid-position of the spindle, is quantified in the graph. Note that CEN sequences segregate earlier than MID sequences, as indicated by arrows. Images by Gabriel Neurohr.*

PUBLICATIONS

Mendoza M, Norden C, Durrer K, Rauter H, Uhlmann F and Barral Y (2009).
"A Mechanism for Chromosome Segregation Sensing by the NoCut Checkpoint."
Nat Cell Biol 11(4):477-83 (2009). (*)

(*) This publication is the result of the work of Dr. Manuel Mendoza at the Institute of Biochemistry, Biology Department, ETH Zurich, Zurich, Switzerland.

CELL AND DEVELOPMENTAL BIOLOGY

Group: Biomechanics of Morphogenesis

Group structure:

Group leader: Jérôme Solon (from September 2009)

Technician: Eulàlia Belloc (from October 2009)

SUMMARY

A developing organism undergoes dramatic tissue reshaping and rearrangements. These tissue movements require the precise coordination in space and time of hundreds of cells. This coordination is achieved by the strong interplay between expression of regulatory genes and mechanical forces exerted by the cells. Our goal is to reveal the mechanisms driving tissue rearrangements during morphogenesis.

RESEARCH PROJECTS

1. Mechanisms driving tissue constriction and fusion during *Drosophila* development at the cellular scale.

We are focusing on tissue constriction, a major morphogenetic process occurring several times during the development of an organism. This consists in the apical constriction of an acto-myosin meshwork. During Drosophila embryogenesis, tissue constriction leads to the fusion and sealing of the embryo's epidermis in a process called dorsal closure (DC). Intriguingly, dorsal closure presents many similarities with wound healing processes in humans. My group is interested in revealing the mechanisms driving DC.

DC consists in the closing of a gap in the epidermis on the dorsal side of the embryo (Fig 1 A). The process takes place after the germ band retraction. It starts with the combined contraction of a monolayer of cells covering the gap, the amnioserosa tissue, with the reinforcement of an actin cable surrounding the contracting tissue. Eventually, at the end of DC, once the two epidermal layers are close enough, they will fuse with a zippering occurring at the two canthis of the opening. The interplay between these three forces, amnioserosa constriction, actin cable reinforcement and filopodia zippering, and their regulation are still poorly understood.

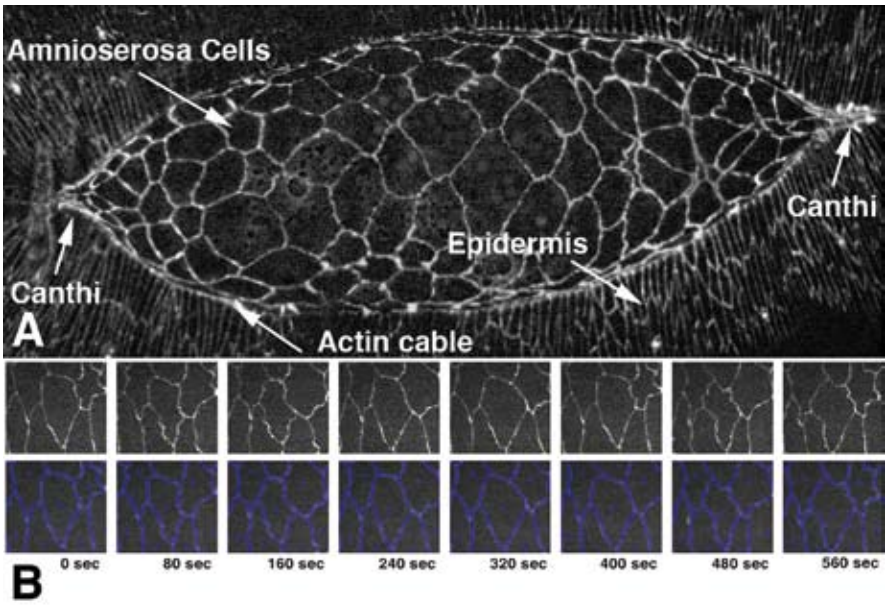


Fig. 1.
(A) Armadillo-GFP highlighting cell membranes during Drosophila dorsal closure. The epidermal tissues converge from the lateral part (top and bottom of the picture) to the dorsal part (in the center of the picture) of the embryo. The dorsal part is covered with amnioserosa cells, which provide the initial force for tissue movement. (B) Typical apical surface area pulsations of an AS cell in a GFP-Arm expressing embryo. The upper panel shows raw data, the lower panel shows the superimposed segmented image. (extracted from Solon et al, 2009).

Recently, we found that the progression of the epidermis toward the dorsal part of the embryo is due to complex pulsed contractions of the amnioserosa cells coupled with the reinforcement of the actin cable, stabilizing the whole structure in a ratchet-like manner (Fig 1 B).

We want to investigate the molecular regulation of the amnioserosa contractions, what are the mechanisms triggering and regulating the amnioserosa cells, and what molecular mechanisms translate these contractions into tissue movement?

2. Spatial and temporal coordination of the cellular contractions in the amnioserosa tissue.

We are also interested in establishing the connection between the intracellular contractions and the tissue organization and constriction. The amnioserosa tissue is composed of about hundred cells; each of them contracts individually. How does this sum of individual contractions generate a global constriction of the tissue and how does the global shape and tension of the tissue influence the individual cell contractions? We are addressing these questions by combining different imaging techniques available at the ALMU (spinning disk, confocal and two-photon microscopy) with automated image analysis and physical modeling.

PUBLICATIONS

Solon J, Kaya-Copur A, Colombelli J, Brunner D.
"Pulsed forces timed by a ratchet-like mechanism drive directed tissue movement during dorsal closure."
Cell 137(7):1331-42. (2009). (*)

(*) This publication is the result of the work of Dr Jérôme Solon at the European Molecular Biology Laboratory, Heidelberg, Germany.

SYSTEMS BIOLOGY

Coordinator: Luis Serrano

During 2009 we have reached the total capacity of the program with 6 group leaders (5 of them part of the CRG-MICINN-EMBL agreement). Now we have a vibrant unit, completely international, made up of 4 young group leaders, a senior scientist and the program director. EMBL/CRG Systems Biology Research Unit covers a wide range of topics in systems biology, providing a rich and stimulating environment. Our targets for the next years, apart from doing science of international quality, will be to create an annual program of workshops about Systems Biology specific topics, in order to create systems biology curricula and to train our doctorate students and our pre-doctoral fellows in simulation tools and mathematical approaches. We will also strengthen our collaborations inside the institute and we will establish an interaction network with other systems biology institutes around Europe.

SYSTEMS BIOLOGY

Group: Design of Biological Systems
Luis Serrano has a Senior ICREA Group Leader position.

Group Structure:

Group Leader: Luis Serrano

Staff Scientists: Christina Kiel, François Stricher

Postdoctoral Fellows: Tobias Maier, Raik Grünberg, Eva Yus, Almer Van der Sloot, Alejandro Nadra, Andreu Alibés, Kostas Michalodimitrakis, Cedrik Magis

Students: Marc Güell, Anne Campagna, Judith Wodke, Bernhard Paetzold, Kiana Toufighi (compartida con Ben Lehner), Erik Verschueren

Technicians: Justine Leigh-Bell, Sira Martinez, Tony Ferrar, Clair Portugal

SUMMARY

Our group is interested in the rational engineering of living systems (ranging from gene networks to organisms). For this purpose we use a combination of tools that involve software for protein design and simulations of networks and experimental approaches. Our approach is based on first understanding a system and then engineering it to obtain the properties we want. Our philosophy is also whenever possible identifying the possible practical applications for human health and biotechnology of our work.

RESEARCH PROJECTS

1. Quantitative understanding of *M. pneumoniae*

The idea of harnessing living organisms for treating human diseases is not new and has been in fiction books since a long time ago. So far the majority of the living vectors used in human therapy are viruses, which have the disadvantage of the limited number of genes and networks that can contain. Bacteria have the advantage of allowing the cloning of complex networks and the possibility of making a large plethora of compounds either naturally or through careful re-design. One of the main limitations for the use of bacteria to treat human diseases is their complexity, the existence of a cell wall that difficult the communication with the target cells, the lack of control over its growth and the immune response that will elicit on its target. Ideally one would like to have a very small bacterium (of the size of a mitochondria), with no cell wall, which could be grown “In Vitro”, could be genetically manipulated, for which we will have enough data to allow a complete understanding of its behavior and which could live as a human cell parasite. Such a microorganism could in principle be used as a living vector in which genes of interests, or networks producing organic molecules of medical relevance, could be introduced under “In Vitro” conditions and then inoculated either on extracted human cells or in the organism, and then become a new organelle in the host. Once the living vector enters inside the host cells it could then produce and secrete into the host proteins which will be needed to correct a genetic disease, or drugs needed by the patient. Putting it into engineering terms, the living vector will be alike a processor which will have a complicated set of instructions and circuits but will only communicate with the host through input and output outlets. Thus the processor could be reprogrammed but the interface with the hosting cell will remain the same. For some particular applications it will not be needed to integrate the bacteria as an organelle, but rather have it inside the cell for a limited amount of time to achieve its goal and then eliminate it by antibiotic treatment for example.

In order to achieve the above goals we need to understand in excruciating detail the Biology of the target bacterium as well as how to interface with the host cell cycle (Systems biology aspect). Then we need to have the engineering tools (network design, protein design, simulations …) in order to modify the target bacterium to behave like an organelle once inside the cell (Synthetic biology aspect). Thus this project has two objectives:

a) Obtain a complete quantitative understanding of a free-living organism (a bacterium in this case).

b) Engineer the bacterium to enter into a mammalian cell line, adapt to the host so as to keep a fixed number of bacteria per host, respond to the host environment and secrete into the host proteins or organic molecules that will provide missing functionalities.

2. Signal Transduction and Disease

Cancer therapy

The efficacy of current treatments for some types of solid tumours is disappointingly poor. Thus, new therapies using novel tumour-selective anti-cancer agents are necessary. A major aim of anti-tumour therapies is to inhibit proliferation and induce death of tumour cells without affecting normal cells. In this regard, members of TNF ligand/receptor family are of interest since they regulate both apoptosis and cell proliferation. One TNF family member, TRAIL, is of particular interest since it selectively induces death of tumour cells without affecting normal cells. Currently, TRAIL and TRAIL-specific antibodies are being investigated as anti-cancer agents. However, one drawback to their efficacy is that they bind to multiple receptors, not all of which transduce an apoptotic signal. Previously, we developed DR5 receptor-selective TRAIL variants, which are potent inducers of apoptosis in various tumour cells, are more efficacious than native TRAIL, and display synergistic effect in combination with other chemotherapy treatments or radiotherapy.

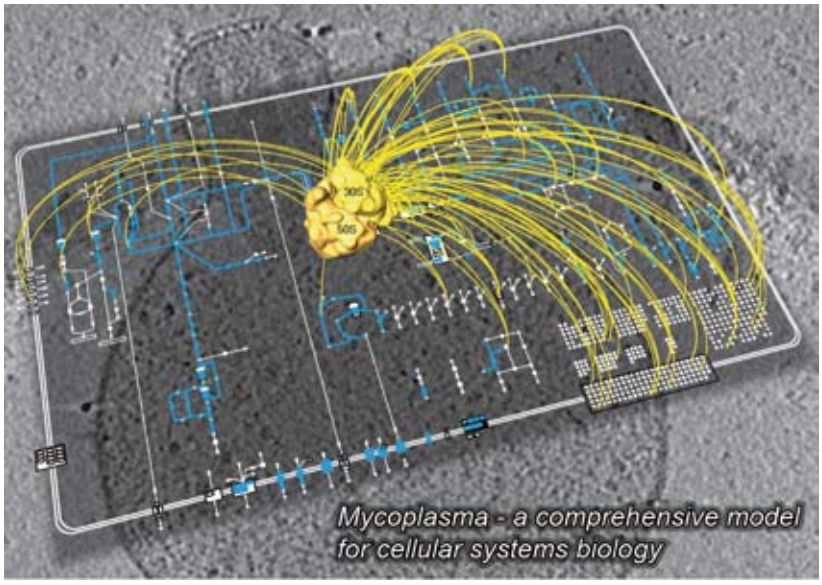


Fig. 1.
Diagramme showing the involvement of different ribosomal proteins in complexes with metabolic enzymes. In yellow the ibosome and the links to metabolic proteins represented as white dots on M. pneumoniae metabolic map.

Structural Systems Biology (Understanding Signal Transduction)

Understanding signal transduction pathways is capital for human health. Current efforts to do so involve knock out experiments and the use of small molecules or antibodies to interfere with selected protein-protein interactions along the pathway. However, signal transduction pathways in higher eukaryotes are characterized by the existence of multiple interactions for any of its particular components. Therefore, it is difficult to elucidate when knocking out a protein or blocking an activity with a drug what is the relationship between phenotype and the interaction affected. Moreover, there is ample evidence that the majority of the drugs in the market are not as specific as thought and they block more than one activity. Here we want to use a different tool that could allow the specific selection and activation, or inhibition, of specific routes in a pathway and also provide much higher specificity than small molecules. This tool is rational protein design. Structures of macromolecules and especially of molecular machines in combination with protein design could provide quantitative parameters, help to elucidate functional networks, or allow rational designed perturbation experiments for reverse engineering. Computer aided protein design has shown to be able to modify in a rational way protein-protein interactions, tuning specificity, affinity and interestingly kinetics of binding. Protein design can also be used to predict using structural information the partners of a particular domain, or protein.

As a scientific target we have selected the MAPK pathway and our final goal is to obtain a global quantitative understanding with the idea of designing better therapies for diseases involving its deregulation. In 2009 we published a paper where we quantitatively analyze the effect of mutation in Ras Raf interaction during ERK phosphorylation (Kiel, C. & Serrano, L., Cell type-specific importance of ras-c-raf complex association rate constants for MAPK signaling. Sci Signal 2 (81), ra38 (2009)) (Figure 2).

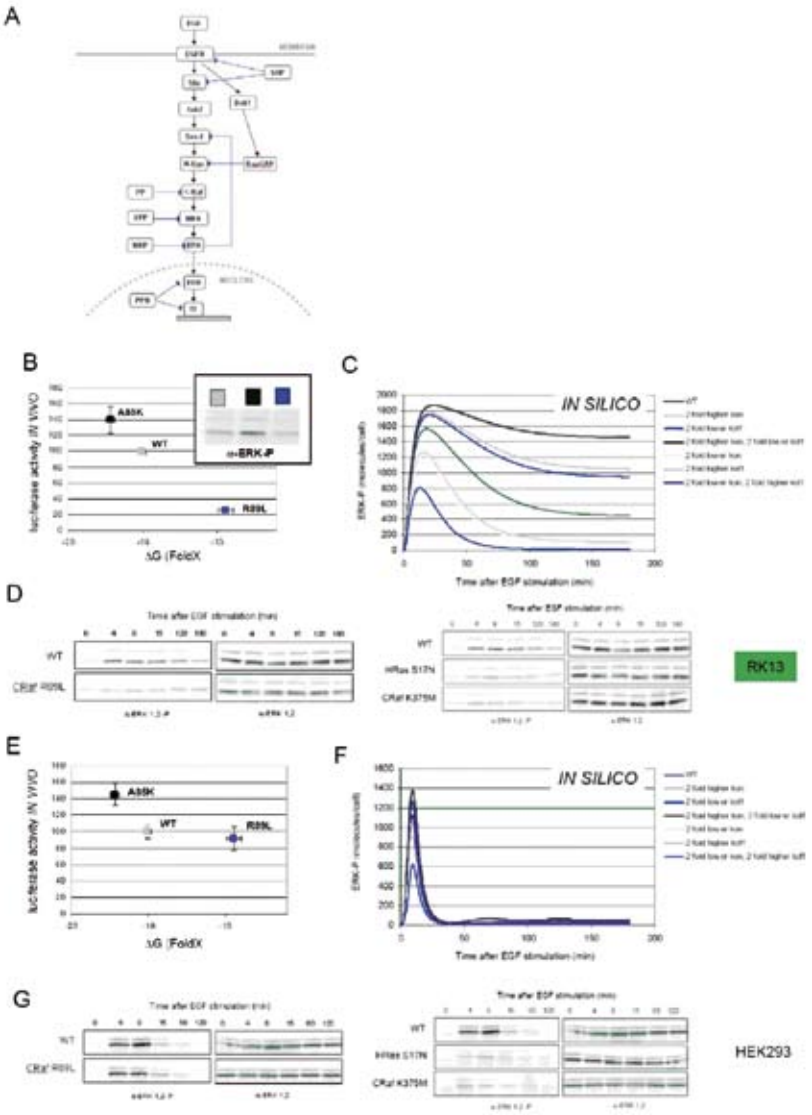


Fig. 2.

Experimental analysis and computer modeling in RK13 and HEK293 cells. (A) Flow chart representation of the EGF-Ras-MAP kinase pathway as used for model generation. (B) and (E) Comparing predicted affinities with luciferase activities in RK13 and HEK293 cells for CRaf WT and two mutants with higher and lower affinity, respectively (A85K, R89L). (D) and (G) ERK-phosphorylation kinetics in RK13 and HEK293 cell, which have been transiently transfected with CRaf WT or R89L, starved, and stimulated with EGF. (C) and (F) Comparing the effect of changing kinetic rate constants for the Ras-Raf complex on the simulated ERK-phosphorylation changes after EGF stimulation in the HEK-like and the RK13-like model.

PUBLICATIONS

Yus E et al.
"Impact of genome reduction on bacterial metabolism and its regulation."
Science 326(5957):1263-1268 (2009).

Vanhee P et al.
"Protein-peptide interactions adopt the same structural motifs as monomeric protein folds."
Structure 17(8):1128-1136 (2009).

Vanhee P et al.
"PepX: a structural database of non-redundant protein-peptide complexes."
Nucleic Acids Res, Epub 2009 Oct 30.

Van der Sloot AM, Kiel C, Serrano L and Stricher F.
"Protein design in biological networks: from manipulating the input to modifying the output."
Protein Eng Des Sel 22 (9):537-542 (2009).

Sanchez-Hidalgo M et al.
"Conformational Stability and Activity of Circular Enterocin AS-48 Derivatives."
Protein Pept Lett. 2009 Oct 7. [Epub ahead of print]

Reis CR et al.
"Enhancement of antitumor properties of rhTRAIL by affinity increase toward its death receptors."
Biochemistry 48(10):2180-2191 (2009).

Petukhov M et al.
"Design of stable alpha-helices using global sequence optimization."
J Pept Sci 15(5):359-365 (2009).

Markson G et al.
"Analysis of the human E2 ubiquitin conjugating enzyme protein interaction network."
Genome Res 19(10):1905-1911 (2009).

Maier T, Guell M and Serrano L.
"Correlation of mRNA and protein in complex biological samples."
FEBS Lett 583(24):3966-3973 (2009).

Liew CW et al.
"Interaction of the human somatostatin receptor 3 with the multiple PDZ domain protein MUPP1 enables somatostatin to control permeability of epithelial tight junctions."
FEBS Lett 583(1):49-54 (2009).

Kuhner S et al.
"Proteome organization in a genome-reduced bacterium."
Science 326(5957):1235-1240 (2009).

Kiel C and Serrano L.
"Cell type-specific importance of ras-c-raf complex association rate constants for MAPK signaling."
Sci Signal 2(81):ra38 (2009).

Guell M et al.
"Transcriptome complexity in a genome-reduced bacterium."
Science 326(5957):1268-1271 (2009).

Fernandez-Ballester G et al.
"Structure-based prediction of the Saccharomyces cerevisiae SH3-ligand interactions."
J Mol Biol 388(4):902-916 (2009).

Encinar JA et al.
"ADAN: a database for prediction of protein-protein interaction of modular domains mediated by linear motifs."
Bioinformatics 25(18):2418-2424 (2009).

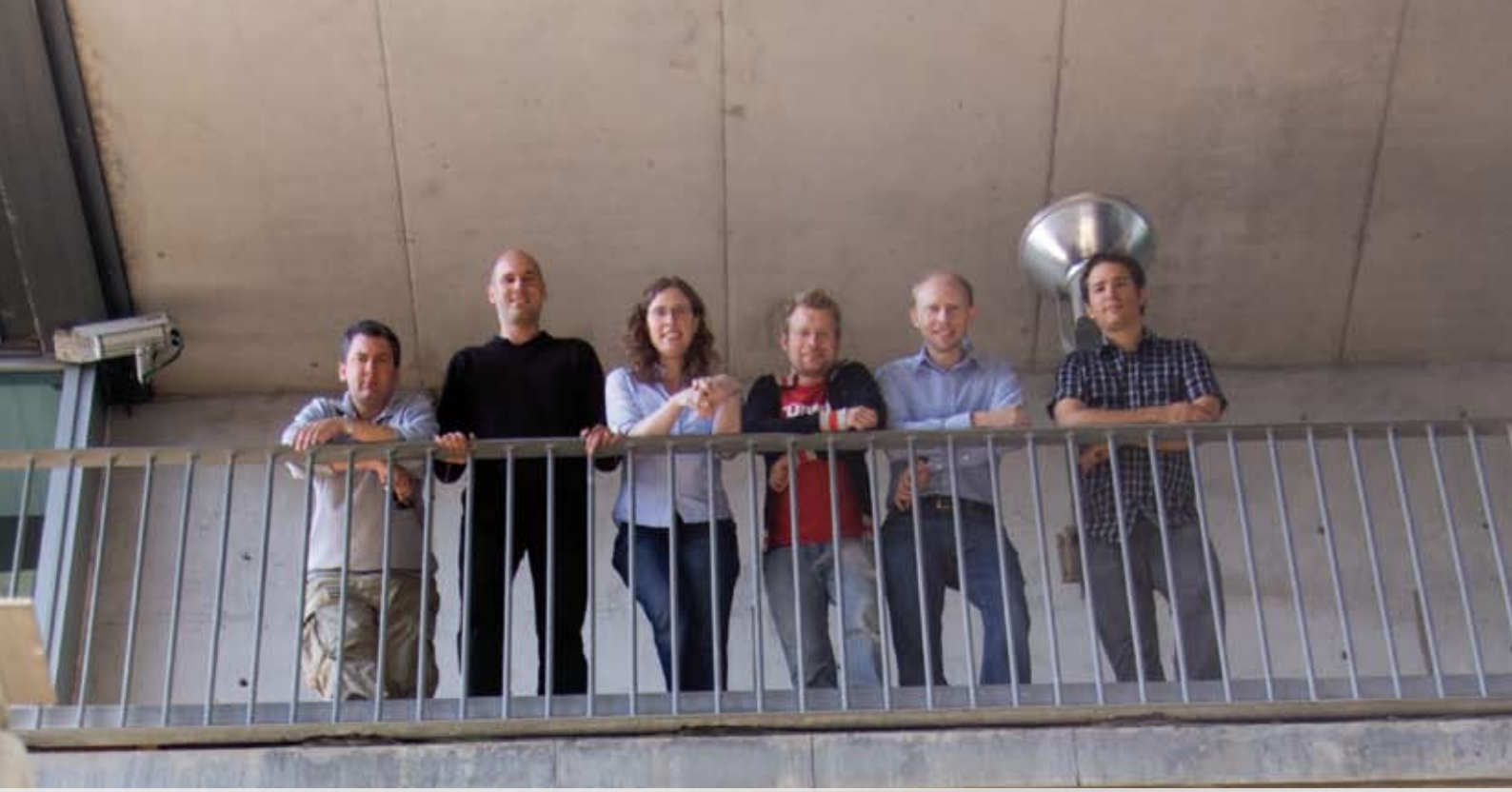
Maurer-Stroh S, Debulpaep M, Kuemmerer N, de la Paz ML, Martins IC, Reumers J, Morris KL, Copland A, Serpell L, Serrano L, Schymkowitz JW, Rousseau F.
"Exploring the sequence determinants of amyloid structure using position-specific scoring matrices."
Nat Methods, in press.

Kiel C, Yus E, Serrano L.
"Engineering signal transduction pathways."
Cell, in press. Review.

Kiel C, Filchtinski D, Spoerner M, Schreiber G, Kalbitzer HR, Herrmann C.
"Improved binding of Raf to Ras-GDP is correlated with biological activity."
J Biol Chem 284(46):31893-902 (2009).

Bouvier B, Grünberg R, Nilges M, Cazals F.
"Shelling the Voronoi interface of protein-protein complexes reveals patterns of residue conservation, dynamics, and composition."
Proteins 76(3):677-92 (2009).

Michalodimitrakis K, Isalan M.
"Engineering prokaryotic gene circuits."
FEMS Microbiol Rev 33(1):27-37 (2009). Review.



SYSTEMS BIOLOGY

Group: Systems Analysis of Development
James Sharpe has a Senior ICREA Group Leader position.

Group structure:

Group Leader:	James Sharpe
Staff Scientist:	Jim Swoger
Postdoctoral Fellows:	Jean-François Colas, Henrik Westerberg, James Cotterell, Niamh Nowlan
Technician:	Laura Quintana
PhD Students:	Bernd Boehm, Luciano Marcon, Gaja Lesnicar-Pucko, Jelena Raspopovic

SUMMARY

The Sharpe lab has 2 primary goals:

(1) To further our understanding of ***organogenesis as a complex system***, by bringing together a diverse range of techniques from biology, physics, imaging and computer science. The larger part of the lab focuses on a well-characterised standard model of development – the vertebrate limb (using both mouse and chick). For this project we constructing a collection of computer simulations which are based on high-quality quantitative empirical data generated by our own new 3D and 4D imaging technologies. A smaller but equally important project within this topic is a more abstract exploration of the patterning potential of gene network motifs.

(2) Building on the success of the 3D imaging technique developed within the lab called Optical Projection Tomography (OPT – ***Science*** 296:541, 2002), the other major goal of the lab is to continue developing and improving 3D and 4D imaging technology. Recent success in this direction includes the development of time-lapse OPT imaging of mouse limb development in-vitro (***Nature Methods*** 5:609-12, 2008).

RESEARCH PROJECTS

1. Development of novel 3D and 4D optical imaging technologies

In addition to OPT, which remains a central technology for the lab, we have also now designed and constructed a SPIM apparatus (selective plane illumination microscope), for which Dr. Jim Swoger was one of the original co-inventors (Huiskin et al, 2004 ***Science*** 305:1007). We are currently exploring a wide range of imaging applications for this technology, including samples from mouse, chick, zebrafish and *Drosophila*.

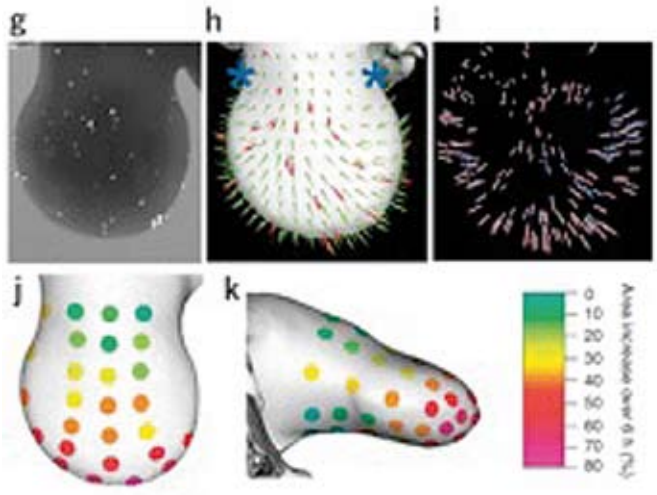


Fig. 1.
New 4D data-capture technologies developed by Sharpe lab (LiveOPT) to provide accurate quantification of limb growth dynamics. (Fig.2 from Boot et al. Nature Methods (2008)).

2. OPT imaging for quantitative assessment of mouse models of diabetes

We have explored many new applications for OPT, and one of the most exciting has been the ability to quantify the number of Islets of Langerhans in an intact adult mouse pancreas in a single scan. Until now this has been performed using the time-consuming approach of traditional histology (cutting hundreds of thin paraffin sections for each pancreas). By contrast, we have demonstrated that the speed of OPT makes it feasible to compare many pancreata in a single study, ***Nature Methods*** (4:31-33, 2007).

During 2009 we started our involvement in a 7FP EU integrated project entitled VIBRANT, with the goal of further improving the resolution of OPT and SPIM. In particular, we aim to use 3D optical imaging as a verification technique for new non-optical in-vivo imaging techniques (such as MRI) being developed within the consortium, to estimate total beta-cell mass within living animals, and ultimately human patients.

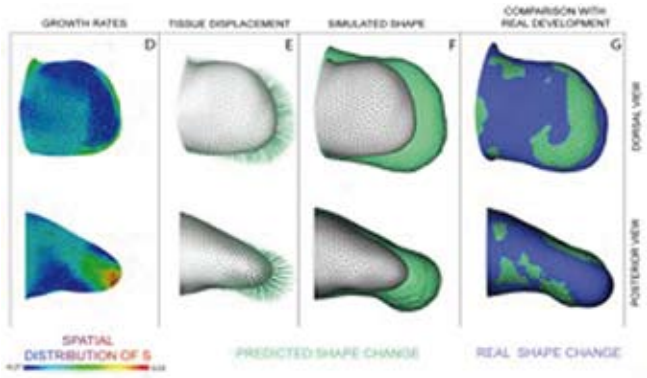


Fig. 2.
Finite element modelling of limb bud growth. *The results of our parameter optimisation are shown, in which a theoretical distribution of proliferation rates was calculated (D) that could successfully generate the correct shape changes (E-G).*

3. 4D Time-lapse Optical Projection Tomography

Optical projection tomography has proven to be a powerful tool for developmental biologists. But until recently OPT has been performed almost exclusively on fixed specimens which have been optically cleared to increase the quality of 3D images obtained. Recently we have explored various improvements to allow the 4D imaging of the developing mouse limb bud in culture (***Nature Methods*** 5:609, 2008), and this technique now provides us with data on tissue movements and dynamic gene expression patterns. During 2009 we made further improvements to this approach (***Organogenesis*** 5:129, 2009) to start mapping internal movements of the mesenchymal tissue.

4. Computer model of limb mechanics

We are exploring the ways in which cell behaviours combine with various physical models of tissue to generate the observed shape changes. In particular, it has become clear from early modelling results that explaining phenomena as apparently simple as the dorso-ventral flattening of the limb bud, may be more complicated than previously thought. We have created a finite-element model (FEM) of limb development within which we are exploring different hypotheses. An essential aspect of this project is that we also perform lab-work to generate our own empirical data for the simulation – the model therefore serves as a framework for combining different types of information. While various types of biological material have previously been mechanically modelled over a short time periods (for example stress analysis on bones and cartilage) a mathematical/physical description of 3D embryonic tissue displaying volumetric growth over a period of hours or days has not previously been achieved, and this is therefore one of the general goals of this project.

5. Computer models of spatially-patterning gene networks

Enough is known about the genetics of limb development to be sure that it involves many signalling molecules (Shh, BMPs, FGFs, Wnts) and many transcription factors (Msx, Hox, Meis) which are wired together into a complex gene network. We are currently exploring how these networks function within the computer model – both in control of the cell behaviours that govern the limb bud shape, and also in another famous patterning case: the spatial organisation of the skeletal elements.

For skeletal patterning, different patterning strategies have been proposed within the literature to explain the process, ranging from pure "Turing-type" reaction-diffusion models at one extreme, to morphogen gradient models on the other. Using our new visualisation tools (above) are exploring the ways that these different principles might operate within the limb bud. We are now extending this analysis into a 3D simulation within the context of the correct limb bud shape, and we are therefore using OPT and confocal microscopy to gather accurate 3D expression data on genes thought to be involved in, for example, the initiation of mesenchymal condensation.

6. Morphometric Staging of limb buds

Another important project in the lab relates to measurements of time. Almost all the gene expression data which exists about limb development has been recorded from fixed specimens. However, the spatial patterns of certain genes (eg. Sox9) are extremely dynamic – apparently changing hour-by-hour. We have therefore created a new staging system which is based on morphometric measurements from the profile of the limb bud. It can determine the stage of a limb to a high temporal accuracy, and we have turned this method into a web-based java application that will soon be openly available to the whole limb community.

PUBLICATIONS

Vendrell V, Summerhurst K, Sharpe J, Davidson D, Murphy P.
"Gene expression analysis of canonical Wnt pathway transcriptional regulators during early morphogenesis of the facial region in the mouse embryo."
Gene Expr Patterns 9(5):296-305 (2009).

Hajihosseini MK, Duarte R, Pegrum J, Donjacour A, Lana-Elola E, Rice DP, Sharpe J, Dickson C.
"Evidence that Fgf10 contributes to the skeletal and visceral defects of an apert syndrome mouse model."
Dev Dynam 238(2):376-85 (2009).

Colas JF, Sharpe J.
"Live OPT."
Organogenesis 5(4):129-42 (2009).

Lemos MC, Harding B, Reed AA, Jeyabalan J, Walls GV, Bowl MR, Sharpe J, Wedden S, Moss J, Ross A, Davidson S, Thakker RV.
"Genetic background influences embryonic lethality and the occurrence of neural tube defects in Men1 null mice: relevance to genetic modifiers."
J Endocrinol 203(1):133-42 (2009).

SYSTEMS BIOLOGY

Group: Gene Network Engineering
This group is part of the EMBL/CRG Research Unit in Systems Biology

Group structure:

Group Leader: Mark Isalan

Postdoctoral Fellows: Mireia Garriga, Emmanuel Fajardo, Frank Herrmann, Phil Sanders

Students: Andreia Carvalho, Marco Constante

Technician: Rebecca Baumstark

SUMMARY

In order to test our understanding of biological systems, one approach is to attempt to build synthetic biological networks. This allows us to see whether they behave as we would predict - and whether we truly understand them. Such "synthetic" approaches can be done either by engineering small-scale artificial gene networks, or by tinkering with large-scale networks. Even small networks are not necessarily straightforward. For example, "Gene A activates gene B; gene B represses gene A." can be as confusing as "The next statement is true. The previous statement is false." Therefore, we have spent some time developing logical frameworks for thinking about such systems (see Isalan et al. ***Nature*** 458:969, 2009; ***Bioessays*** 31 (10): 1110-5, 2009). For larger networks, our recent work on rewiring bacterial gene networks at highly-connected hub genes, answered the question, "what happens if you add lots of new links to an existing biological network?" The surprising result was that even highly-connected genes could be "rewired" and were generally very well-tolerated by the bacteria (Isalan et al. ***Nature*** 452:840-5, 2008). Overall, as well as building artificial gene networks in cells, our group also develops the basic technology to achieve this in mammalian cells (zinc finger-based genome engineering and gene repair).

RESEARCH PROJECTS

1. Liar Paradoxes and gene networks

Mark Isalan and Matthew Morrison (collaborator).

Much of 2009 was spent on two articles that combined ancient Greek philosophy with the study of gene networks. We developed a collaboration with a pure philosophy department, and discussed how genetic interactions such as "Gene A represses itself" or "Gene A makes gene B. Gene B represses gene A." are very similar to the classical liar paradoxes, "This statement is false" or "The next statement is true. The previous statement is false." The result is that by adding time or spatial dimensions explicitly, oscillations, waves or stripes can intuitively form through these types of interaction ("it is true then false then true then false"). Verbal analogies are thus a way of communicating intuitively to a non-mathematical audience why a Turing-pattern makes spots or stripes on the skin of a discus fish, or why the p53 tumour suppressor protein oscillates.

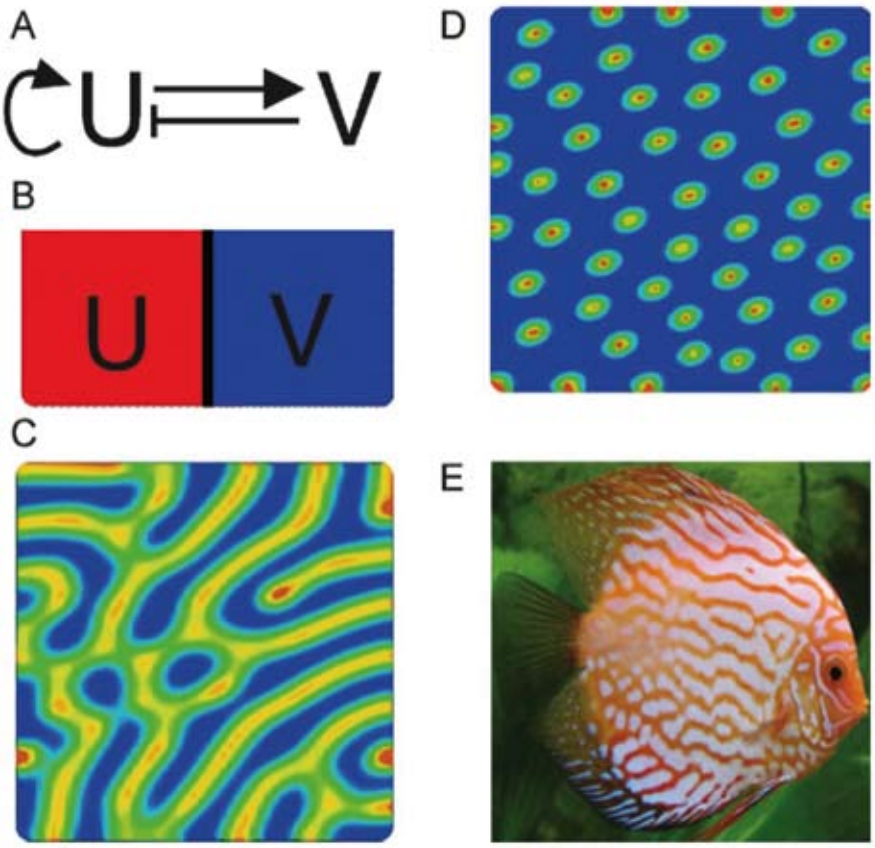


Figure 1.
From liar paradoxes to Turing patterns. A: 'The next statement is true. The previous statement is false' is a zero-dimensional analogue of a stripe-pattern forming system, based on an activator U and inhibitor V (2-D space and 1-D time). Such systems rely on reaction and diffusion to make patterns (local autoactivation, long-range inhibition. B: Rainbow colour coding shows local high expression of activator U (more red) and inhibitor V (more blue). With high autoactivation and weak repression, the activator dominates giving a uniform stability (red). With high transactivation and strong repression, the inhibitor dominates (blue). C: By balancing production, diffusion, reaction and degradation, spots, stripes or waves spontaneously emerge. D: This simulation has the same parameters as (C), except with a lower autocatalysis saturation constant, resulting in spots. E: The discus fish is thought to employ just such a body-patterning mechanism, based on reaction-diffusion. Images and calculations kindly provided by Luciano Marcon and James Sharpe. Discus image by Anka Zolnierzak

We managed to communicate this liar paradox analogy to a wide audience through a non-technical essay in Nature (Nature 458:969 (2009)), as well as expanding the ideas in a longer essay for Bioessays (Gene networks and Liar Paradoxes Bioessays 31 (10): 1110-5 (2009)), where the ideas were even illustrated on the cover.

2. Engineering pattern forming gene networks in eukaryotic cells

Mireia Garriga, Phil Sanders, Marco Constante and Andreia Carvalho

We are carrying out several spatial and temporal pattern-engineering projects in parallel, corresponding to different scales of cellular organisation. These start from engineering localised gradients and patterns in single cells and move towards designing gene networks using diffusible factors that operate over fields of cells.

In 2009, major progress was achieved in building a synthetic gene network that secretes a diffusing 'morphogen' molecule that then activates distal cells in a 3D-model of morphogenesis: the MDCK (Madine-Darby Canine Kidney Cell) cyst system. Briefly, MDCK cells spontaneously form hollow spheres of cells in 3D culture and tubulate upon exposure to HGF. By building an HGF-secretion and reaction-diffusion system, with suitable reporters, we aim to see if we can engineer a synthetic developmental patterning system in 3D.

Fig. 2.
A synthetic gene network for studying morphogen gradient-type pattern formation. Cells locally express HGF (red regions), and then secrete HGF extracellularly. HGF diffuses and activates a reporter gene (green - GFP) in distal cells. The HGF also has a cell scattering activity in 2D culture, or a tubule-formation-inducing activity in 3D culture. The aim is to measure and control diffusion and HGF-activity to convert this cell-cell communication into a controlled, self organising pattern.

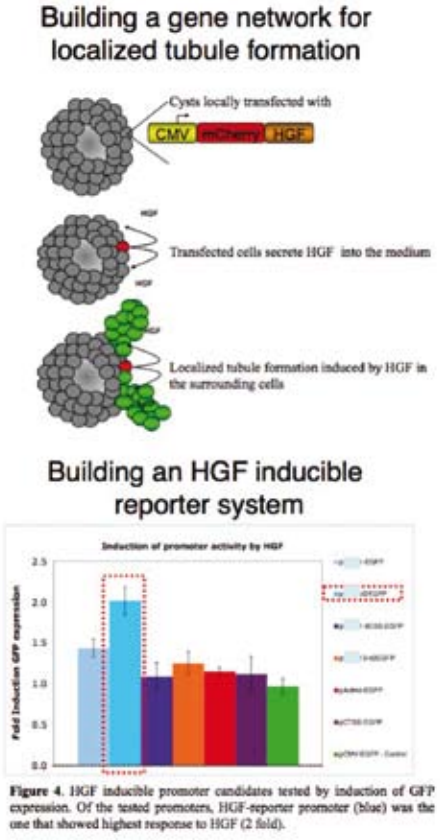


Figure 4. HGF inducible promoter candidates tested by induction of GFP expression. Of the tested promoters, HGF-reporter promoter (blue) was the one that showed highest response to HGF (2 fold).

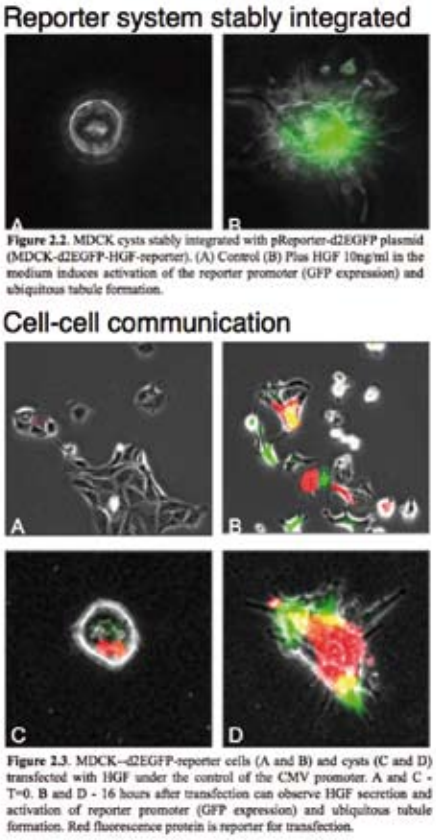


Figure 2.2. MDCK cysts stably integrated with pReporter-d2EGFP plasmid (MDCK-d2EGFP-HGF-reporter). (A) Control (B) Plus HGF 10ng/ml in the medium induces activation of the reporter promoter (GFP expression) and ubiquitous tubule formation.

Figure 2.3. MDCK--d2EGFP-reporter cells (A and B) and cysts (C and D) transfected with HGF under the control of the CMV promoter. A and C - T=0. B and D - 16 hours after transfection can observe HGF secretion and activation of reporter promoter (GFP expression) and ubiquitous tubule formation. Red fluorescence protein is reporter for transfection.

3. Synthesising zinc fingers for genome engineering and gene repair

Frank Herrmann, Rebecca Baumstark and Emmanuel Fajardo

As part of three EU-funded projects (Netsensor, EC Contract No. 012948 http://netsensor.embl.de/; Integra, EC Contract No. FP6 - 29025; FP7 ERC Zinc-Hubs) we are building a number of artificial sequence-specific DNA-binding proteins using our established protocol (Isalan, M., Klug, A. & Choo, Y. *Nature Biotechnology*, 19, 656-60; 2001). We aim to develop upon the recent reports of endogenous gene repair using zinc finger nucleases (Bibikova et al. *Science* 300, 764, 2003; Urnov et al., *Nature* 435, 646-51, 2005; Fig 2).

During 2009, our main pilot project, gene repair of p53 (>50% of all human cancers have mutations in p53) reached a final phase. Our engineered zinc finger nucleases for cancer hot-spots (z771 and z1166) achieved gene repair both episomally and chromosomally. In 2010, we will precisely quantitate the rate of repair events using solexa ultrasequencing technology.

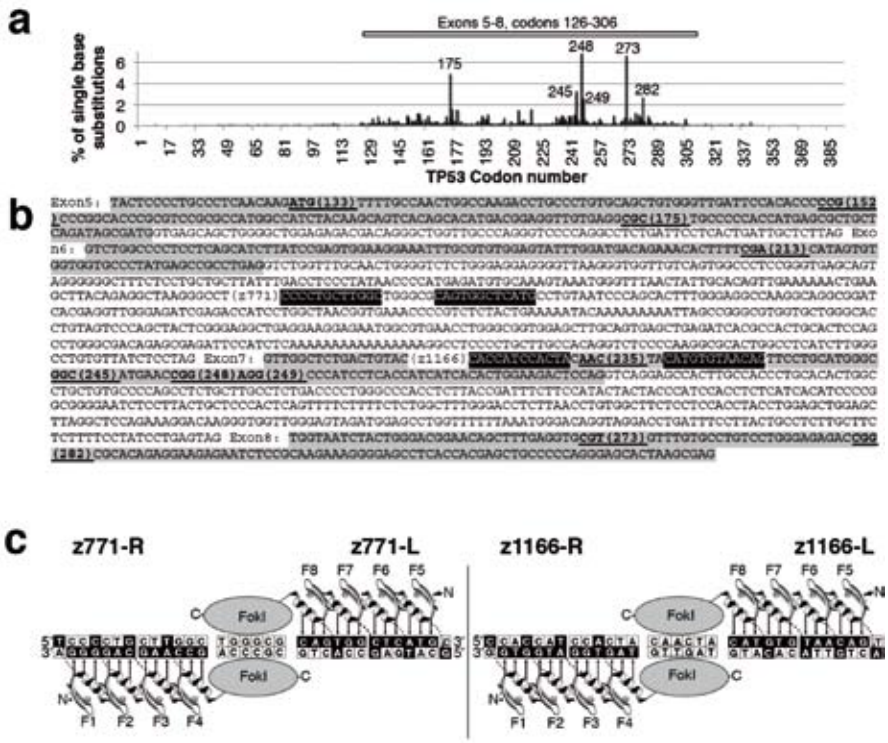


Fig. 3.
Zinc finger nucleases (ZFN) for the human p53 gene. a, Mutation hotspots in somatic cancers from the IARC TP53 Mutation Database R13. b, Exons 5-8 from the p53 gene (Human Genome: NW_001838403) are highlighted in grey and contain nearly all mutation hotspots (underlined black; codon number in brackets). ZFN binding sites are highlighted in black with white letters. c, Canonical model of designed zinc finger nucleases (z771 and z1166) that we engineered against the best two target sites in the p53 gene. Arrows indicate possible base contacts (F1, Finger 1, etc.).

PUBLICATIONS

Isalan M and Morrison M.
"This title is false."
Nature 458:969 (2009).

Isalan M.
"Gene networks and liar paradoxes."
Bioessays 31(10):1110-5 (2009).

Michalodimitrakis K and Isalan M.
"Engineering prokaryotic gene circuits FEMS Microbiol."
FEMS Microbiol Rev. 33(1):27-37 (2009). Review.

Group:	Genetic Systems This group is part of the EMBL/CRG Research Unit in Systems Biology
Group structure: Group Leader:	Ben Lehner, ICREA Researcher
Postdoctoral researchers (experimental):	Jennifer Semple, Olivia Casanueva
Postdoctoral researchers (computational):	Tanya Vavouri, Rob Jelier, Benjamin Schuster-Böckler
PhD students (experimental):	Alejandro Burga Ramos, Angela Krüger, Adam Klosin
Technician (experimental):	Rosa Garcia-Verdugo
Visiting PhD student (computational):	Mirko Francesconi
Masters student (computational):	Said Aktas

SYSTEMS BIOLOGY

SUMMARY

We use experiments and computational analysis to answer basic questions in genetics. These include, but are not limited to the following. How do mutations combine to cause disease? Why are genes harmful when they are overexpressed? Can we make useful personal phenotypic predictions from personal genome sequences? Can we predict all of the genes important for any process or disease? Can we predict how mutations in two genes combine to cause disease? Why do the effects of mutations vary between cell types? Can we understand which non-coding mutations are most likely to be detrimental? How can we fully quantify the effects of mutations on the phenotype of an organism? How does gene regulation evolve? How does noise influence evolution? Our work involves three aspects: computational analysis of existing data, quantitative experiments in ***C. elegans*** and yeast, and the development of new methodology when required.

In particular, this year we published two studies where considering expression and concentration in cellular networks helped us to understand some basic questions in genetics.

RESEARCH PROJECTS

1. Why are genes harmful when their expression is increased?

By testing possible causes of dosage sensitivity in yeast, we identified intrinsic protein disorder as an important determinant of dosage sensitivity [1]. Disordered regions are prone to make promiscuous molecular interactions when their concentration is increased, and we demonstrated that this is a likely cause of pathology when genes are overexpressed. We showed that our model is also consistent with data from two animals, ***Drosophila melanogaster*** and ***Caenorhabditis elegans***.

In mice and humans we found that the same properties are strongly associated with dosage-sensitive oncogenes, such that promiscuous non-physiological mass-action-driven molecular interactions may be relevant mechanistically in cancer.

Further, we found that dosage-sensitive genes are tightly regulated at the transcriptional, RNA, and protein levels, which may serve to prevent harmful increases in their concentration under physiological conditions.

In summary, we proposed mass-action-driven interaction promiscuity as a theoretical framework to be used to understand, predict, and possibly treat the effects of increased gene expression in evolution and disease.

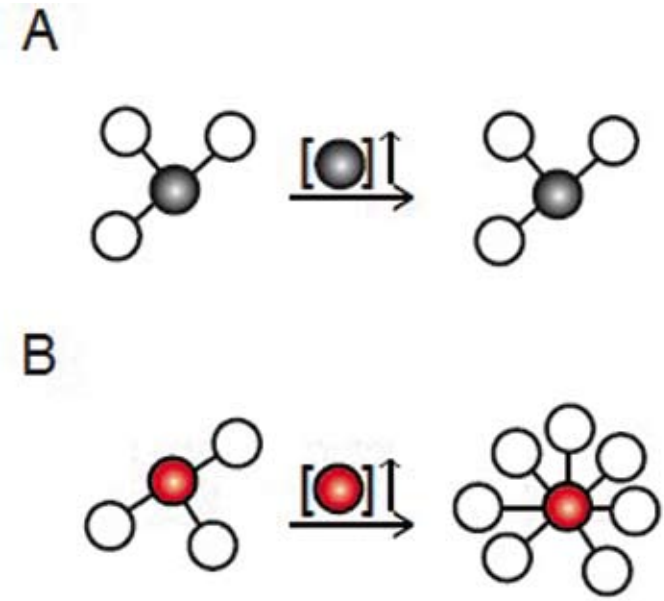


Fig 1.
Interaction promiscuity: proteins that interact through large globular domains and with few "off-target" interactions tend not to be dosage-sensitive (A). In contrast, increasing the expression of a protein that interacts through short, degenerate linear motifs will result in promiscuous non-physiological interactions, many of which may be detrimental (see ref [1]).

2. How do mutations have cell-type specific effects, even if a protein is ubiquitously expressed?

By integrating human protein-protein interaction and expression data we addressed several questions about the relationship between the expression of a protein and its physical interactions [2]. We found that most (and possibly all) universally expressed "housekeeping" proteins have different interaction partners in different cell types. Thus most housekeeping proteins are actually important for many tissue-specific functions, which may be an important reason why mutations in these proteins generally have quite tissue-specific effects.

Further, we found that most proteins that are only expressed in restricted cell types contact the core cellular machinery, and so can function by recruiting or modifying core cellular processes. We also observed that tissue-specific proteins tend to be recent evolutionary innovations and to have fewer known interaction partners. Taken together, our findings suggest that many tissue-specific processes evolve by the addition of metazoan-specific proteins that make a limited number of interactions with core cellular components.

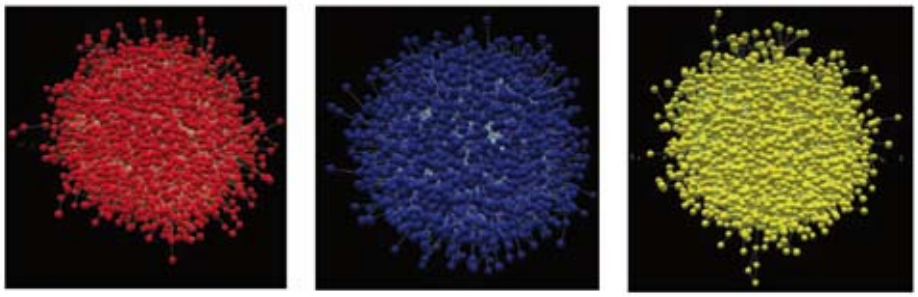


Fig 2.
Maps of possible interactions in three human cell types: skeletal muscle, B-cells, and pancreatic islet cells (see ref [2]).

3. Other work

In addition, in collaboration with the group of Erik Meijering (Erasmus MC) we described some preliminary work developing methods for automated and quantitative phenotyping in ***C. elegans*** [3]. We also further expounded our ideas about the evolution of gene regulatory networks and animal body plans [4].

PUBLICATIONS

Articles

Vavouri T, Semple JI, Garcia-Verdugo R and Lehner B.
"Intrinsic protein disorder and interaction promiscuity are widely associated with dosage sensitivity."
Cell 138:198–208 (2009).

Bossi A, Lehner B.
"Tissue specificity and the human protein interaction network."
Molecular Systems Biology 5:260 (2009).

Dzyubachyk O, Jelier R, Lehner B, Niessen W, Meijering E.
"Model-based approach for tracking embryogenesis in Caenorhabditis elegans fluorescence microscopy data."
Conf Proc IEEE Eng Med Biol Soc1:5356-9 (2009).

Review

Vavouri T, Lehner B.
"Conserved noncoding elements and the evolution of animal body plans."
Bioessays 31(7):727-35 (2009).

SYSTEMS BIOLOGY

Group: Sensory Systems and Behaviour
This group is part of the EMBL/CRG Research Unit in Systems Biology

Group structure:

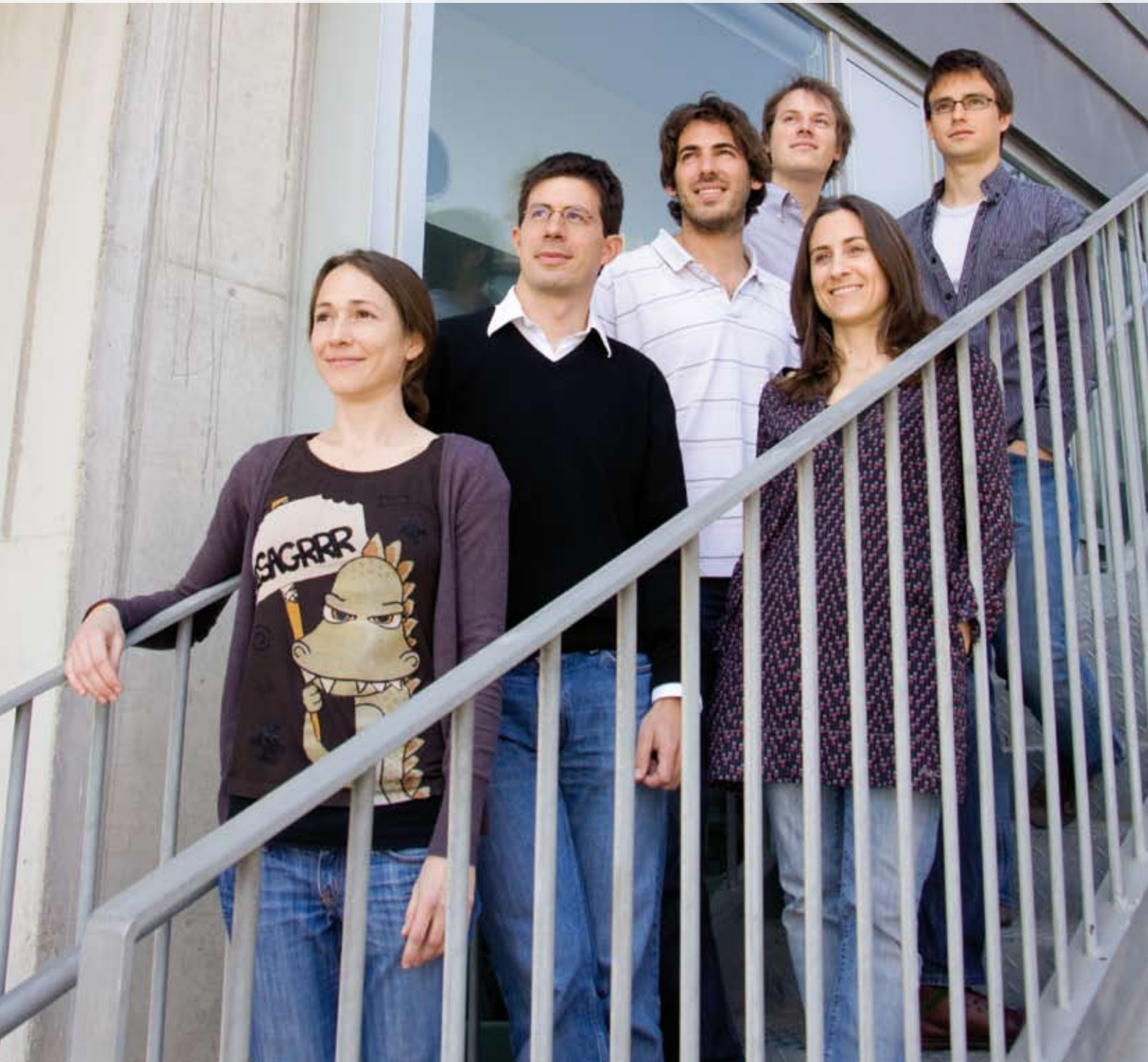
Group leader: Matthieu Louis

Technicians: Verena Hotz, Mariana Lopez-Matas, Moraea Philipps, Jordi Sanchez

Postdoctoral fellow: Alex Gomez-Marin

Students: Julia Riedl, Aljoscha Schulze

Visiting students: Nicolas Partoune, Eliana Pouchard



SUMMARY

The mission of our group is to unravel structure-function relationships between neural circuits, sensory coding and adaptive behaviour. We are interested in understanding how odour tracking comes about in terms of circuit computation. This problem is tackled in the olfactory system of the fruit fly ***Drosophila melanogaster*** larva. Our research combines a variety of experimental and computational techniques aiming to define how naturalistic odorant stimuli are encoded and processed in the olfactory system. In particular, we are developing new computational tools to annotate stereotypical behaviours automatically. We are seeking to identify circuits involved in the processing of olfactory information and the making of decisions underlying chemotaxis.

RESEARCH PROJECTS

1. Peripheral representation of odour stimuli

Mariana Lopez-Matas, Moraea Philipps, Jordi Sanchez, Julia Riedl, Aljoscha Schulze

The larval 'nose' is composed of 21 olfactory sensory neurons (OSNs) expressing typically one type of odorant receptors along with a ubiquitously expressed co-receptor (Or83b) (Fig. 1). Previously, we have demonstrated that individual odorant receptors have overlapping, yet distinct, ligand tuning properties (Asahina, Louis et al., 2009). Accordingly, each OSN can be viewed as distinct information channel to the olfactory system. We have also shown that the information transmitted by a single functional OSN is sufficient for the detection of minute changes in odor intensity. Whether one type of odorant receptors alone is capable of encoding the quality of an odor remains an open question in the larva. We have undertaken to disentangle the contribution of single OSNs to the representation of static and dynamic odor stimuli. In associative learning experiments, we found evidence that a single functional OSN is sufficient to mediate odor quality discrimination. This finding highlights the surprisingly high coding potential of single sensory neurons.

Odours are represented by dynamical patterns of activity originating from the OSNs and transmitted to higher brain centres (Fig. 1). We have devised a preparation to carry out extracellular and intracellular recordings from the peripheral olfactory system. Using this novel technique, we have begun to compare the activity elicited in identified OSNs. Our current goal is to explain the ability (or inability) of larvae with a single functional OSN to discriminate between distinct odours. We are also interested in characterizing how odour-specific activity patterns evolve across different layers of the olfactory system. Functional imaging is applied to monitor the activity in neuronal ensembles. By expressing genetically encoded calcium sensors (GCaMP) in targeted neurons, we are establishing precise correlations between sensory input and circuit activity. We aim to clarify the role of local inhibitory interneurons in the antennal lobe (Fig. 1), a circuit which is thought to achieve gain control and decorrelation. Our long-term goal is to extend this analysis to motor control, and propose a model for how ethologically relevant sensory input (e.g., increase or decrease in the concentration of an attractant) induce specific behavioural output (e.g., orientation towards an odour source).

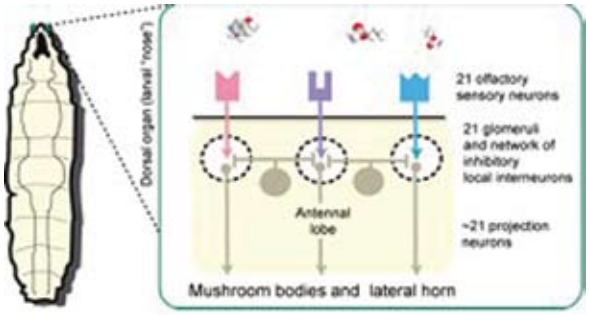


Figure 1.

2. Behavioural algorithms underpinning chemotaxis

Alex Gomez-Marin and Nicolas Partoune

Bacteria and nematodes chemotax according to an ***indirect*** orientation mechanism which consists in biased random walks. In contrast, ***Drosophila*** larvae employ a direct mechanism where the direction of motion is locally aligned with the odour gradient. Previously, we have demonstrated that ***Drosophila*** larvae do not require bilateral olfactory input to perform chemotaxis (Louis, Huber et al., 2008). Having ruled out a mechanism purely based on stereo sampling (comparisons between the left and the right 'noses'), we are now seeking to understand how larvae with unilateral olfactory function are capable of extracting directional information from spatially distributed odour cues.

To address this question, we are dissecting the odour tracking-strategy employed by larvae in detail. We have developed customized software for high-resolution image recognition. These programs extract morphological and kinematic features describing the state of freely moving larvae. Computerized analysis permits us to generate large datasets for wild type and mutant larvae engaged in chemotactic tasks. Using dimensionality reduction methods, we are now investigating quantitative relationships between stimulus input and behavioural output. Our objective is to extract patterns in the stimulus changes which trigger stereotypical actions such as turns, reversals, and lateral head sweeps. We are working on a predictive model that will capture the sensory computation featured in larval chemotaxis. Our long-term goal is to delineate algorithms underpinning sensory-motor responses, and to manipulate these behaviours in a predictable way.

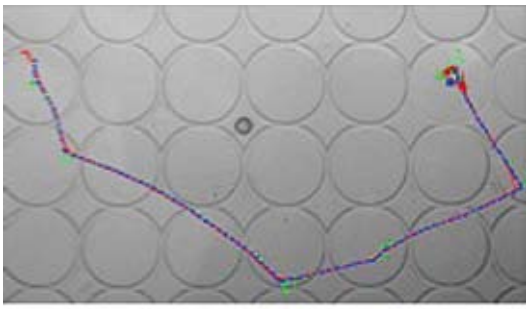


Figure 2.
Trajectory of a wild type larva approaching an odour source (blue square, top right corner). Successive positions of the head are shown in green. Positions of the tail are shown in blue.

3. New centres participating to orientation behaviour

Verena Hotz and Julia Riedl

Larvae appear to make use of active sampling to track odours. By moving forward and casting their head sideways, individuals are able to locally infer the direction of the vector pointing towards the source. How this decision-making process is achieved is still unknown. Our working hypothesis is that concentrations detected at different positions are held in a spatial short-term memory. To identify the circuits participating to this process, we have initiated a behavioural screen. We take advantage of thousand of existing Gal4-drivers to study the effects of disabling sub-regions of the larval brain by expressing the tetanus toxin light chain (synaptic transmission inhibitor). While specific neurons are silenced, we test the orientation capabilities of mutants to directional sensory stimuli. The anatomy and function of neuronal subsets associated with interesting phenotypes are inspected further to establish their involvement in specific defined behaviors.

PUBLICATIONS

Asahina K*, Louis M* et al.
"A circuit supporting concentration-invariant odor perception in Drosophila."
J Biol 8(1): 9 (2009). (*)

* *equal contribution*
(*) This article is mainly the result of the work of Dr. Matthieu Louis at the Rockefeller University, New York, USA.

SYSTEMS BIOLOGY

Group: Comparative Analysis of Developmental Systems

This group is part of the EMBL/CRG Research Unit in Systems Biology

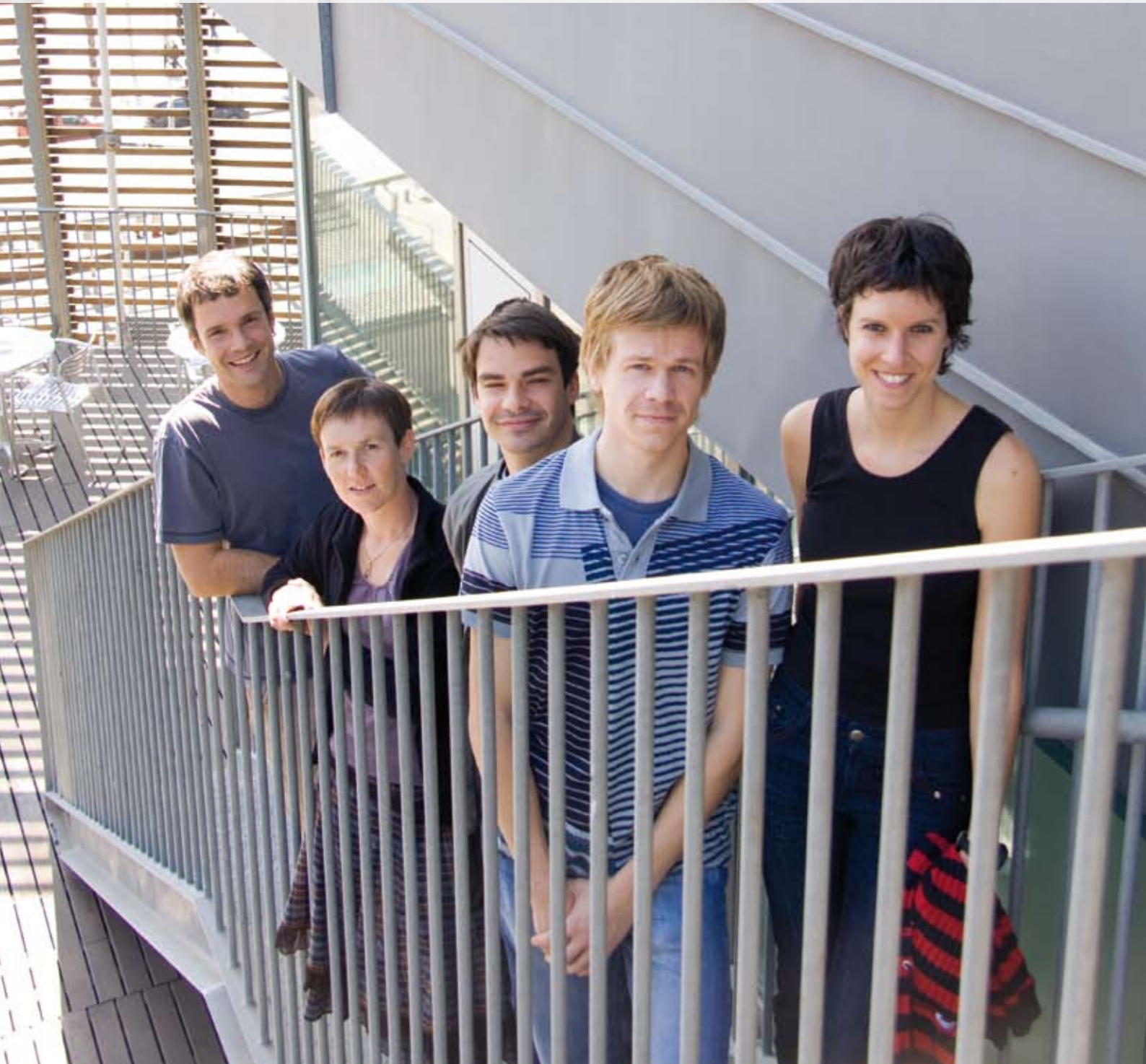
Group structure:

Group Leader: Johannes Jäger

Technician/Lab Manager: Hilde Janssens

Postdoctoral Fellows: Anton Crombach, Karl Wotton

PhD Student: Astrid Hörmann



SUMMARY

Natural selection acts on phenotypic variability within populations. However, we still lack a coherent view of how such variability arises during development, and how it reflects molecular variation in the genome. The relationship between genotype and phenotype is complex and non-linear. Therefore, we need a systems-biology approach to address this question. We are carrying out an integrative, comparative analysis of a real evolving developmental gene regulatory network using a novel reverse-engineering approach (the gene circuit method). Gene circuits are computational tools to extract regulatory information from quantitative spatial gene expression data (Fig. 1A). Our system of choice is the gap gene network involved in pattern formation in the early embryo of dipterans (flies, midges and mosquitoes; Fig. 1B). We are establishing three dipteran species—the fruit fly ***Drosophila melanogaster***, the scuttle fly ***Megaselia abdita*** and the moth midge ***Clogmia albipunctata***—as model systems to experimentally and quantitatively test hypotheses derived from systems-biology approaches to evolutionary developmental biology. Our models allow us to infer the regulatory interactions necessary and sufficient to explain the observed expression patterns by fitting models to data. Models from different species can be compared to reveal which interactions are conserved and which have diverged during evolution. In addition, we study evolutionary transitions between species using an ***in silico*** evolution approach. We will test these predictions by using RNA interference (RNAi) in various species and reporter assays in ***Drosophila***. Our approach provides an integrative view of network evolution across multiple levels, from the molecular to the phenotypic. To our knowledge, this has not yet been achieved for any real developmental system.

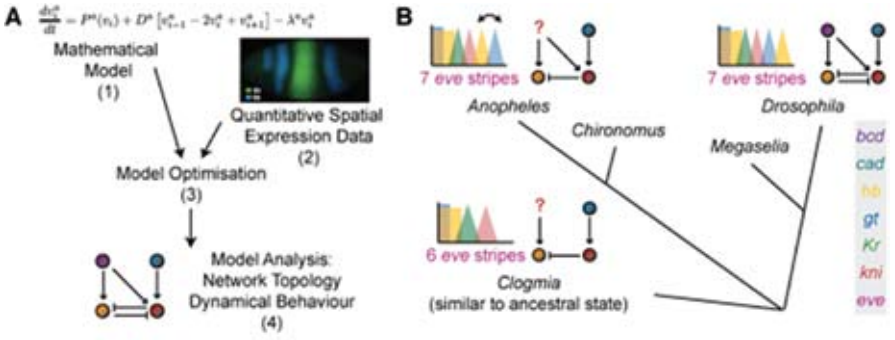


Fig. 1: *(A) Reverse-Engineering: The Gene Circuit Method. (B) Convergent evolution of the gap gene network. (B) shows a simplified phylogenetic tree of the dipterans (flies, mosquitoes and midges). Anopheles: malaria mosquito; Chironomus: chironomid midge; Clogmia: moth midge; Drosophila: fruit fly; Megaselia: scuttle fly. Relative position of gap gene expression domains and number of eve stripes at gastrulation time are shown schematycally for selected species (anterior is to the left). Network diagrams show hypothetical examples of regulatory changes: '?' highlights unknown anterior determinant. Posterior hb an gt domains have been swapped in Anopheles compared to Drosophila, and are missing in Clogmia.*

RESEARCH PROJECTS

1. A systems-level analysis of *giant (gt)* regulation in *Drosophila melanogaster*

Astrid Hörmann, Hilde Janssens

Aim. Due to the lack of a quantitative ***in vitro*** transcription assay, we still do not have a precise mechanistic understanding of eukaryotic gene regulation. For this reason, our current gene network models do not include molecular details, such as specific transcription factor binding sites and ***cis***-regulatory elements. As a first step towards resolving this issue, we will look into the regulation of a particular gap gene (***gt***) using our reverse-engineering approach with a model of transcriptional regulation. This enables us to identify and analyze contributions of particular binding sites to the expression pattern of ***gt***. In particular, we want to address important open questions such as how individual binding sites constitute a ***cis***-regulatory element or how such elements interact to result in the expression of a whole, endogenous gene.

Results. Preliminary analysis of selected ***gt cis***-regulatory elements shows that multiple elements contribute to most of gt's expression domains in a non-additive way, and that there are differences in regulation at early vs. late stages. We are currently creating strains of ***Drosophila*** carrying reporter constructs for these elements using site-specific transgenesis. We will quantify gene expression in each of these

strains, and will use the resulting data sets to fit a model of transcriptional regulation. This will allow us to predict the contribution of particular binding sites to the expression of specific gt domains. These predictions will then be tested experimentally.

2. A quantitative study of gap gene mutants in *Drosophila melanogaster*

Hilde Janssens

Aim. If we are to study evolutionary transitions using mathematical models of gene networks, we need a modeling formalism that captures the variational properties of the network. However, it remains unclear whether our current gene circuit models are able to correctly reproduce expression in mutant embryos. We will investigate these issues systematically by generating quantitative gene expression data for *Drosophila* gap gene mutants (with J. Reinitz, Stony Brook, USA), and by testing various modeling formalisms with regard to their ability to reproduce wild-type/mutant patterns correctly.

Results. Preliminary quantitative data sets for mutants of *Kr* and *kni* (Reinitz) are available, and we are currently completing a high-quality data set for mutants of the terminal gap gene *tll*. It reveals extensive embryo-to-embryo variation as some individuals have six, some seven stripes of the pair-rule gene *even-skipped (eve)*. This data set is currently being analyzed in detail, and will be used to fit and test various network modeling formalisms in the future.

3. A quantitative, comparative study of gap gene regulation in dipterans

Karl Wotton, Anton Crombach

Aim. We aim to create gene circuit models of the gap gene network in *Drosophila*, *Megaselia* and *Clogmia*. To achieve this, we require detailed characterization of the early stages of development, as well as quantitative spatial gene expression patterns for all three species. Such data sets can be acquired using immunofluorescence (for protein) or in situ hybridization (for mRNA expression patterns). These data will be analyzed to characterize spatial variability of gene expression domains and gene expression dynamics. Gene network models will be obtained by fits to data (Fig. 1A). We will characterize and compare the dynamical behavior of these networks. The resulting gene network topologies will be used to predict, which aspects of the network have been conserved, and which ones diverged during evolution. These predictions will be verified using RNAi and reporter assays.

Results. We have established colonies and experimental protocols for all three species, and are characterizing their early development by DIC time-lapse microscopy. We have cloned the complete set of gap genes from *Megaselia*. We are raising antibodies against gap proteins in *Megaselia* and *Clogmia*, and have created a preliminary data set for gap protein expression in *Clogmia* (with Ken Siggens, in Cambridge, UK; Fig. 2A). In parallel, we are using colorimetric in situ hybridization to systematically characterize gap gene expression at the mRNA level in both *Drosophila* and *Megaselia* (similar data for *Clogmia* were acquired previously by Mónica García Solache in Cambridge, UK). We have developed a novel data quantification pipeline for our mRNA data (Fig. 2B), which is currently used to create expression databases and integrated expression patterns for all three species. These data will be used to fit gene circuit models, and will serve as a basis for our *in silico* evolution study (see below). We are currently implementing an improved optimization algorithm (parallel Island-based Evolution Strategy, pIES) for model fitting.

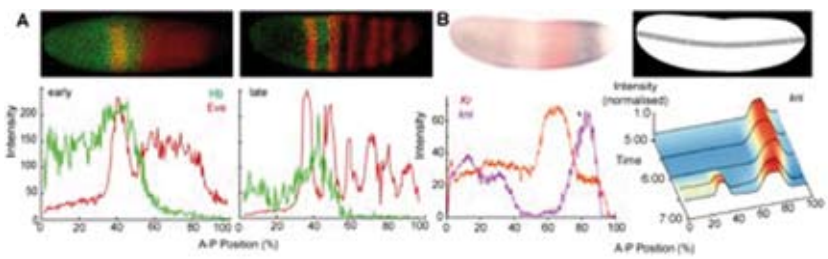


Fig. 2: *Quantified protein (A) and mRNA (B) expression patterns from the moth midge Clogmia.*
(A) Embryo images and quantified protein expression patterns for Hb and Eve at early & late blatodrem.
(B) Data processing pipeline for colorimetric in situ hybridization. Original image, embryo mask with region to be quantified, extracted profiles and reconstituted data (by spline approximation) are shown.

4. Modeling the evolutionary dynamics of the gap gene network

Anton Crombach

Aim. We want to explore the possibility that intermediate stages of evolution could be predicted (or reconstructed) using *in silico* evolution. In particular, we are interested whether these evolutionary transitions are constrained by network topology, or whether they depend on selection alone. Predictions from such an analysis will be tested against the data sets described above—or against qualitative expression data obtained from other suitable dipteran species.

Results. We are currently adapting evolutionary simulation code for use with our gene circuit models. We have obtained access to the Mare Nostrum computer at the Barcelona Supercomputing Center (BSC), where calculations will be carried out as soon as the required quantitative data sets are available (see above).

PUBLICATIONS

Ashyraliyev M, Siggens K, Janssens H, Blom J, Akam M & Jaeger J.
"Gene Circuit Analysis of the Terminal Gap Gene huckebein."
PLoS Comp Biol 5:e1000548 (2009).

Reviews

Jaeger J.
"Modelling the Drosophila Embryo."
Mol BioSyst 5:1549–1568 (2009). Review.

Jaeger J and Martinez-Arias A.
"Getting the Measure of Positional Information."
PLoS Biology 7: e1000081 (2009). Review.

Book Chapters

Jaeger J and Monk NAM.
"Reverse Engineering of Gene Regulatory Networks."
In: Learning and Inference in Computational Systems Biology, Lawrence ND, Girolami M, Rattray M & Sanguinetti G (eds.), MIT Press, Cambridge, MA (2009).

CORE FACILITIES

Director: Doris Meder

SUMMARY

The core facilities and the technologies they offer have become central to the science produced at the CRG. In order to account for this, and to convert the units into internationally competitive technology platforms, the CRG core facilities have been significantly restructured during this year. With the taking office of Dr. Doris Meder as Head of Core Facilities in February, the core facilities were unhinged from the research programmes they had previously been associated with, and were instead integrated into a separate structure under the coordination of the Head of Core Facilities. Now the core facilities are the third main column of CRG, next to research and administration. Not only does this remove the risk of core facilities becoming monopolized by individual opinions and instead make them equally accessible to all researchers at CRG and PRBB, but it also facilitates the establishment of common procedures.

OVERVIEW

The program currently comprises six core facility units: Genomics with the three subunits Genotyping, Microarrays and Ultrasequencing, Proteomics, Microscopy, High-throughput Screening, FACS and Bioinformatics. The Histology Service that was set up within the Differentiation and Cancer program and provides internal service to CRG researchers has been associated to the Core Facilities Program as well.

MISSION

The Core Facilities constitute the centralized scientific research infrastructure at CRG and allow research scientists shared access to sophisticated and expensive technologies that would be hard or impossible for the Institute to provide to each research group independently. CRG Core Facilities are open to all researchers at the PRBB and also to external users anywhere in the world. They are meant to be a meeting point for scientists from different institutes and disciplines, to foster exchange and integration of expertise, and to create and conduct interdisciplinary projects.

CONTRIBUTIONS AND ACHIEVEMENTS

The first half of 2009 was dedicated to consolidating procedures in this newly created program. Besides internal organizational measures, a "User Advisory Committee" was created for each Core Facility whose function is to monitor user satisfaction and to tailor the technology development towards the needs of CRG researchers. The committees consist of 4-5 regular users from CRG, or CRG and UPF who will meet quarterly with the Head of Unit and the Head of Core Facilities and give feedback on the performance of the Unit, as well as discuss new requirements.

In the second half of 2009 quite some effort was put into presenting the core facilities, the potential of the available technologies, and their services to the research community at the CRG and the PRBB. A "Core Facilities Presentation Day" was organized as a morning symposium with short presentations of each core facility, explaining to the researchers, which are the available technologies, which projects can be taken on, and how they should interact with the Unit. The event was attended by students, postdocs and group leaders from the CRG and other PRBB institutes. In order to visualize the workflow in each core facility, short video clips were prepared in collaboration with a local movie editor. These were first released during the "Core Facilities Presentation Day" and afterwards posted on the Units' web pages.

A Technology Symposium Series was established, the aim of which is to present to the researchers the cutting-edge technologies of today and tomorrow and how they will impact research. The first one was held on December 11th on Next-generation sequencing with Ivo Gut from France and Hans Lehrach from Germany presenting their latest research using ultradeep sequencing, and Stephen Turner from Pacific Biosciences in the USA presenting their 3rd generation sequencing technology. The symposium was extremely well attended, with more than 200 participants from the PRBB and other institutes in Barcelona and Catalunya. For 2010, symposia on Proteomics, High-content screening and super-resolution microscopy are planned.

The biggest achievement within Core Facilities in the second half of 2009 has been the establishment of the Bioinformatics Core. This has been long awaited since the demand for bioinformatics support of different types and levels is high, in the research groups as well as in the core facilities. A strategic concept for a Bioinformatics Core was elaborated and several high-profile candidates for the Head of Bioinformatics Core position were interviewed. Dr. Guglielmo Roma accepted the offer and joined the CRG in September. Together with him, the concept for the Bioinformatics Core was revised and the immediate strategies agreed upon. The first member joined the Bioinformatics Core in October, interviews for Bioinformatics Scientists were conducted in December and offers were made to three candidates. The Bioinformatics Core is already involved in a variety of projects and we are eagerly awaiting for it to take off with full resources in the next year.

CORE FACILITIES

Core Facility: Genomics Unit
Genotyping Unit

Unit Structure:

Unit Leader: Mònica Bayés

Project Manager: Magda Montfort

Technicians: Carles Arribas, Sílvia Carbonell, Cecília García, Sebastian Moran, Anna Puig

SUMMARY

Single nucleotide polymorphisms (SNPs) are the most common type of genetic variation in the human genome. A fraction of this genetic variation is likely to explain the majority of the differences between individuals, including their predisposition to develop common human disorders, such as cardiovascular disease, diabetes, asthma and cancer, and their differences in response to drugs. Other goals of SNP research include population genetics and trait selection for agricultural, cattle farming or aquaculture applications. Some of the genotyping technologies also enable to accurately characterize copy number variants (CNVs), loss of heterozygousity (LOH) and DNA methylation status. Our research activity, in some cases arising from the involvement of the Genotyping Unit in collaborative projects, is mainly focused on the identification of genes that contribute to psychiatric disorders such as attention-deficit/hyperactivity disorder and obsessive-compulsive disorder.

SERVICES

The Genotyping Unit, supported by "Genoma España" through the National Genotyping Centre (CeGen, *www.cegen.org*) (Barcelona Node Director Xavier Estivill, Genes and Disease Program), provides support to PRBB users and external users from public or private institutions for genotyping projects in every aspect of research, from experiment design, DNA extraction, genotyping, data interpretation, through to statistical analysis.

At the CeGen Barcelona Node several medium and high throughput genotyping and related services are available:

1. Automated DNA extraction from blood or other tissues (Chemagen)

2. DNA quantification using Picogreen (Molecular Probes)

3. Whole Genome Amplification using GenomiPhi (Amersham)

4. Custom Genotyping by SNPlex (Applied Biosystems):
genotyping of 24-48 SNPs selected by customer

5. Custom Genotyping by Veracode (Illumina):
genotyping of 48-384 SNPs selected by customer

6. Custom Genotyping with GoldenGate technology (Illumina):
genotyping of 96-1,536 SNPs selected by the customer

7. Focused-content SNP Genotyping with GoldenGate technology on all catalogue Illumina products: African American admixture panel, DNA test panel, Cancer SNP panel and MHC panel set, Human Linkage V panel, Mouse LD Linkage panel and Mouse MD Linkage panel

8. Focused-content SNP Genotyping with Infinium technology on all catalogue Illumina products (Figure 1): HumanOmni1-quad (NEW), Human1M-quad, Human660W-quad, HumanCytoSNP-12 (NEW), BovineSNP50, CanineSNP20 and EquineSNP50.

9. DNA methylation profiling using GoldenGate or Infinium technology on custom designed arrays and all catalogue Illumina products (Figure 2): HumanMethylation27 and Methylation Cancer Panel I.

All services are integrated with robust software tools for experimental design, management of data and analyses (Figure 3). Extensive quality control measures (both human and computational) let us further refine the quality of data. Additional information about services and SNP array can be found at the Genotyping Unit webpage (*http://www.crg.es/genotyping_unit*).

During 2009, liquid handling robots and Illumina scanner have been upgraded in order to be able to process the new high density Illumina BeadChips. Other improvements include the development of an in-house PCR based quality control assay for bisulfite treatment in methylation services, the implementation of specific software to identify non-Mendelian errors in high-density SNP data and the use of micronic barcoded tubes and plates for traceable DNA sample storage.

In addition, we have established an agreement with GATC Biotech so that researchers can outsource their Sanger sequencing reactions to this company. Personnel at the Genotyping Unit take care of sample shipments to GATC Biotech.

During 2009 the Unit has finished 55 genotyping projects for 31 different researchers, produced more than 1 million genotypes with SNPlex, more than 4 million genotypes for Illumina custom designs (Veracode and GoldenGate) and has processed 244 Beadchips through Infinium technology. It is worth to notice that the number of Veracode services increased almost 10 fold with respect to the previous period. The number of processed samples and the number of workorders for each service are detailed in table 1.

Services	Samples	Workorders
DNA quantification	32,724	46
DNA Extraction	1,326	6
Whole Genome Amplification	5,137	11
Custom Genotyping Illumina	16,660	17
Custom Genotyping SNPlex	14,496	24
Infinium	704	14

Fig 1.
Loading DNA samples on Ilumina Human1M-Duo BeadChips.

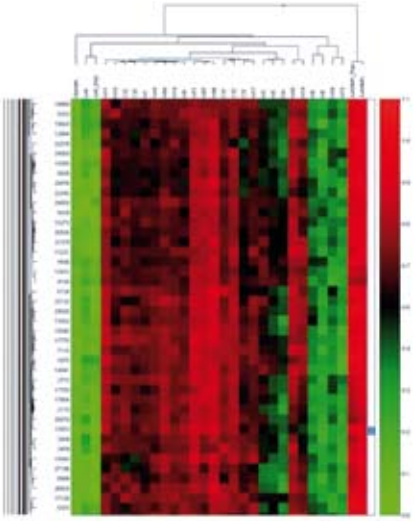


Fig 2.
Methylation profiles of bisulfite-converted genomic DNAs analyzed with the HumanMethylation27 BeadChip. Green, black and red colors correspond to low, medium and high methylation levels, respectively.

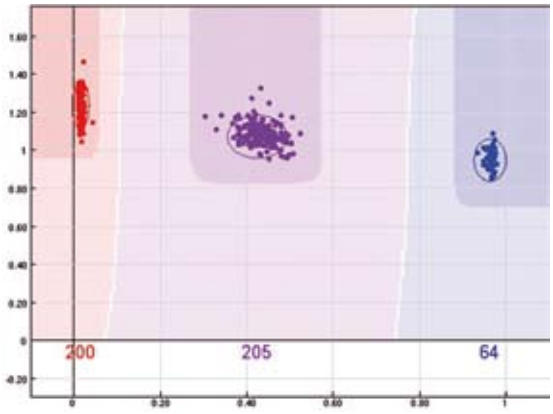


Fig 3.
Genotype calling with GenomeStudio software, with color coded data points (red = AA, purple = AB, blue = BB).

PUBLICATIONS

Franke B, Vasquez AA, Johansson S, Hoogman M, Romanos J, Boreatti-Hümmer A, Heine M, Jacob CP, Lesch KP, Casas M, Ribasés M, Bosch R, Sánchez-Mora C, Gómez-Barros N, Fernàndez-Castillo N, Bayés M, Halmøy A, Halleland H, Landaas ET, Fasmer OB, Knappskog PM, Heister AJ, Kiemeney LA, Kooij JJ, Boonstra AM, Kan CC, Asherson P, Faraone SV, Buitelaar JK, Haavik J, Cormand B, Ramos-Quiroga JA, Reif A.
"Multicenter analysis of the SLC6A3/DAT1 VNTR haplotype in persistent ADHD suggests differential involvement of the gene in childhood and persistent ADHD."
Neuropsychopharmacology 35(3):656-64 (2009).

Sánchez-Mora C, Ribasés M, Ramos-Quiroga JA, Casas M, Bosch R, Boreatti-Hümmer A, Heine M, Jacob CP, Lesch KP, Fasmer OB, Knappskog PM, Kooij JJ, Kan C, Buitelaar JK, Mick E, Asherson P, Faraone SV, Franke B, Johansson S, Haavik J, Reif A, Bayés M, Cormand B.
"Meta-analysis of brain-derived neurotrophic factor p.Val66Met in adult ADHD in four European populations."
Am J Med Genet B Neuropsychiatr Genet. 2009 Jul 14. [Epub ahead of print]

Cuenca-León E, Corominas R, Montfort M, Artigas J, Roig M, Bayés M, Cormand B, Macaya A.
"Familial hemiplegic migraine: linkage to chromosome 14q32 in a Spanish kindred."
Neurogenetics 10(3):191-8 (2009).

PUBLICATIONS BY CRG GENOTYPING UNIT USERS

Albacar G, Sans T, Martín-Santos R, García-Esteve L, Guillamat R, Sanjuan J, Cañellas F, Carot JM, Gratacòs M, Bosch J, Gaviria A, Labad A, Zotes AG, Vilella E.
"Thyroid function 48h after delivery as a marker for subsequent postpartum depression."
Psychoneuroendocrinology. 2009 Nov 23. [Epub ahead of print].

Baca-Garcia E, Vaquero-Lorenzo C, Perez-Rodriguez MM, Gratacòs M, Bayés M, Santiago-Mozos R, Leiva-Murillo JM, de Prado-Cumplido M, Artes-Rodriguez A, Ceverino A, Diaz-Sastre C, Fernandez-Navarro P, Costas J, Fernandez-Piqueras J, Diaz-Hernandez M, de Leon J, Baca-Baldomero E, Saiz-Ruiz J, Mann JJ, Parsey RV, Carracedo A, Estivill X, Oquendo MA.
"Nucleotide variation in central nervous system genes among male suicide attempters."
Am J Med Genet B Neuropsychiatr Genet 153B(1):208-13 (2009).

Castellano S, Andrés AM, Bosch E, Bayes M, Guigó R, Clark AG.
"Low exchangeability of selenocysteine, the 21st amino acid, in vertebrate proteins."
Mol Biol Evol 26(9):2031-40 (2009).

Castro-Giner F, Bustamante M, González JR, Kogevinas M, Jarvis D, Heinrich J, Antó JM, Wjst M, Estivill X and de Cid R.
"A pooling-based genome-wide analysis identifies new potential candidate genes for atopy in the European Community Respiratory Health Survey (ECRHS)."
BMC Medical Genetics 10(1):128 (2009).

Castro-Giner F, de Cid R, Gonzalez JR, Jarvis D, Heinrich J, Janson C, Omenaas ER, Matheson MC, Pin I, Anto JM, Wjst M, Estivill X, Kogevinas M.
"Positionally cloned genes and age specific effects in asthma and atopy: an international population-based cohort study (ECRHS)."
Thorax. 2009 Dec 8. [Epub ahead of print]

Castro-Giner F, Kogevinas M, Imboden M, de Cid R, Jarvis D, Mächler M, Berger W, Burney P, Franklin KA, Gonzalez JR, Heinrich J, Janson C, Omenaas E, Pin I, Rochat T, Sunyer J, Wjst M, Antó JM, Estivill X, Probst-Hensch NM.
"Joint effect of obesity and TNFA variability on asthma: two international cohort studies."
Eur Respir J 33(5):1003-9 (2009).

Castro-Giner F, Künzli N, Jacquemin B, Forsberg B, de Cid R, Sunyer J, Jarvis D, Briggs D, Vienneau D, Norback D, González JR, Guerra S, Janson C, Antó JM, Wjst M, Heinrich J, Estivill X, Kogevinas M.
"Traffic-related air pollution, oxidative stress genes, and asthma (ECHRS)."
Environ Health Perspect 117(12):1919-24 (2009).

de Cid R, Riveira-Munoz E, Zeeuwen PL, Robarge J, Liao W, Dannhauser EN, Giardina E, Stuart PE, Nair R, Helms C, Escaramís G, Ballana E, Martín-Ezquerra G, den Heijer M, Kamsteeg M, Joosten I, Eichler EE, Lázaro C, Pujol RM, Armengol L, Abecasis G, Elder JT, Novelli G, Armour JA, Kwok PY, Bowcock A, Schalkwijk J, Estivill X.
"Deletion of the late cornified envelope LCE3B and LCE3C genes as a susceptibility factor for psoriasis."
Nat Genet 41(2):211-5 (2009).

Gratacòs M, Costas J, de Cid R, Bayés M, González JR, Baca-García E, de Diego Y, Fernández-Aranda F, Fernández-Piqueras J, Guitart M, Martín-Santos R, Martorell L, Marti A, Santos JL, Gratacos M, Moreno-Aliaga MJ, Maiz A, Martinez JA, Estivill X.
"Association between leptin receptor (LEPR) and brain-derived neurotrophic factor (BDNF) gene variants and obesity: a case-control study."
Nutr Neurosci 12(4):183-8 (2009).

Menchón JM, Roca M, Sáiz-Ruiz J, Sanjuán J, Torrens M, Urretavizcaya M, Valero J, Vilella E, Estivill X, Carracedo A; Psychiatric Genetics Network Group.
"Identification of new putative susceptibility genes for several psychiatric disorders by association analysis of regulatory and non-synonymous SNPs of 306 genes involved in neurotransmission and neurodevelopment."
Am J Med Genet B Neuropsychiatr Genet 150B(6):808-16 (2009).

Morales E, Julvez J, Torrent M, de Cid R, Guxens M, Bustamante M, Künzli N, Sunyer J.
"Association of early life exposure to household gas appliances and indoor nitrogen dioxide with cognition and attention behavior in preschoolers."
Am J Epidemiol 169(11):1327-36 (2009).

Muiños-Gimeno M, Montfort M, Bayés M, Estivill X, Espinosa-Parrilla Y.
"Design and evaluation of a panel of single-nucleotide polymorphisms in microRNA genomic regions for association studies in human disease."
European Journal of Human Genetics. Epub 2009 Oct 7.

Real E, Gratacòs M, Soria V, Escaramís G, Alonso P, Segalàs C, Bayés M, Cid RD, Menchón JM, Estivill X.
"A Brain-Derived Neurotrophic Factor Haplotype Is Associated with Therapeutic Response in Obsessive-Compulsive Disorder."
Biol Psychiatry 66(7):674-80 (2009).

Ribasés M, Ramos-Quiroga Ja, Hervás A, Bosch R, Bielsa A, Gastaminza X, Artigas J, Rodriguez-Ben S, Estivill X, Casas M, Cormand B, Bayés M.
"Exploration of 19 serotoninergic candidate genes in adults and children with attention-deficit/hyperactivity disorder identifies association for 5HT2A, DDC and MAOB."
Mol Psychiatry 14(1):71-85 (2009).

Ribasés M, Bosch R, Hervás A, Ramos-Quiroga JA, Sánchez-Mora C, Bielsa A, Gastaminza X, Guijarro-Domingo S, Nogueira M, Gómez-Barros N, Kreiker S, Gross-Lesch S, Jacob CP, Lesch KP, Reif A, Johansson S, Plessen KJ, Knappskog PM, Haavik J, Estivill X, Casas M, Bayés M, Cormand B.
"Case-control study of six genes asymmetrically expressed in the two cerebral hemispheres: association of BAIAP2 with attention-deficit/hyperactivity disorder."
Biol Psychiatry 66(10):926-34 (2009).

Tortajada S, García-Gomez JM, Vicente J, Sanjuán J, de Frutos R, Martín-Santos R, García-Esteve L, Gornemann I, Gutiérrez-Zotes A, Canellas F, Carracedo A, Gratacos M, Guillamat R, Baca-García E, Robles M.
"Prediction of postpartum depression using multilayer perceptrons and pruning."
Methods Inf Med 48(3):291-8 (2009).

Ponsa I, Ramos-Quiroga JA, Ribasés M, Bosch R, Bielsa A, Ordeig MT, Morell M, Miró R, de Cid R, Estivill X, Casas M, Bayés M, Cormand B, Hervás A.
"Absence of cytogenetic effects in children and adults with attention-deficit/hyperactivity disorder treated with methylphenidate."
Mutat Res 666(1-2):44-9 (2009).

Soria V, Martínez-Amorós E, Escaramís G, Valero J, Pérez-Egea R, García C, Gutiérrez-Zotes A, Puigdemont D, Bayés M, Crespo JM, Martorell L, Vilella E, Labad A, Vallejo J, Pérez V, Menchón JM, Estivill X, Gratacòs M, Urretavizcaya M.
"Differential Association of Circadian Genes with Mood Disorders: CRY1 and NPAS2 are Associated with Unipolar Major Depression and CLOCK and VIP with Bipolar Disorder."
Neuropsychopharmacology, in press.

SOME PUBLICATIONS BY EXTERNAL GENOTYPING UNIT USERS

Bosch E., Laayouni H., Morcillo-Suarez C., Casals F., Moreno-Estrada A., Ferrer-Admetlla A., Gardner M., Rosa A., Navarro A., Comas D., Graffelman J., Calafell F., Bertranpetit J.
"Decay of linkage disequilibrium within genes across HGDP-CEPH human samples: most population isolates do not show increased LD."
BMC Genomics 10:338 (2009).

Corominas R, Ribases M, Camiña M, Cuenca-León E, Pardo J, Boronat S, Sobrido MJ, Cormand B, Macaya A.
"Two-stage case-control association study of dopamine-related genes and migraine."
BMC Med Genet 10:95 (2009).

Corominas R, Ribasés M, Cuenca-León E, Narberhaus B, Serra SA, del Toro M, Roig M, Fernández-Fernández JM, Macaya A, Cormand B.
"Contribution of syntaxin 1A to the genetic susceptibility to migraine: a case-control association study in the Spanish population."
Neurosci Lett 455(2):105-9 (2009).

Corominas R, Sobrido MJ, Ribasés M, Cuenca-León E, Blanco-Arias P, Narberhaus B, Roig M, Leira R, López-González J, Macaya A, Cormand B.
"Association study of the serotoninergic system in migraine in the spanish population."
Am J Med Genet B Neuropsychiatr Genet 153B(1):177-84 (2009).

Del Valle J, Feliubadaló L, Nadal M, Teulé A, Miró R, Cuesta R, Tornero E, Menéndez M, Darder E, Brunet J, Capellà G, Blanco I, Lázaro C.
"Identification and comprehensive characterization of large genomic rearrangements in the BRCA1 and BRCA2 genes."
Breast Cancer Res Treat. 2009 Nov 6. [Epub ahead of print]

Montesinos A, Tonsor SJ, Alonso-Blanco C, Picó FX.
"Demographic and genetic patterns of variation among populations of Arabidopsis thaliana from contrasting native environments."
PLoS One 4(9):e7213 (2009).

Moreno-Estrada A, Tang K, Sikora M, Marquès-Bonet T, Casals F, Navarro A, Calafell F, Bertranpetit J, Stoneking M, Bosch E.
"Interrogating eleven fast-evolving genes for signatures of recent positive selection in worldwide human populations."
Molecular Biology and Evolution 26(10):2285-2297 (2009).

CORE FACILITIES

Core Facility: Genomics Unit
Ultrasequencing Unit

Unit Structure:

Unit Leader: Heinz Himmelbauer

Laboratory Technicians: Ana Vivancos, Ester Castillo, Anna Menoyo, Maria Luisa Campos (from October 2009), Maik Zehnsdorf (GABI-FUTURE), Hui Kang (GABI-FUTURE; from October 2009)

Bioinformaticians: Matthew Ingham, Debayan Datta, Robert Kofler (GABI-FUTURE; until June 2009), Darek Kedra (GABI-FUTURE; from December 2009), Juliane Dohm (GABI-FUTURE)

PhD student: Andrè Minoche (GABI-FUTURE)

SUMMARY

The sequencing field has experienced a major boost by the introduction of innovative, second-generation sequencing technologies. The Ultrasequencing Unit utilizes two different high-throughput sequencing platforms. At present, the Unit is equipped with two Illumina Genome Analyzer IIx sequencers, and one Roche 454 FLX sequencing instrument. A wide range of different sequencing services is provided to the CRG, PRBB, and to external users, including state-of-the art technologies like ChIP-Seq, mRNA-Seq, and small RNA sequencing.

SERVICES

Next generation sequencing has paved the way for answering a large number of different questions in biology and in biomedical sciences. The protocols currently in use in the Unit are genomic sequencing (***de novo*** sequencing, re-sequencing, ChIP-Seq, mate-pair sequencing), transcriptome analysis (mRNA-Seq, strand-specific mRNA Seq, small RNA sequencing, RIP-Seq), and amplicon sequencing. In addition, procedures for indexing samples are in use, allowing the pooling of samples. The Illumina and 454 platforms in the Ultrasequencing Unit complement each other: Illumina sequencing produces large numbers of short reads per run (up to 160 million reads per run, read length 36-76 bases). With the 454 technology, we generate approximately 1 million reads in a single run (read length 500 bases, 300-500 Mb sequence output per run). Thus, Solexa is very well suited for applications such as ChIP-Seq, miRNA detection, and genome re-sequencing, while long-read 454 sequencing is very appropriate for projects such as ***de novo*** sequencing of complex genomes and transcriptome characterization in non-model species lacking a reference genome.

The services offered to users include the preparation of samples ready for sequencing on the GA IIx and the 454-FLX instruments, and performing the sequencing runs. After basecalling, the read sequences are quality filtered and mapped against a reference genome (if applicable) and are made available to the users as sequence files, quality files, and alignment files.

RESEARCH PROJECTS

The research of the Ultrasequencing Unit focuses on the development of new procedures, both in the lab, and for data analysis. Ongoing projects include the development of DSSS (Direct Strand Specific Sequencing, a protocol for strand-specific sequencing of prokaryotic and eukaryotic RNA samples), developing robust procedures for miRNA profiling, and the analysis of structural variation in cancer genomes. With funding provided by the German Ministry for Education and Research (BMBF) in the context of the GABI-FUTURE program, we are in the process of sequencing the genome and transcriptome of the crop plant sugar beet (***Beta vulgaris***) using 454 and Solexa technology. Working on a genome about one quarter of the size of the human genome, this project enables us to establish and test the workflows required to process and analyse sequence data from organisms with no reference genome sequence available.

PUBLICATIONS

Doss MX, Wagh V, Schulz H, Kull M, Kolde R, Pfannkuche K, Nolden T, Himmelbauer H, Vilo J, Hescheler H, Sachinidis A.
"Comprehensive Analysis of Murine ES Cell Derived Brachyury+ (T) Cell Specific Transcriptome."
Genes to Cells, in press.

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

Schulz H, Kolde R, Adler P, Aksoy I, Anastassiadis K, Bader M, Billon N, Boeuf H, Bourillot PY, Buchholz F, Dani C, Xavier Doss MX, M Forrester L, Gitton M, Henrique D, Hescheler J, Himmelbauer H, Hübner N, Karantzali E, Kretsovali A, Lubitz S, Pradier L, Reimand J, Rolletschek A, Sachinidis A, Savatier P, Stewart F, Storm MP, Trouillas M, Vilo V, Welham M, Winkler J, Wobus AM, Hatzopolous AK.
"The FunGenES database: a genomics resource for mouse embryonic stem cell differentiation"
PLoS ONE 4(9):e6804 (2009).

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

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

Tsend-Ayush A., Dodge N., Mohr J., Casey A., Himmelbauer H., Kremitzki C.L., 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. (*)

(*) All these publications are the result of the work of Dr. Heinz Himmelbauer at the Max Planck Institute for Molecular Genetics, Berlin, Germany.

PUBLICATIONS BY CRG ULTRASEQUENCING UNIT USERS

Guell M et al.
"Transcriptome complexity in a genome-reduced bacterium."
Science 326:1268-1271 (2009).

PUBLICATIONS BY EXTERNAL ULTRASEQUENCING UNIT USERS

Lalueza-Fox C, Gigli E, de la Rasilla M, Fortea J and Rosas A.
"Bitter taste perception in Neanderthals through the analysis of the TAS2R38 gene."
Biol Lett 5:809-811 (2009).

CORE FACILITIES

Core Facility: Genomics Unit
Microarrays Unit

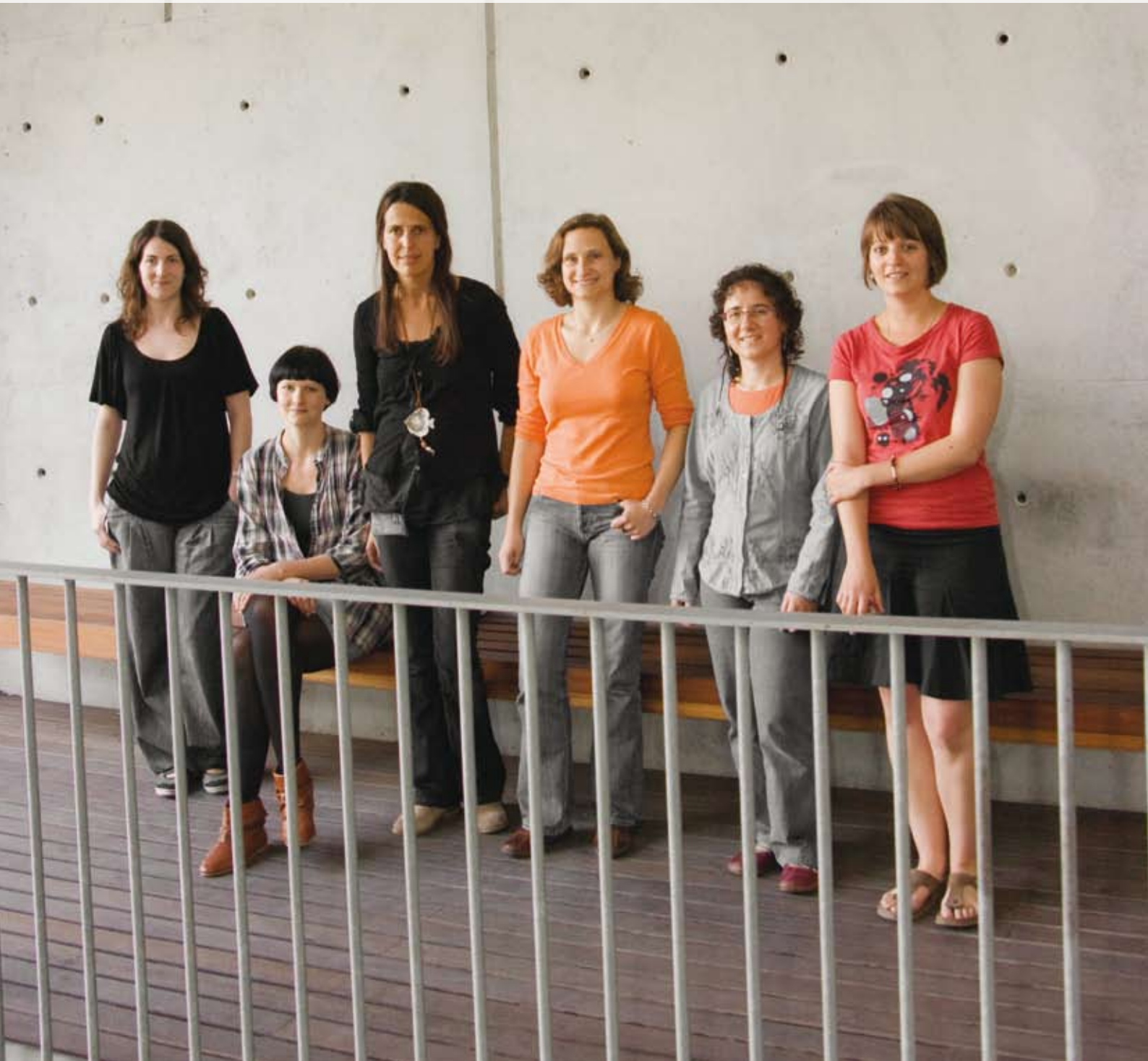
Unit structure:

Unit Leader: Mònica Bayés

Lab Manager: Anna Ferrer

Technicians: Heidi Mattlin, Maria Aguilar

Bioinformaticians: Sarah Bonnin, Manuela Hummel



SUMMARY

Microarrays allow us to analyze expression profiles (mRNA and microRNA levels) and structural variation (DNA copy number) on a genome-wide level. The Microarrays Unit is equipped with all the instruments to carry out microarray analysis with the latest version of Agilent and Illumina technologies. For both technologies, the Microarrays Unit provides a full service, including technical advice in experimental design, quality control of starting material, sample and array processing, and data analysis.

SERVICES

The laboratory provides microarray methodologies as a service at established rates to scientists from the CRG, other PRBB institutions and other external public and private institutions. Services offered include: sample quality control and amplification, RNA and DNA sample labeling, hybridization of microarrays and data processing and analysis, as well as array design and fabrication by spotting.

The Microarrays Unit offers a range of options suitable for a variety of applications including:

> mRNA expression profiling on Agilent microarrays using the 2-color and 1-color protocols, and Illumina BeadArrays (1-color protocol).
> miRNA expression profiling on Agilent microarrays.
> array based comparative genomic hybridization (aCGH) on spotted BAC and Agilent microarrays.
> chromatin immunopreciptation on array (ChIP-on-chip) on spotted promoter, CpG island and Agilent microarrays.

At the CRG, the most common microarray application is mRNA expression profiling. During 2009 we have set up the Agilent 1-color protocol, which offers more flexibility in the experimental design. For particular samples, we have also developed a new labeling protocol based on ***random priming***. The Agilent scanner at the Unit has been upgraded so that it has better resolution and dynamic range and can process the high-resolution 1-million-feature microarrays.

We are currently including one commercial reference sample in every microarray experiment, ***Universal Human Reference RNA*** for mRNA profiling and ***Placenta total RNA*** for miRNA profiling. They provide a high-quality standard for accurate and consistent data comparison. Bioinformaticians at the Unit have developed new tools to monitor the performance of these RNA reference samples across different experiments. Parameters such as signal and background intensity ratios and the number of detected genes are automatically plotted for each experiment and are extremely helpful for identifying potential failures in the array processing protocol (Figure 1).

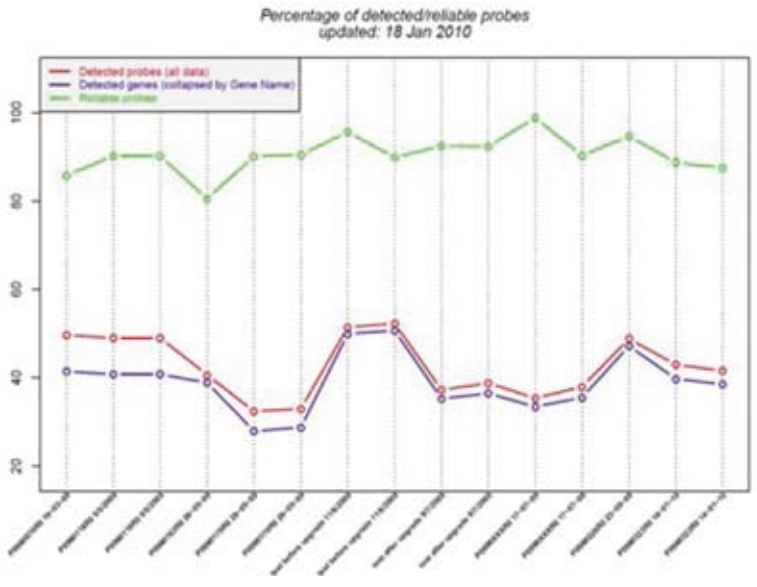


Figure 1.
Bioinformatic tools for monitoring microarrays quality based on the use of a reference RNA sample (UHRR, Stratagene). The percentage of detected (intensity above a certain threshold) microarray probes (red) and genes (blue) and percentage of reliable (intensity sufficiently larger than background) probes (green) across 15 different Agilent 2-color experiments are shown.

We have evaluated the performance of three microarray platforms (Agilent 1-color, Agilent 2-colors and Illumina) in profiling the expression of two commercially available reference RNA samples and compared the results with those obtained by the Microarray Quality Control Consortium (MAQC) (Figure 2). We have conducted a total of 36 hybridizations, including at least 4 technical replicates. Our results show relatively low technical variability in the intra-platform measurements (median CV of 7–12% for the quantitative signal), as well as high inter-platform concordance (>70%) in terms of genes identified as differentially expressed between 2 reference RNA samples. The results from the Agilent 1-color platform are more comparable to TaqMan data, whereas the others demonstrate higher signal compression. In summary, our study indicates and overall good performance of all 3 platforms tested and high agreement with MAQC results in terms of reproducibility, sensitivity and accuracy.

During 2009 the Unit has finished 49 projects for 29 different researchers, processing 385 samples for mRNA expression, 192 samples for miRNA profiling, 10 samples for aCGH and 14 samples for ChIP-on-chip. In addition, we have fabricated 625 arrays using the Versarray spotter. The Unit also provides access and support to real-time qPCR instruments and has set up a new service for automated 384 plate preparation using liquid handling robots.

A. Intra-platform reproducibility

B. Microarray and qPCR data comparison

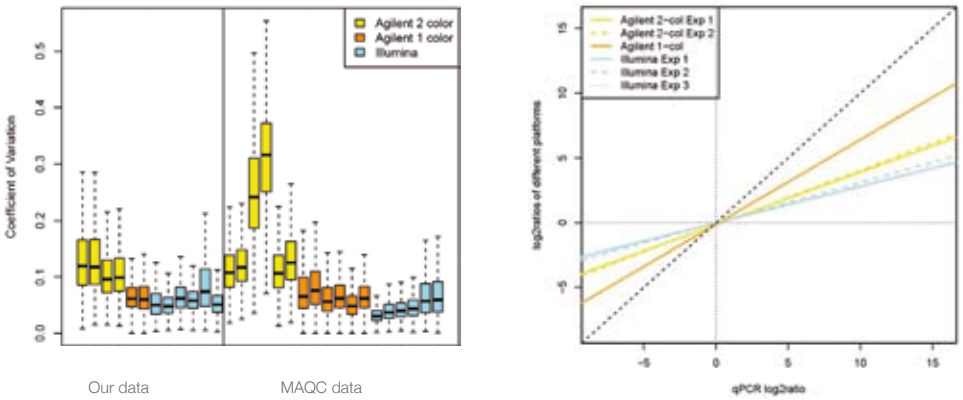


Figure 2. *Microarray cross-platform comparison.*

a) Gene-wise coefficients of variation of the expression signal values between replicates of the same sample were calculated for both for the different test experiments performed at the Unit (left panel) and for the results reported by the MAQC study (right panel). The boxplots are highlighted to distinguish the different experiments performed using different platforms: Agilent 2-colors (yellow), Agilent 1-color (orange) and Illumina (blue). b) Comparison of the log ratio differential expression values from each microarray platform relative to values obtained by TaqMan assays. The solid lines represent a linear regression fit. Microarrays compression effect is visualized as a deviation from the dashed line (ideal slope=1).

PUBLICATIONS BY CRG MICROARRAY UNIT USERS

Brunet A, Armengol L, Heine D, Rosell J, García-Aragonés M, Gabau E, Estivill X, Guitart M.
"BAC array CGH in patients with Velocardiofacial syndrome-like features reveals genomic aberrations on chromosome region 1q21.1."
BMC Med Genet 10:144 (2009).

Armengol L, Villatoro S, González JR, Pantano L, García-Aragonés M, Rabionet R, Cáceres M, Estivill X.
"Identification of copy number variants defining genomic differences among major human groups."
PLoS One 4(9):e7230 (2009).

Buschbeck M, Uribesalgo I, Wibowo I, Rué P, Martin D, Gutierrez A, Morey L, Guigó R, López-Schier H, Di Croce L.
"The histone variant macroH2A is an epigenetic regulator of key developmental genes."
Nat Struct Mol Biol 16(10):1074-9 (2009).

Servitja JM, Pignatelli M, Maestro MA, Cardalda C, Boj SF, Lozano J, Blanco E, Lafuente A, McCarthy MI, Sumoy L, Guigó R, Ferrer J.
"Hnf1alpha (MODY3) controls tissue-specific transcriptional programs and exerts opposed effects on cell growth in pancreatic islets and liver."
Mol Cell Biol 29(11):2945-59 (2009).

Quiles I, Millán-Ariño L, Subtil-Rodríguez A, Miñana B, Spinedi N, Ballaré C, Beato M, Jordan A.
"Mutational analysis of progesterone receptor functional domains in stable cell lines delineates sets of genes regulated by different mechanisms."
Mol Endocrinol 23(6):809-26 (2009).

PUBLICATIONS BY EXTERNAL MICROARRAY UNIT USERS

Cuscó I, Medrano A, Gener B, Vilardell M, Gallastegui F, Villa O, González E, Rodríguez-Santiago B, Vilella E, Del Campo M, Pérez-Jurado LA.
"Autism-specific copy number variants further implicate the phosphatidylinositol signaling pathway and the glutamatergic synapse in the etiology of the disorder."
Hum Mol Genet 18(10):1795-804 (2009).

Rangel A, Madroñal N, Gruart i Massó A, Gavín R, Llorens F, Sumoy L, Torres JM, Delgado-García JM, Del Río JA.
"Regulation of GABA(A) and glutamate receptor expression, synaptic facilitation and long-term potentiation in the hippocampus of prion mutant mice."
PLoS One 4(10):e7592 (2009).

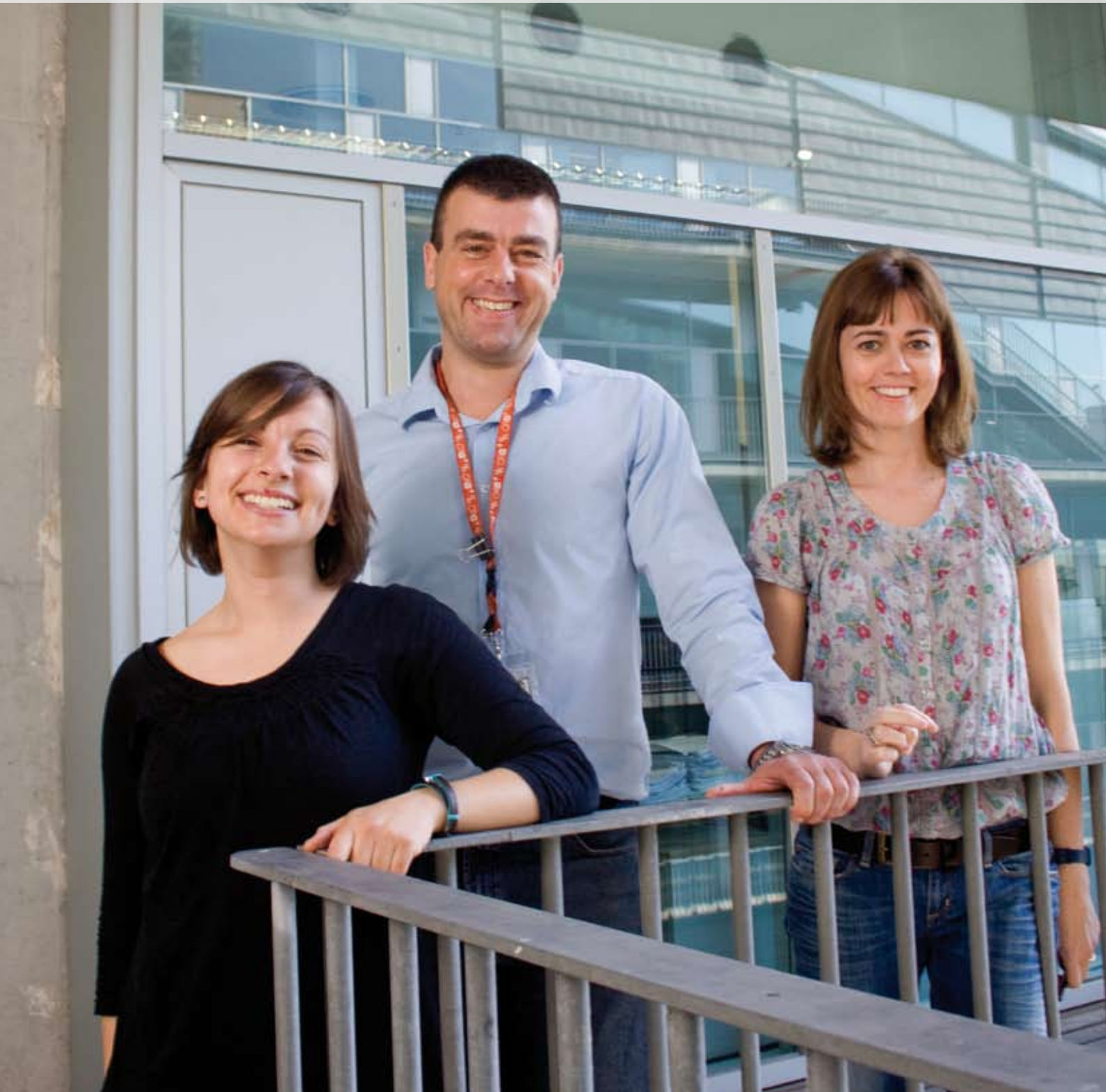
CORE FACILITIES

Core Facility: Proteomics Unit

Unit Structure:

Unit Leader: Henrik Molina (CRG)

Staff Technicians: Cristina Chiva (UPF), Eva Borras (UPF), Carolina de la Torre (ProteoRed), Guadalupe Espadas Garcia (CRG)



SUMMARY

Proteome & Proteomics: The word proteome was first conceived by Marc Wilkins in 1984 and is a mix of ***proteins*** and ***genome***. Proteomics is a contraction of ***proteome*** and ***–omics*** and translate into "the study proteins".

In a simplified way, the genome can be seen as the cook book for an organism…with the proteome being the dish (Figure 1A). Generally speaking, the genome is a relative static entity which is in sharp contrast to the very dynamic proteome. The proteome changes over time, between different cellular compartments, or in response to stimulation. An example of this is a recent study by Julia von Blume/Vivek Malhotra (von Blume et al. ***Journal of Cell Biology***, 2009). Here, a proteomics 'snapshot' is taken of the proteins secreted from HeLa cells (the ***secretome***) when genes coding for ADF and cofilin are suppressed. Even when focusing a study to only one organism, the term proteome might not make much sense. Because of the ever changing conditions, the plural form of proteome, proteomes, is much more appropriate.

Fig 1A.
The relation between the genome and the proteome can be compared to a recipe (here; crème brulée) and the product that emerges when the recipe is followed. In the photo, the crème brulée is placed on top of a scale which can represent mass spectrometry used in proteomics.

Proteomics is the tool used to decipher proteomes and proteomics is thus a very broad area that includes the analysis of:

I) How genes are combined to give rise to alternative gene products (splicing) or isoforms
II) Where specific proteins are expressed,
III) When proteins are expressed,
IV) How proteins interact (protein-protein interactions),
V) How proteins are being modified (post-translational modifications), and
VI) What are the dynamics of proteins with respect to stimuli or other cellular states.

Mass spectrometry has proven a valuable technology for studying proteomics and it is safe to say that mass spectrometry has been the driving force for proteomics, and that today close to half of all proteomics studies rely on mass spectrometry. The CRG/UPF Joint Proteomics Unit is also relying heavily on these instruments. The Unit is equipped with four such instruments where two of these are ***state-of-the-art*** hybrid ion trap / Orbitrap type mass spectrometers (Figure 1B). These ***state-of-the-art*** instruments offer high resolution, fast acquisition speed and excellent sensitivity, factors which are crucial for in-depth proteomic studies.

Fig 1B.
Two of the work-horse mass spectrometers in the Proteomics Unit. Both Mass Spectrometers are connected to Liquid Chromatography systems that are used to separate digested proteins prior to analysis by the mass spectrometers. Such setups are abbreviated: LC-MS.

Where proteomics is often associated with global or larger studies, exemplified by e.g. biomarker screens, the analysis of complex signal transduction pathways or mapping phosphorylations in a cell, the Proteomics Unit is also taking part in smaller and much more focused studies. Such specialized examples include Almer Vandersloot's /Luis Serrano's (CRG) elucidation of a proteins 3D structure by combining proton/deuterium exchange and mass spectrometry. Figure 2A shows a mass spectrum of a peptide originating from a protein, before (top panel) and after (bottom panel) protons in the proteins have been allowed to exchange with deuterium. The degenerated pattern in the lower spectrum suggests that amino acids in this peptide have taken part in a deuterium exchange and it can be concluded that residues in the sequence are located at the surface of the protein and therefore exchange could take place. Another example of a focused study is supplied by the group of Elena Hidalgo (Universitat Pompeu Fabra - UPF). This group studies the ***redox-ome*** in yeast cells focusing on the amino acid cysteine. Using stable isotopes and a specific enrichment strategy, cysteines are targeted and compared to a control system. The strategy allows filtering for innocent bystanders and pinpointing cysteines that are members of the studied redox system. The active cysteines are detected by displaying a differential signal between the control and the treated systems. The spectrum (Figure 2B) shows the mass spectrum of a reduced cysteine containing peptide which shows a 6-fold difference between the control and the treatment.

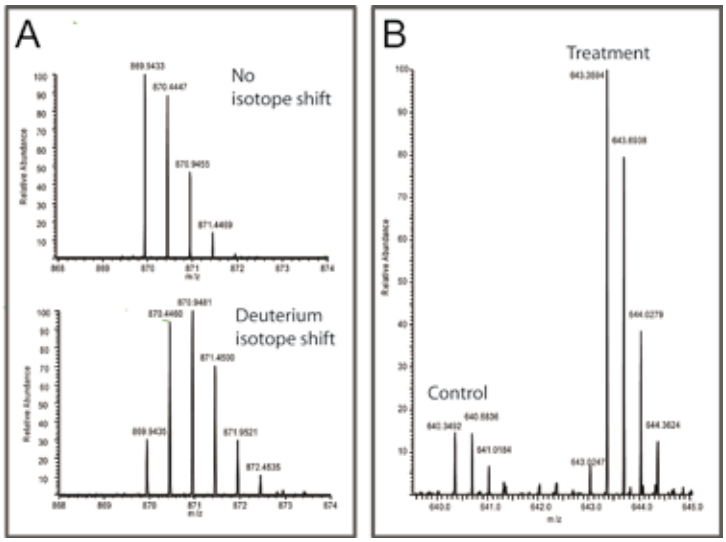


Fig 2.
Three mass spectra are shown. Horizontal axes show mass divided by charge (m/z) units, whereas vertical axes are relative intensities. A) Top panel shows a spectrum of a peptide from a protein that has not yet been allowed to exchange protons with the 1 Dalton heavier deuteriums. The lower spectrum shows the same spectrum after the exchange has taken place. The mass increases (in denominations of 1 Dalton) illustrates that residues in this particular sequence have been accessible for exchange, meaning they have been exposed on the surface of the protein. B) Spectrum of a pair of cysteine containing peptides with identical chemical properties. The mass difference between the two peptides is caused by chemical tags and the mass difference is created by using tags that differ in their content of a normal and a heavier carbon isotope (13C). The intensities (vertical axis) can directly be converted into fold differences in-between the cysteines in the two peptides (control versus treatment).

SERVICES TYPE OF OPERATIONS OFFERED

In addition to specialized studies, as exemplified above, typical standard services provide by the unit are:

I) Identifications of proteins by either peptide mass finger printing (PMF) or liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS)
II) Determination of molecular mass of molecules.
III) Phosphorylation enrichments and analysis by LC-MS/MS
IV) Relative quantitation of proteins based on stable isotopes (iTRAQ or SILAC)
V) Chromatographic separation of peptides and proteins
VI) 1D and 2D gel electrophoresis

AFFILIATIONS

The Joint UPF/CRG Proteomics Core Facility is a member of ProteoRed, the Spanish Proteomics Network.

PUBLICATIONS

Kandasamy K, Pandey A, Molina H.
"Evaluation of several MS/MS search algorithms for analysis of spectra derived from electron transfer dissociation experiments".
Anal Chem 81(17):7170-80 (2009).

CORE FACILITIES

Core Facility: Advanced Light Microscopy Unit

Unit structure:

Unit Leader: Timo Zimmermann

Technicians: Raquel Garcia Olivas, Arrate Mallabiabarrena

SUMMARY

The Advanced Light Microscopy Unit (ALMU) of the CRG serves as a core facility for high-end light microscopy for CRG researchers. A range of instruments with unique capabilities covers the spectrum of advanced imaging applications from thick tissue reconstruction to fast in-vivo imaging to the sensitive detection of very faint signals. The staff of the facility provides advice in the initial experiment planning, training of the researchers on the instruments and assistance with the subsequent data analysis. It is the aim of the facility to provide a link for the biological questions of researchers to the full capabilities of advanced light microscopy at the organismic, cellular and molecular level. Methods available in the facility include optical sectioning (single photon and multi-photon microscopy), spectral imaging, in-vivo timelapse imaging, Total Internal Reflection Fluorescence (TIRF) Microscopy and methods for the study of molecular properties and interactions like Fluorescence Correlation Spectroscopy (FCS), Fluorescence Lifetime Imaging Microscopy (FLIM), Fluorescence Resonance Energy Transfer (FRET) detection and Fluorescence Recovery after Photobleaching (FRAP). Additionally, dedicated software packages for data visualization and analysis are available for 3D rendering, particle tracking and image analysis.

In this year the ALMU has extended its capabilities for thick tissue imaging and for molecular studies by installing a unique platform comprising two multiphoton microscopes and detection modules for FCS, FLIM and combinations of these technologies. The unit is used regularly by researchers from all CRG programs and additionally by researchers from other PRBB institutes. Applications range from immunofluorescence imaging of fixed samples to timelapse observations spanning several days.

FACILITY OVERVIEW

In 2009, the Advanced Light Microscopy Unit continued to provide instrumentation at the forefront of imaging technology through the installation of a worldwide unique imaging platform. The complete system consists of an inverted Leica SP5 confocal microscope with resonant scanner, the new Leica Single Molecule Detection Platform for FLIM, FCS and FLCS (Fluorescence Lifetime Correlation Spectroscopy) and an upright fixed stage confocal microscope Leica SP5 CFS (Figure 1). Both microscope systems are connected to a multiphoton laser that can be shared between the systems. Above the instrument table, a system of instrumentation racks on gliding rails provides a freely configurable workspace to house additional equipment that should be located away from the imaging table to avoid vibrations. The combination of an inverted and an upright microscope stand allow the full range of applications from live cell imaging through the bottom of petri dishes to complex electrophysiology measurements in animal models. Through an optical path inverter, the systems can also be simultaneously operated as two upright or two inverted microscopes, thus providing full flexibility for experimental planning.

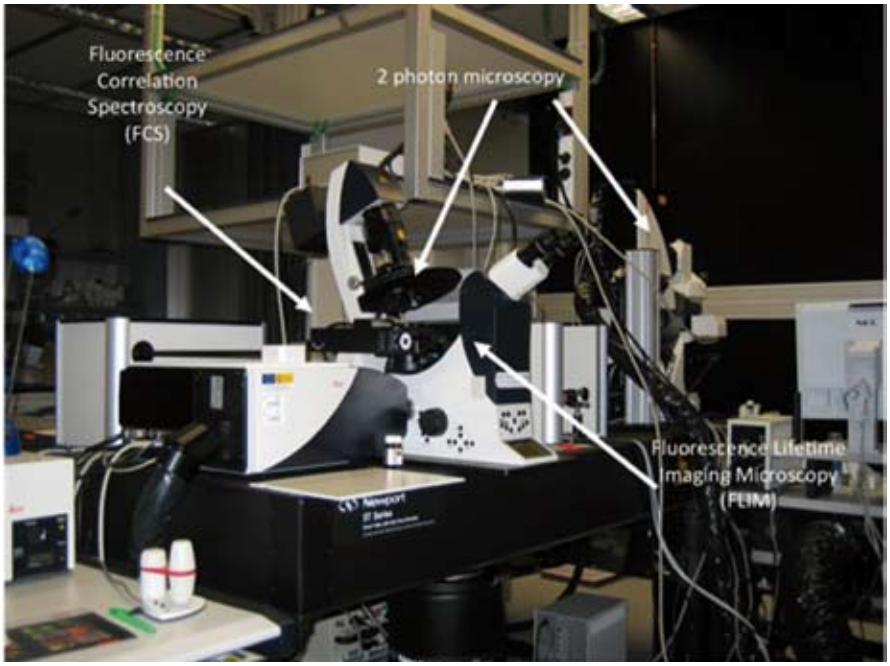


Fig 1.
A system combination of two confocal microscopes (Leica TCS SP5 AOBS inverted and Leica TCS SP5 CFS upright), one MaiTai multiphoton laser and the Leica Single Molecule Detection Platform for FCS, FLIM and FLCS

The systems were co-financed by the Ministry for Science and Innovation (MICINN) from funds available in the “Fondo Europeo de Desarrollo Regional” (FEDER) program of the European Union. As the instruments are very complex and sensitive to user errors, an additional staff position was created for them through the ISC III initiative “Contratos de Técnicos de Apoyo a la Investigación en el Sistema Nacional de Salud” that was filled in February with Dr. Arrate Mallabiabarrena, a highly experienced light microscopist.

Among the six available microscope systems of the unit, no two are identical in their features. Because of this, a wide range of microscopy applications can be covered. However, most applications can be performed on at least two systems. This redundancy ensures that experiment planning is not impaired by the limited availability of a single system. Reflecting the variety in available instrumentation, the experiments performed in the year ranged from in-vivo timelapse experiments spanning several days to the high-resolution 3D imaging of multiple intranuclear components.

The total microscope usage time of the unit in 2009 surpassed 14000 hours, corresponding to almost eight hours of daily usage per microscope system. The usage has grown by 50% compared to the year before, reflecting the high need for light microscopy by CRG researchers and the additional capacity that became available by the addition of two more systems to the existing three. During the year, 93 in-house users from 24 of the CRG research groups have used the unit. On average, 38 users from 16 groups worked on the microscopes every month. Additionally the unit was used by 14 users from 8 groups from UPF and IMIM.

During the year, the ALMU staff has participated in teaching masters courses of the Universitat Pompeu Fabra (UPF), as well as in microscopy courses in other Barcelona universities. They have continued to participate as speakers and instructors in courses, workshops and seminars at institutes in Spain and in other European countries.

Together with Isabelle Vernos from the Cell and Developmental Biology Programme, the head of the unit, Timo Zimmermann, co-organised in November a two day microscopy course and a two-day international scientific symposium on “Imaging Approaches to study cytoskeletal dynamics” with Leica Spain.

The ALMU has also participated in the “Zeiss on your campus” initiative, a showcase of the latest Zeiss microscopy developments for several days in selected institutes in Europe in February. Carl Zeiss Spain and Olympus Spain installed demo systems for thorough evaluation for up to several weeks. The ALMU also performed on-site application testing of systems in development.

In the ongoing effort to streamline core facility infrastructures inside PRBB, CRG and the UPF Department of Experimental and Health Sciences have agreed to closely link the confocal microscopy facility of the UPF with the ALMU. Accordingly, the UPF confocal microscope was moved in November into a location inside CRG and the UPF microscopy specialist, Dr. Xavier Sanjuan, moved into the ALMU office. The current plan is that also the operating procedures and instrument billing of the two units will be unified in the course of the next year.

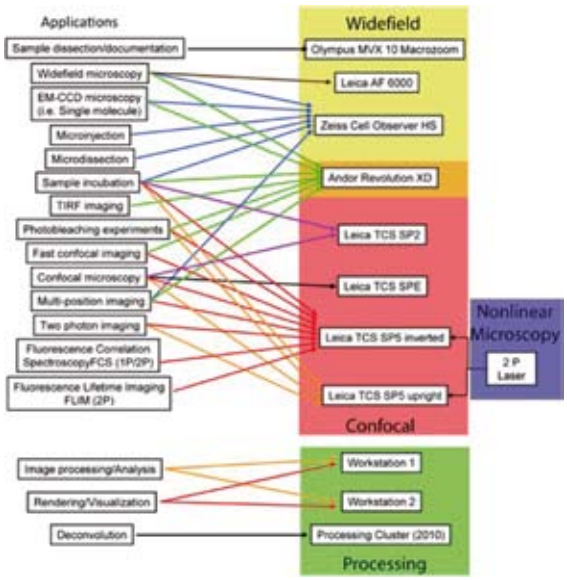


Fig 2.
Overview of the available applications in the facility and their distribution on the corresponding instruments.

PUBLICATIONS

Buntru A, Zimmermann T and CR Hauck.
"Fluorescence resonance energy transfer (FRET)-based subcellular visualization of pathogen-induced host receptor signaling."
BMC Biol 7:81 (2009).

Bussmann LH, Schubert A, Vu Manh TP, De Andres L, Desbordes SC, Parra M, Zimmermann T, Rapino F, Rodriguez-Ubreva J, Ballestar E and T Graf.
"A robust and highly efficient immune cell reprogramming system."
Cell Stem Cell 5(5):554-66 (2009).

Book chapters

Zimmermann T, Morrison J, Hogg K, O'Toole P.
"Clearing up the signal: Spectral imaging and linear unmixing in fluorescence microscopy."
In: Confocal microscopy methods and protocols, Ed. Stephen Paddock, in press.

PUBLICATIONS BY CRG ALMU USERS

Sardon T, Cottin T, Xu J, Giannis A and Vernos I.
"Development and Biological Evaluation of a novel Aurora A Kinase inhibitor"
ChemBioChem 10(3):464-78 (2009).

Vanneste D, Takagi M, Imamoto N and Vernos I.
"The role of Hklp2 in the stabilization and maintenance of spindle bipolarity."
Current Biology 19(20):1712-7 (2009).

Saito K, Chen M, Bard F, Chen S, Zhou H, Woodley D, Polischuk R, Schekman R and Malhotra V.
"TANGO1: a protein required for collagen VII packing into COPII transport carriers."
Cell 136:891-902 (2009).

von Blume J, Duran JM, Forlanelli E, Alleaume AM, Egorov M, Polishchuk R, Molina H, Malhotra V.
"Actin remodeling by ADF/cofilin is required for cargo sorting at the trans-Golgi network."
J Cell Biol 187(7):1055-69 (2009).

Azkona G, Amador-Arjona A, Obradors-Tarragó C, Varea E, Arqué G, Pinacho R, Fillat C, de la Luna S, Estivill X, Dierssen M.
"Characterization of a mouse model overexpressing beta-site APP-cleaving enzyme 2 reveals a new role for BACE2."
Genes Brain Behav. 2009 Sep 22. [Epub ahead of print]

Salichs E, Ledda A, Mularoni L, Albà MM, de la Luna S.
"Genome-wide analysis of histidine repeats reveals their role in the localization of human proteins to the nuclear speckles compartment."
PLoS Genet 5(3):e1000397 (2009).

Faucherre A, Pujol-Martí J, Kawakami K, López-Schier H.
"Afferent neurons of the zebrafish lateral line are strict selectors of hair-cell orientation."
PLoS One 4(2):e4477 (2009).

CORE FACILITIES

Core Facility: High-Throughput Screening Unit

Unit structure:
Unit Responsible: Raúl Gómez

Technicians: Anja Leimpek

SUMMARY

The High-Throughput Screening (HTS) and Robotics Facility is dedicated to providing automation and high-content screening services for a number of research projects. The facility staff provides advice in the initial experiment planning, training of the researchers on the instruments and assistance with the subsequent data analysis. The mission of the HTS and Robotics Facility is to automate a variety of simple and complex assays for high-content biochemical, cell-based, siRNA screening of a plethora of target classes.

FACILITY OVERVIEW

During January and February the HTS and Robotics Unit underwent its final phase of refurbishments that was completed with the installation of a laminar air flow cabinet from TELSTAR. It now comprises a platform for automated cell-based screens in an environment that complies with ISO 5 requirements and provides sterile and temperature controlled conditions for working with mammalian cell lines and tissues. The platform is a unique design - a customized assembly of sophisticated and sensitive pieces of equipment, including precise liquid handling systems, a cell incubator, a plate reader, and a microscope for automated acquisition and analysis of images (figure 1).

Fig 1.
Current HTS robotic platform, comprising a multimode plate reader, a widefield fluorescence screening microscope, as well as liquid handlers and a computer server for image processing

Taking over the management of the Dharmacon SMARTPool siRNA library, the HTS Unit was fully operational for performing its first genome-wide screen in April 2009. The first genome-wide siRNA screen was performed for the lab of Vivek Malhotra, with the goal of identifying genes implicated in mucus secretion. The primary screen gave more than 400 hits, which will be validated during 2010.

In addition to the siRNA screen, the HTS Unit performed automated ELISAs and automated immunofluorescence stainings for the quality control of the antibodies generated within the CRG Antibody Project.

A number of protocols have been automated and are now available for other projects:

1) Library aliquotation
2) siRNA transfection
3) Fluorescent staining procedure for nuclei detection and counting
4) Luminescence procedure for sensitive determination of peroxidase-labelled conjugates in immunoassays
5) ELISA
6) Immunofluorescence on fixed mammalian cells

All of the protocols listed above were implemented to be executed for a maximal capacity of 32 plates at a throughput of 18 plates/day (figure 2A).

For data analysis, a custom script with the following features was written in MATLAB (Figure 2B):

> Data normalization (B-score)
> Hits rank
> Hits distribution

A)

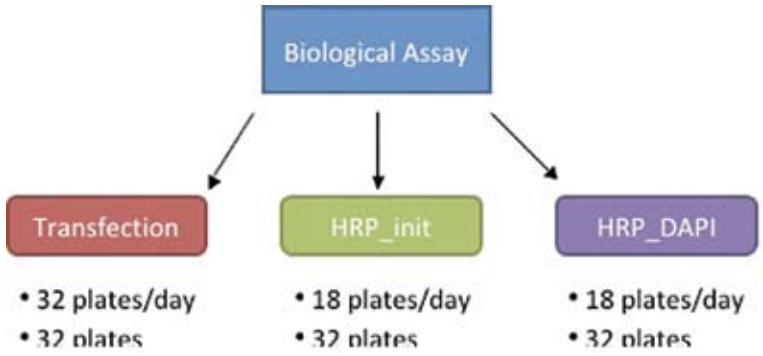
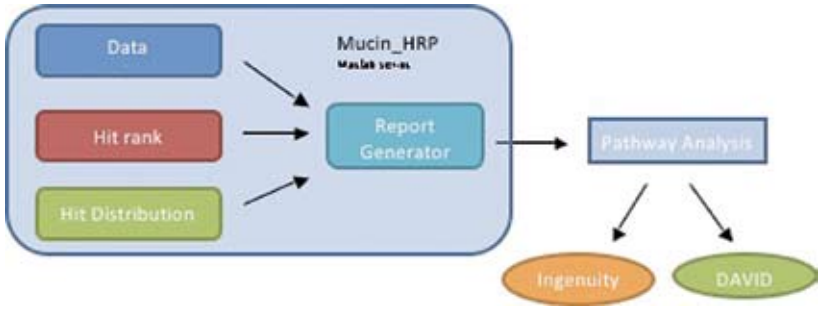


Fig 2.
General overview of the workflow of the mucin siRNA screen. A) Programs implemented for the robotic platform. B) MATLAB application for data analysis (Mucin_HRP)

B)



CORE FACILITIES

Core Facility: Bioinformatics Unit

Unit structure:

Unit Leader: Guglielmo Roma

Technician: Antonio Hermoso

SUMMARY

The Bioinformatics Core Facility aims to provide researchers at CRG, PRBB, and other external institutions with services of consultation, data analysis and management, software development, and access to computing resources. As further important commitment, the core will periodically organize courses and workshops in order to train biologists on the implementation of the main bioinformatics tools in their research.

FACILITY OVERVIEW

Starting in September 2009, the Bioinformatics Unit currently comprises two first members. In its initial stage, it has initiated communication with other core facilities and research groups in order to explore current bioinformatics and data analysis needs. Meanwhile, the team has provided ad-hoc collaborative support to existing projects, as well as started the implementation of a computational infrastructure that will be further deployed next year in close collaboration with the Bioinformatics and Genomics Programme and the IT department. As further development, additional staff members have been recently recruited and will join the team at the beginning of 2010.

SERVICES

The main goal of the unit is to set up a collaborative environment within CRG/PRBB and contribute to research projects providing support to scientists on several bioinformatics aspects.

Basic services:

> Consultation on bioinformatics methods and resources
Assistance with the computational aspects of research by providing support and consultation in relation to data management, statistical data analysis, and interpretation of experimental data, focusing primarily on high-throughput genomics technologies.

> Software evaluation, implementation, and training
Support on the evaluation and implementation of new bioinformatics software (gathering of users requirements, scheduling, implementation, validation, documentation, training, and end-users support).

> Customization of bioinformatics resources
Local implementation of genome-related software (e.g. genomics browsers like Ensembl, UCSC, and GBrowse) and most common bioinformatics tools (Galaxy, EMBOSS, among others).

Advanced services:

> High-throughput data analysis of genomics, transcriptomics, and proteomics datasets

> Downstream analysis for the interpretation of gene lists, including gene set enrichment and over-representation of biological functions using Gene Ontology terms;

> Identification of functional genomics elements, such as genes and regulatory sequences, through comparative genomics approaches;

> Construction of ad-hoc databases to allow systematically storage, management, analysis, and search of biologically relevant information;

> Development of bioinformatics scripts and more advanced pipelines to automate sequence analyses and genome annotation tasks

The unit will be also focused on the education of biologists in the analysis and interpretation of their experimental data through the provision of custom advice and training in bioinformatics as well as in the acquisition of basic programming skills.

Fig 1.
Genomic region surrounding human c-myc gene

PUBLICATIONS

Sanseverino W*, Roma G*, De Simone M, Faino L, Melito S, Stupka E, Frusciante L, Ercolano MR.
"PRGdb: a bioinformatics platform for plant resistance gene analysis."
Nucleic Acids Res. Epub 2009 Nov 11. [Epub ahead of print] (* joint first authors)

Romito A, Lonardo E, Roma G, Minchiotti G, Ballabio A and Cobellis G.
"Lack of Sik1 in Mouse Embryonic Stem Cells Impairs Cardiomyogenesis by Down-Regulating the Cyclin-Dependent Kinase Inhibitor p57kip2."
Plos One, in press.

HIGHLIGHTS

HIGHLIGHTS

GENE REGULATION PROGRAMME

"Luxury packages for genome messages"

The genome is the code that determines the biological characteristics of living things. Scientists generally assume that the genomic code is linear, that is to say its significance is determined by the specific DNA sequence. However, the linear DNA sequence of our chromosomes is more than 2 metres long. To accommodate all of this information in the diminutive space of the cellular nucleus, it is necessary to package it in the most efficient way. A molecular structure called chromatin is responsible for this packaging. The research of various groups from the Center for Genomic Regulation (CRG) suggests that, contrary to that which is generally accepted, the DNA packaging into chromatin also plays a key role in the process of gene regulation.

The instructions to make a living thing work, written in the language of DNA, are translated in the protein synthesis to be undertaken by cellular functions. The genetic material, or DNA, is found packaged in the cell in the form of chromatin. The DNA in chromatin is wrapped around proteins in units known as nucleosomes. In order to express a gene and achieve protein synthesis it is necessary to transcribe DNA into RNA. The RNA then undergoes a process called splicing in which some fragments (exons) link in such a way that they are used together to synthesise an RNA messenger and a protein.

This work has revealed that the position of the nucleosomes on the DNA coincides with that of the exons, facilitating the splicing process to generate RNA messengers, which are translatable into proteins. In this way, the architecture of the DNA packaging predicts the architecture of the RNA messengers: the genome fragments which contain important information appear to be more delicately packaged than other areas with less important contents. The packaging section of the DNA of our cells seems to be very busy differentiating between the most essential and accessory messages of our genome.

The findings are the result of the interdisciplinary collaboration between groups from the Bioinformatics and Genomics and Regulation of Gene Expression programmes of the Center for Genomic Regulation (CRG), making this a 100% "home grown" effort. Thanks to the collaboration between the diverse research groups, the work demonstrates that knowing the genetic code is not enough and that other factors which until now have not been accorded enough importance such as the architecture of DNA in chromatin, can play a very important role.

Reference:

Hagen Tilgner, Christoforos Nikolaou, Sonja Althammer, Michael Sammeth, Miguel Beato, Juan Valcárcel and Roderic Guigó.
"Nucleosome positioning as a determinant of exon recognition."
Nature Structural and Molecular Biology, September 3 2009, DOI: 10.1038/nsmb.1658 *http://www.nature.com/nsmb*

HIGHLIGHTS

DIFFERENTIATION AND CANCER PROGRAMME

"A step forward in cell reprogramming"

There are increasingly more research groups trying to discover the mechanisms of cell differentiation in order to reprogramme differentiated cells. Thomas Graf and his team have described a process of cell reprogramming which results in morphologically and functionally distinct cells with a 100% efficiency rate.

The researchers have used B lymphocyte precursor cells and reprogrammed them in order to transform them into macrophages using the inducible expression of only one transcription factor. Both types of cells form part of the immune system but are morphologically, structurally and functionally very different. Whereas type B lymphocytes are antibody-producing cells, macrophages are responsible, via phagocytosis, for the elimination of both foreign agents introduced into the body and dead tissue cells. In this way, the new macrophages induced by the system described by the investigators are bigger than the original cells, they contain different cellular organelles and the structure of their cytoskeleton is modified, so they have phagocytic capacity and respond to inflammation stimuli. Moreover, it was observed that the cells acquire their new form and function a few hours after induction. Two or three days later these lymphocytes have converted into completely autonomous macrophages.

The speed and efficiency of the reprogramming reached by Thomas Graf and collaborators has not been achieved in any other type of cell. For example, in studies of reprogramming of somatic cells to iPS cells there always exists a high percentage of the cellular population that is not reprogrammed. This makes the study of the reprogramming mechanism at a molecular level very difficult. For this reason, the new system published in the Cell Stem Cell journal, is a unique tool for the study of biochemical and biological aspects of cell reprogramming.

This study is part of the cell reprogramming project of the CRG, a project in which all of the research programmes of the centre collaborate. The ultimate goal of this project is to understand cell differentiation in various tissues of the body, a fundamental process in all multicellular organisms and one which we still understand very little of. In the future these studies will enable us to generate differentiated cells "a la carte" from cultured cell biopsies. In this way, not only will immunologic rejection be avoided, but the application of this in regenerative medicine would also be much simpler, and perhaps less dangerous, than the use of embryonic stem cells.

Reference:

Bussman et al.
"A Robust and Highly Efficient Immune Cell Reprogramming System."
Cell Stem Cell. 2009. DOI: 10.1016/j.stem.2009.10.004 *http://www.cell.com/cell-stem-cell/*

HIGHLIGHTS

GENES AND DISEASE PROGRAMME

"Genetic causes for psoriasis"

Psoriasis is a chronic inflammatory disease of the skin, affecting over 15 million people in Europe, and usually appearing between the ages of 15 and 40. It causes abnormal proliferation and differentiation of skin cells (excessive growth), due to changes in the skin, specifically the keratinocytes. Although the exact cause of psoriasis is still unknown, the disease is due to a combination of genetic and environmental factors. Certain drugs, infection, trauma, cold weather and stress are the main triggers of this skin disease, characterised by the presence of quite extensive skin lesions, although it may also affect the joints.

Psoriasis is most common among identical twins (monozygotic) and there are families in which psoriasis is highly prevalent, the evidence of which demonstrates that genetic causes are involved. Although scientists have identified some of the genetic factors involved in susceptibility to psoriasis (such as the major histocompatibility system or HLA), only a few factors are common to different populations, probably reflecting the exposure of humans to different environments.

The Genetic Causes of Disease Group led by Xavier Estivill at the CRG looked at samples from patients and control subjects in a comprehensive study of the genome. Researchers have explored regions of the genome, which vary in the number of copies of the genes they contain, known as CNVs (copy number variants). The investigation discovered that patients with psoriasis lack, with a significantly high frequency, both copies of two genes known as LCE3B and LCE3C. Samples from more than 2,500 subjects from Spain, Holland, Italy and the United States were studied, with similar results being obtained from all groups. This research, published in the journal Nature Genetics, showed that the absence of the LCE3B and LCE3C genes contributes to over 20% of the risk of developing psoriasis, it being the main genetic factor for susceptibility after the HLA locus.

The project gathered together a group of investigators from different countries (Italy, the United Kingdom, the Netherlands and the United States). The investigation is a milestone in the identification of the genes involved in multifactor diseases. CRG investigators Rafael de Cid and Eva Riveira have played a major role in developing the work, which has relied upon many clinical collaborators. Researchers are now studying additional processes that alter the entry barrier for microorganisms and further agents which can damage the skin and other tissues for which these genes may have a protective role. The work has immediate diagnostic implications and opens new avenues for the exploration of preventive and curative treatments for a disease that affects 2% of people and for which only palliative treatments are available.

The Catalan Government, the Ministry of Science and Innovation, and the European Commission supported this research. The results represent an advance in the implementation of the genome sequencing of various diseases, in which Spain has decided to play an important role by creating a new sequencing centre and financing projects in this field.

Reference:

De Cid, R. et al.
"Deletion of the late cornified envelope LCE3B and LCE3C genes as a susceptibility factor for psoriasis."
Nature Genetics, DOI: 10.1038/ng.313. *http://www.nature.com/ng/index.html*

"The bovine genome: new information about mammalian evolution and clues to efficient cattle"

The genomics era provides us with new decoded genomes. The Bovine Genome Sequencing project involved over 300 scientists from 25 countries, and researchers from the Centre for Genomic Regulation (CRG) played a major role. Francisco Cámara, Tyler Alioto and Roderic Guigó, from the CRG, contributed to the gene identification while Roderic Guigó led the international team for the experimental validation of the gene set. He also oversaw the studies of sequencing other genomes (human, mouse & rat) and his laboratory collaborated in this new international study published in *Science*.

The cattle sequence took six years to complete, annotate and analyse. The Bovine Genome Sequencing Project sequenced the genome of a female Hereford cow named L1 Dominette. The Hereford breed originated in England and is one of the most important food producing breed in the world. The previously studied mammals were laboratory animals and humans. The bovine genome sequencing is the first livestock genome to be studied. The genome, consisting of at least 22,000 genes, is more similar to that of humans than mice or rats; most cattle chromosomes correspond to part or all of human chromosomes, although the DNA is rearranged in certain areas.

The work showed that cattle and humans still conserve a high degree of similarity in the organisation of their chromosomal architecture, far more so than with humans and mice. It seems that, like humans, duplicated segments appear to have played a major role in the rearrangements of chromosomes in the cattle lineage.

Some of these rearrangements appear important to cattle biology as they impact on genes involved in immunity, lactation, metabolism and digestion. These changes could help explain the amazing ability of cattle to efficiently convert low-quality fodder into energy-dense meat and milk, processes long exploited by man. One particular genetic rearrangement involves the histatherin gene in mammary tissue, which undergoes special regulation during lactation. It produces a novel protein in milk which has anti-microbial activity. Other gene changes result in the transfer of immunoglobulin G to milk which passes on innate immunity to suckling calves from their dams.

The international consortium of researchers involved in the project agreed that the challenge for the future is the exploration of the bovine genome sequence in greater depth in order to fully understand the genetic basis of the evolutionary success of ruminants as this will provide opportunities to address some of the crucial current issues.

Reference:

The Bovine Genome Sequencing and Analysis Consortium, Christine G. Elsik, Ross L. Tellam, Kim Worley. *"The Genome Sequence of Taurine Cattle: A Window to Rumiant Biology and Evolution."* Science, vol. 324, doi:10.1126/science.1169588 (2009). *www.sciencemag.org*

HIGHLIGHTS

CELL AND DEVELOPMENTAL BIOLOGY PROGRAMME

"Tango1: a guide for protein transport inside the cell"

At least 15% of the genes in our genome encode for proteins that are secreted from cells. But, who is in charge of controlling traffic within the cell? Who decides the best travel programme for each newly made protein? What prevents these proteins from travelling to the wrong place? What averts traffic jams?

Until recently, it was thought that the proteins destined for secretion were packed into transport carriers without any specific guides or loaders. The transport of proteins is mediated by small containers, which package them in the Endoplasmic Reticulum and deliver them to the next station within the cell. This process is repeated; the load moves forwards and is ultimately secreted from the cell. However, recent evidence suggests that some proteins require special guides for their transport.

The work, published in the Cell journal, was directed by Vivek Malhotra, coordinator of the Cell and Developmental Biology Programme at the Centre for Genomic Regulation (CRG) and reveals the identity of a new protein called TANGO1 that places the load into transport carriers via a novel mechanism. TANGO1 is responsible for the transport of proteins such as collagen VII, which is too big to fit into a standard transport carrier. TANGO1 helps in creating a special larger transport carrier for collagen VII but, strangely, employs the same machinery and the same materials used for generating the standard small transport carriers. To summarise, using the same materials, TANGO1 builds a transatlantic liner in place of a rowing boat.

TANGO1 captures collagen VII with one end and holds the transport carrier with the other, preventing it from leaving the Endoplasmic Reticulum. In this state the transport carrier grows in size to accommodate the bulky collagen VII. Once a transport carrier big enough to accommodate collagen VII has been created, TANGO1 lets go of the collagen VII and the newly forming mega-transport carrier. The transport carrier is then allowed to depart and travel in the forward direction of secretion. These findings provide strong evidence that cells use specific guides for loading cargo and can make different sized carriers depending on the size of the load.

Upon secretion collagen VII interacts with other adhesive proteins and helps in the attachment of the epidermis to the dermis. Patients with genetic defects in collagen VII have a number of severe skin disorders which are collectively called Dystrophic Epidermolysis Bullosa (DEB) and Epidermolysis Bullosa Acquisita (EBA), respectively. TANGO1-mediated collagen VII export could reveal important insights into the process of skin biogenesis.

It is important to note that collagen VII is not only secreted by the skin cells and TANGO1 is present in all cells. Malhotra and colleagues suggest that TANGO1 is required for the secretion of proteins that are generally required for cell-cell attachment. Cancer cells disrupt this attachment and revealing the mechanism of TANGO1-dependent protein secretion of adhesive proteins could help the understanding of cancer metastasis.

Reference:

Saito K, Chen M, Bard F, Chen S, Zhou H, Woodley D, Polischuck R, Schekman R and Malhotra V. *"TANGO1 facilitates cargo loading at Endoplasmatic Reticulum exit sites."* Cell, March 6 2009, DOI: 10.1016/j.cell.2008.12.025. ***http://www.cell.com***

HIGHLIGHTS

SYSTEMS BIOLOGY PROGRAMME

"Even the simplest form of life is more complex than expected"

What are the bare essentials of life, the indispensable ingredients required to produce a cell that can survive on its own? Can we describe the molecular anatomy of a cell, and understand how an entire organism functions as a system? These are just some of the questions that scientists in a partnership between the Centre for Genomic Regulation (CRG) in Barcelona and the European Molecular Biology Laboratory (EMBL) in Heidelberg, set out to address.

Three papers, published back-to-back in Science and which provide the first comprehensive picture of a minimal cell, based on an extensive quantitative study of the biology of a bacterium, are the results of the work. The bacterium studied was ***Mycoplasma pneumoniae***, a small, single-cell bacterium that causes atypical pneumonia in humans. It is also one of the smallest prokaryotes that does not depend on a host's cellular machinery to reproduce. This is why the researchers, who set out to characterise a minimal cell, chose ***M. pneumoniae*** as the model: it is complex enough to survive on its own, but also small and, theoretically, simple enough to represent a minimal cell and allow a global analysis.

The investigation went beyond the genome, and the researchers approached the bacterium on three different levels. One team of scientists described the transcriptome of ***M. pneumoniae***, identifying all the RNA molecules, or transcripts, produced from its DNA, under various environmental conditions. Another defined all of the metabolic reactions that occurred within it, collectively known as its metabolome, under the same conditions. A third team identified every multi-protein complex the bacterium produced, thus characterising its proteome organisation.

The results revealed that more than genome sequencing is necessary to understand how organisms work and that even the simplest of cells is more complex than expected. Moreover, the study uncovered fascinating new facts relevant to bacterial biology. For example, when studying both its proteome and its metabolome, the scientists found many molecules were multifunctional, with metabolic enzymes catalysing multiple reactions. They also showed that the regulation of this bacterium's transcriptome is much more similar to that of eukaryotes than previously thought. Another surprise was the fact that, although it has a very small genome, this bacterium is incredibly flexible and readily adjusts its metabolism to drastic changes in environmental conditions. This adaptability and its underlying regulatory mechanisms mean that ***M. pneumoniae*** has the potential to evolve quickly, and all the above features are also shared with other, more evolved organisms.

Luis Serrano and the researchers from his laboratory of Design of Biological Systems at the CRG could build a complete overall picture of how the whole organism functions as a system. This study required a wide range of expertise to understand the molecular organisation of ***M. pneumoniae*** at such different scales and integrate all the resulting information. The group will amalgamate all of the results in an engineering and monitoring project of ***M. pneumoniae*** as a cellular organelle. They continue trying to understand how this bacteria works in order to be able to manipulate it in the future and, for example, use it to cure diseases.

References:

Güell, M. et al.
"Transcriptome Complexity in a Genome-Reduced Baterium."
Science (2009) Vol. 326. DOI: 10.1126/science.1176951. ***www.sciencemag.org***

Yus, E. et al.
"Impact of Genome Rection on Bacterial Metabolism and Its Regulation."
Science (2009) Vol. 326. DOI: 10.1126/science.1177263. ***www.sciencemag.org***

Kühner, S. et al.
"Proteome Organization in a Genome-Reduced Bacterium."
Science (2009) Vol. 236. DOI: 10.1126/science.1176343. ***www.sciencemag.org***

Appendix 1

VIII ANNUAL SYMPOSIUM OF THE CENTRE FOR GENOMIC REGULATION

"Stem Cells, Differentiation and Cancer"

The CRG symposia have been highly successful in bringing together some of the best scientists in their field. For the VIIIth symposium we merged the fields of stem cell research and differentiation with that of cancer. Differentiation and cancer have long been considered as being two sides of the same coin, in particular in the case of leukemias. Thus, many leukemias essentially represent differentiation diseases, in which specific oncogenes (often over-expressed or fused transcription factors) cause the proliferation of aberrantly or incompletely differentiated blood cells. Some of the successful therapies revolve around the concept to enable their differentiation, bringing the cells into a resting state, the best example of which is the treatment of PML RAR induced promyelocytic leukemia with retinoic acid. This field has been pioneered by Pier Giuseppe Pellici.

More recently a strong link has also been established between stem cells and cancer. Studies in the 90s by John Dick and colleagues showed that acute myeloid leukemia is a heterogeneous population, consisting of a small fraction of cells with self-renewal potential (the 'cancer stem cells') and the bulk of the leukemic cells that are essentially non-dividing cells. This concept has now been extended to other types of tumors, such a mammary and colon carcinomas, although it appears not to be a universal principle of all cancers.

The question of whether or not a given tumor contains a subset of self-renewing cells with tumor initiating potential is not just of academic interest, as it strongly influences the approaches chosen for therapy. If a defined subset of cancer stem cells can be identified, all efforts should be directed at eradicating them; if essentially all cells in a tumor proliferate, the approach must be directed at the bulk of the population.

Interestingly, there are now several studies that establish a connection between stem cells, differentiation and cancer. For example, it has been shown in mouse models by Tariq Enver and others, that certain types of leukemias induced by aberrant transcription factors originate not from stem cells, as it had been originally assumed, but from committed progenitors, which are then directed to acquire self-renewal potential and to exhibit a stem cell specific gene expression signature. In addition, as shown by Robert Weinberg and colleagues, mammary epithelial cells induced to become tumorigenic by undergoing an epithelial to mesenchymal transition acquire stem cells properties. Together, these findings suggest that at least some forms of cancer arise by the reprogramming of differentiated cells under the influence of aberrant transcription factors, and undergo extensive remodelling of their epigenetic machinery. This remodelling may also involve specific microRNAs, as has been shown for a number of tumor types. All of the above-mentioned scientists, except Tariq Enver, participated at the conference.

Additionally, the field of epidermal stem cells (and its links with cancer) was represented by one of the leading scientists in the field. Elaine Fuchs and with Fiona Watt (who wasn't finally able to participate), have made seminal contributions to isolate, and characterize epidermal stem cells. They have identified key pathways that control their behaviour during homeostasis and cancer onset and progression. As well, they have studied in recent years how and when epidermal stem cells are specified early during embryonic development, and how they make lineage choices to form the entire epidermal compartment. More specifically, they have described a role for wnt TGFb, Bmp and Notch signalling as essential players in epithelial stem cell differentiation and lineage commitment.

The role of one of the best-characterized oncogenes, Myc, and its contribution to the establishment and maintenance of the cancer phenotype was discussed by Gerard Evan. Finally, the importance of epigenetic gene programming and reprogramming, telomers, either at specific loci or genome-wide was illustrated by Wolf Reik, Maria Blasco and Richard Young.

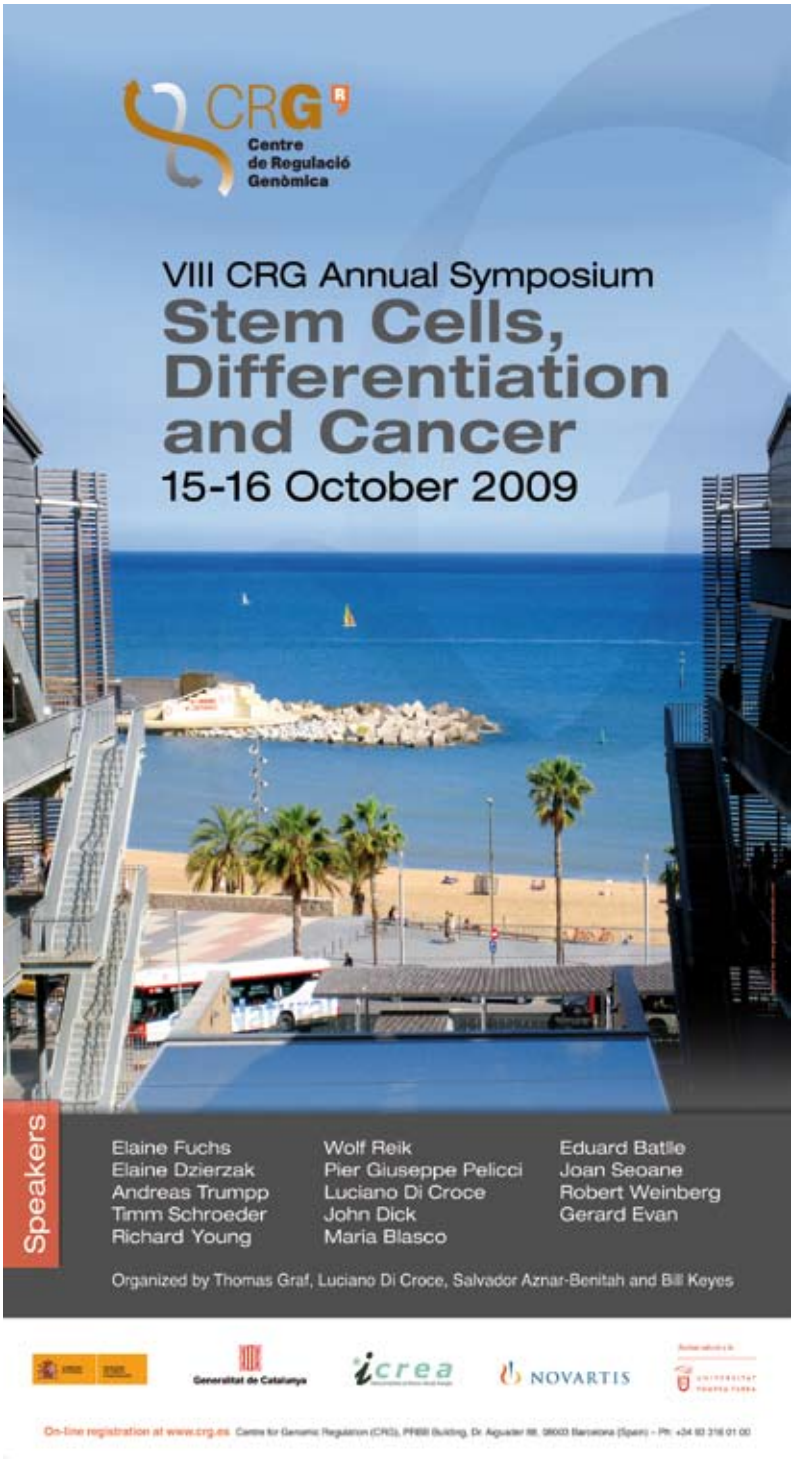
In this scenario, the VIII CRG Symposium titled "STEM CELLS, DIFFERENTIATION AND CANCER" was held on 15th and 16th October 2009, at the PRBB Auditorium in Barcelona. The symposium, gathered some of the worldwide leaders in the fields of stem cells, differentiation and cancer and became a magnet for the attendance of the Spanish and the international scientific community working in these areas. The symposium was open to everyone in the community and attracted a national and international audience and included 16 presentations of high level speakers divided into two blocks.

All attendants agreed that the speakers selected were the best worldwide in their respective areas and according to the different sessions in which the symposium was divided. The topics presented, as well as the discussions afterwards were of high interest for the scientists from the CRG and the entire scien-

tific community in the area of Barcelona, as well as for all the attendants coming from the rest of Spain and Europe. On the other hand, it is very important to highlight that contacts amongst several of the speakers' groups were established, which might crystallize in future collaboration projects.

The high number of attendants (more than 250), the level of the invited speakers and the discussions contributed to internationalize the scientific image of Barcelona, Catalonia and Spain.

We do believe the final result of this symposium was really interesting for the attendants, due to the relevance and prestige of experts in this area. The contents and format of the symposium worked as a forum, clearly suitable for the consecution of these objectives and for this reason, we consider it a great success.



Appendix 2

PRBB-CRG SESSIONS 2009

04-12-09 Thomas Risler
Laboratoire Physico-Chimie Curie, CNRS-UMR 168, Universite Pierre et Marie Curie Paris VI, Institut Curie, Paris, France
"Homeostatic pressure driving tumor growth: From general physics arguments to cellular pathways"

27-11-09 Rudy Dekeyser
VIB, Flanders Institute of Biotechnology, Belgium
"When is your research sexy enough for industry?"

13-11-09 Ari Helenius
Institut f. Biochemie, ETH Zürich, Switzerland
"How viruses enter their host cells"

06-11-09 Irene Bozzoni
Dept. of Genetics and Molecular Biology, Sapienza-University of Rome, Rome, Italy
"Role of small non coding RNAs in the physiopathology and therapy of Duchenne Muscular Dystrophy"

04-11-09 Steve Cohen
Chief Executive Officer, Developmental Biology Program, Temasek Life Sciences Laboratory, Singapore
"MicroRNA Functions"

30-10-09 Weimin Zhong
Department of Molecular, Cellular and Developmental Biology, Yale University, USA
"Mechanisms of Stem-Cell Homeostasis"

09-10-09 Marcos Gonzalez-Gaitan
Departament of Biochemistry Sciences II, Geneva, Switzerland
"Temporal versus spatial computation of DPP during growth control"

02-10-09 Kevin Davies
Editor-in-Chief, Bio-IT World, Needham, USA
"Personal Genomics and Next generation Sequencing"

25-09-09 Jordi Torres-Rosell
Eukaryotic Cell Cycle Lab, IRB Lleida, Spain
"The Smc5/6 complex: a 'resolvin' for DNA-mediated cohesion?"

14-09-09 Paul Fox
Prof. of Molecular Medicine, Department of Cell Biology, Lerner Research Institute / NC10 Cleveland Clinic, USA
"The GAIT system: A gatekeeper of inflammatory gene expression"

04-09-09 Bill Hansson
Department of Evolutionary Neuroethology, Max Planck Institute for Chemical Ecology, Jena, Germany
"Olfactory Evolution"

31-07-09 Pilar Blancafort
Dept. of Pharmacology and Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, USA
"Reprogramming self-renewal in breast cancer with Artificial Transcription Factors"

27-07-09 Pia Cosma
Inborn Errors of Metabolism Research Programme, Telethon Institute of Genetics and Medicine (TIGEM), Napoli, Italy
"Periodic activation of Wnt signalling: a toggle switch for mouse somatic cell reprogramming"

17-07-09 Josep Lluís Sanfeliu
Ysios Capital Partners, Barcelona, Spain
"Ingredients and timing to create a new biotech company"

26-06-09 Haig H. Kazazian, Jr.
Seymour Gray Prof. of Molecular Medicine, Department of Genetics, University of Pennsylvania School of Medicine, USA
"Biology of human L1 retrotransposon"

12-06-09 Manolis Dermitzakis
Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK
"Advances in genomics of human gene expression"

05-06-09 Hermann Bujard
Director, EMBO, Heidelberg, Germany
"Studying gene functions via tetracycline controlled transcription activation: principles, advances, applications"

29-05-09 Lluis Montoliu
Departamento de Biología Molecular y Celular, Centro Nacional de Biotecnologia (CNB-CSIC), Madrid, Spain
"EMMA (The European Mouse Mutant Archive), now in Spain"

18-05-09 Carlos Brody
Dept.of Molecular Biology, Princeton University, Princeton, USA
"Functional separation between "what" and "when" in prefrontal cortex"

08-05-09 Witold Filipowicz
Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland
"Mechanisms and regulation of miRNA-mediated repression in mammalian cells"

22-04-09 Danny Reinberg
HHMI Investigator, Biochemistry Dept., NYU School of Medicine-Smilow Research Center, New York, USA
"A molecular understanding of Epigenetics"

17-04-09 Alexander Borst
Max Planck Institute of Neurobiology, Martinsried, Germany
"In search of the holy grail of fly motion vision"

03-04-09 Hollis Cline
Professor, Depts of Cell Biology and Chemical Physiology, The Scripps Research Insitute, La Jolla, USA
"Activity Dependent Mechanisms controlling Visual System Development"

27-03-09 Antoine H.F.M. Peters
Friedrich Miescher Institute for Biomedical Research (FMI), Novartis Research Foundation, Basel, Switzerland
"Epigenetic control of early mouse development: integrating parental contributions into embryonic programs"

20-03-09 Jürgen Meier
European Patent Attorney, Vossius & Partner, Germany
"Biotech and Patents: rumors, speculations and the truth"

13-03-09 Amos Bairoch
Swiss Institute of Bioinformatics, Swiss-Prot Group, Geneve, Switzerland
"The consequences of the first complete annotation of the human 'proteome': a new database, collaborative efforts and... back to the wet lab!"

05-03-09 Esteban González Burchard
Associate Professor, Depts. of Biopharmaceutical Sciences & Medicine, University of California, San Francisco, USA
"The Importance of Race/Ethnicity in Biomedical Research and Clinical Practice"

27-02-09 Manel Esteller
Programa de Epigenetica y Biología del Cancer, IDIBELL-ICREA, Barcelona, Spain
"Cancer Epigenetics"

20-02-09 Gero Miesenböck
Department of Physiology, Anatomy and Genetics, University of Oxford, UK
"Writing to Memory"

06-02-09 Nikolaus Rajewsky
Max Delbruck Center for Molecular Medicine, Berlin-Buch, Germany
"microRNAs and post-transcriptional gene regulation"

30-01-09 Iiris Hovatta
Department of Medical Genetics, University of Helsinki, Finland
"A cross-species neurogenomics approach reveals novel candidate genes for anxiety disorders"

22-01-09 Juan Martínez
Centro Andaluz de Biología del Desarrollo, Universidad de Olavide - CSIC, Sevilla, Spain
"The genetic control of eye morphogenesis in medakafish"

PROGRAMME & CORE FACILITIES SEMINARS 2009

GENES & DISEASE PROGRAMME

14-12-09 Carmen Sandi
Laboratory of Behavioral Genetics, Brain Mind Institute, Ecole Polytechnique Federale de Lausanne (EPFL), Switzerland
"Understanding the contribution of stress to cognitive function"

16-11-09 Maria V. Sanchez-Vives
ICREA Professor, IDIBAPS, Barcelona, Spain
"Cerebral Cortex: The rhythms and the players"

05-10-09 Mel Slater
ICREA Professor, Universitat de Barcelona, Spain
"Immersive Virtual Environments in the Study of Body Ownership Illusions"

28-05-09 Anne Bowcock
Department of Genetics, Washington University School of Medicine, USA
"Genetics of psoriasis: A complex disease of the skin and joints"

18-05-09 Alexander Cupido
Department of Neuroscience, Erasmus University Medical Center, Rotterdam, The Netherlands
"The Erasmus Ladder: an Automated System to Measure Motor Performance and Associative Motor Learning in Mice."

03-04-09 Luciano Sobrevals
CIMA, Universidad de Navarra, Pamplona, Spain
"sv40 and aav vectors expressing insulin-like growth factor i revert establish cirrosis in rats"

SYSTEMS BIOLOGY PROGRAMME

04-12-09 Tine Arnvig
Division of Mycobacterial Research, NIMR, MRC, London, UK
"Regulatory RNA in Mycobacterium tuberculosis"

03-12-09 Mark Matzas
Febit Holding GmbH, Heidelberg, Germany
"New generation gene synthesis: performance gain by a new combination of high throughput technology"

23-11-09 Jacek Puchalka
Systems and Synthetic Biology, Helmholtz-Centre for Infection Research (HZI), Braunschweig, Germany
"Navigating the microbial metabolic landscape through genome scale constraint-based modelling"

20-11-09 Luke Jostins
Statistical and computational genetics, Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK
"Biologically-Inspired Computing for Biological Problems: A Comparison of Parallel Global Optimisation Algorithms for Reverse Engineering Gene Networks"

13-11-09 Fabio Mohn
Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland
"Epigenome Dynamics During Stem Cell Differentiation As A Function Of DNA Sequence"

28-10-09 Leonardo Trabuco
Theoretical and Computational Biophysics Group, University of Illinois, Urbana, USA
"Investigating the mechanisms of protein synthesis using multi-resolution structural data"

21-10-09 Yolanda Schaerli
Department of Chemistry and Biochemistry, University of Cambridge, Cambridge, UK
"Microfluidic droplets - the unit operation of DNA amplification"

14-10-09 Troy Shirangi
University of Wisconsin-Madison, USA
"Evolution of behavior and pheromone signals in Drosophila"

08-10-09 David Jarriault
INRA, Versailles, France
"Sex pheromone coding and its plasticity in the moth Agrotis ipsilon"

05-10-09 Toru Yao
RIKEN, AIST and JST, Japan
"RIKEN and Systems Biology in Japan"

10-07-09 Chris Sander
Computational and Systems Biology, Memorial Sloan Kettering Cancer, New York City, USA
"Systems Biology of Cancer Cells"

07-07-09 Madan Babu
MRC Laboratory of Molecular Biology, University of Cambridge, Cambridge, UK
"The impact of genomic neighbourhood on human transcriptome evolution"

12-06-09 Xavier Trepat
Unitat de Biofísica i Bioenginyeria, Universitat de Barcelona, and Institute For Bioengineering of Catalonia, Barcelona, Spain
"Cytoskeletal Fragility and Collective Cell Migration"

09-06-09 Thomas Nowotny
CCNR, Informatics, University of Sussex, Brighton, UK
"Large (neural) network simulations in the NVidia CUDA parallel computing framework"

25-05-09 Nick Monk
Division of Applied Mathematics, School of Mathematical Sciences, University of Nottingham, Nottingham, UK
"Delay-driven oscillations in the Notch signalling network"

22-05-09 Alexis Gallagher
Department of Zoology, Oxford University, Oxford, UK
"Can we explain the evolution of complexity? - The prospects of a formal theory of evolvability."

18-05-09 Vincent Hakim
Laboratoire de Physique Statistique, ENS, Paris, France
"What can one learn about the structure of gene networks by simulating evolution in silico?"

20-04-09 Andrew Millar
School of Biological Sciences, The University of Edinburgh, Edinburgh, UK
"Flexible Regulation in Biological Clockworks"

14-04-09 Bertram Gerber
Dept. of Genetics and Neurobiology, University of Wuerzburg, Wuerzburg, Germany
"Neurogenetics of learning in Drosophila"

07-04-09 Leah Herrgen
Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany
"Time and time again - The regulation of segmentation clock period in zebrafish"

26-03-09 Robert Hoffmann
Computational Biology Center, Memorial Sloan-Kettering Cancer Center, New York, USA
"iHOP and WikiGenes - Information Management Strategies for the Life Sciences"

05-02-09 Nicola Abrescia
Structural Biology Unit, CIC bioGUNE, Derio, Spain
"Closer and closer to Eukarya: The crystal structure of the complete 13-multi subunits archaeal RNAP"

26-01-09 Gregory J. Stephens
Lewis-Sigler Institute for Integrative Genomics, Princeton University, USA
"More bits for behavior: from C elegans movement toward the principles of animal action"

BIOINFORMATICS & GENOMICS PROGRAMME

17-12-09 Hideki Innan
Graduate University for Advanced Studies, Hayama 240-0193, Japan
"Coevolution of Duplicated Genes by Gene Conversion"

10-12-09 Andrea Tanzer
Institute for Theoretical Chemistry, TBI University of Vienna, Austria
"Evolution of miRNAs and other small ncRNAs"

10-12-09 Gabriel Valiente
G.V., Technical Universty of Catalonia, Spain
"Accurate Taxonomic Assignment of Short Pyrosequencing Reads"

19-11-09 Steve Mount
Department of Cell Biology and Molecular Genetics, University of Maryland, USA
"Some human polymorphisms affecting splicing"

09-09-09 Michael Galperin
"What Part of the Genome Sequence Do You Not Understand?"
"What Part of the Genome Sequence Do You Not Understand?"

15-07-09 Rory Johnson
Stem Cell and Developmental Biology Group, Genome Institute of Singapore
"Investigating Gene Regulatory Networks in Development, Disease and Evolution"

09-07-09 Angelika Merkel
Molecular Ecology Lab, School of Biological Sciences, University of Canterbury in Christchurch, New Zealand
"Evolution of short tandem repeats (microsatellites) in yeast"

08-07-09 Jesse Shapir
Alm Laboratory for Microbiology, Massachusetts Institute of Technology, USA
"Sympatric speciation in marine vibrios is driven by a few ecologically relevant loci"

07-07-09 Roman Valls Guimerà
esCERT-UPC, Barcelona, Spain
"Cloud computing for Biology"

05-06-09 Michael Laessig
Statistical Physics and Quantitative Biology, Institute for Theoretical Physics University of Cologne, Germany
"From fitness landscapes to seascapes: Non-equilibrium dynamics of selection and adaptation"

28-05-09 Chaysavanh Manichanh
Institut de Recerca Vall d'Hebron, Barcelona, Spain
"Gut Microbiota 'Transplantation' and High-throughput sequencing"

04-05-09 Knut Reiner
Freie Universität Berlin, Institut für Informatik, AG Algorithmische Bioinformatik, Germany
"Algorithms for Next Generation Sequencing. Efficient implementations for Read mapping and Insert Sequencing using the SeqAn C++ library"

30-04-09 Martijn Huynen
Nijmegen Centre for Molecular Life Sciences, The Netherlands
"Evolution of the mitochondrial proteome and its application for protein function prediction"

24-04-09 Daniel Zerbino
PANDA Nucleotides Group, European Bioinformatics Institute (EBI), Cambridge, UK
"Assembling and comparing genomes with de Bruijn graphs"

17-04-09 Ariel Fernández
Department of Bioengineering, Rice University, Houston, USA
"Translationalm ideas in molecularly targeted anticancer therapy"

03-04-09 Charles Chapus
Dept. of Organismic & Evolutionary Biology, Museum of Comparative Zoology, Harvard University, Cambridge, USA
"Genome Evolution in Reptilia: In silico mapping to the chicken genome of 12,000 BAC-end sequences from two reptiles and a basal bird"

02-04-09 Marc. A Martí-Renom
Head of the Structural Genomics Unit, Bioinformatics & Genomics Department, Prince Felipe Research Center, Valencia, Spain
"Towards a third dimension in cell regulation: ligands, RNA and genomes"

02-04-09 Josep Abril
Department of Genetics, UB, Barcelona, Spain
"Planarian Proteomics: A Glimpse on Neoblasts Machinery"

19-03-09 Tom Gingeras
Head of Functional Genomics, Cold Spring Harbor Laboratory, NY, USA
"New Classes of Functional Short RNAs and Chromosome-wide Transcriptional Networks"

06-03-09 Elia Stupka
Program Manager, CBM-Cluster in Molecular Biomedicine, Trieste, Italy
"How much (or little!) do we understand about genomes?"

12-02-08 Thomas Lingner
Department of Bioinformatics, Institute for Microbiology and Genetics, Georg-August-University of Göttingen, Germany
"Alignment-free analysis of protein sequences with machine learning techniques"

GENE REGULATION PROGRAMME

17-12-09 Hidehito Kuroyanagi
Laboratory of Gene Expression, Graduate School of Biomedical Science, Tokyo Medical and Dental University, Japan
"Tissue-specific alternative splicing regulation in C. Elegans"

13-11-09 Laurence Wurth
Institut de Biologie Moléculaire et Cellulaire Centre National de la Recherche Scientifique, Strasbourg, France
"A common assembly pathway for selenoprotein mRNPs and small non coding RNPs"

25-07-09 Reinhard Lührmann
Max Planck Institute for Biophysical Chemistry, Göttingen, Germany
"Structure and function of the spliceosome"

23-07-09 Karla Neugebauer
MPI of Molecular Cell Biology and Genetics, Dresden, Germany
"Co-transcriptional spliceosome assembly and mRNP formation in vivo"

22-07-09 Michael Sattler
TU München and Helmholtz Zentrum, München, Germany
"Structural biology and NMR as a tool for studying protein-RNA interactions"

21-07-09 Bertrand Seraphin
CGM-CNRS, Gif-sur-Yvette, France
"Purification and characterization of RNA-protein and protein-protein splicing factors"

20-07-09 Eric Westhof
IBMC/CNRS, Strasbourg, France
"The elementary modules of RNA architecture and their participation in some RNPs"

20-07-09 Jernej Ule
MRC Laboratory of Molecular Biology, Cambridge, UK
"Mapping RNA-protein interactions with roles in pre-mRNA splicing"

19-05-09 Bassem Hassan
Laboratory of Neurogenetics, Flanders Interuniversity Institute of Biotechnology (VIB), University of Leuven (KUL) School of Medicine, Belgium
"Proneural Genes: Master Switches in Development and Cancer"

14-05-09 Mikel Zaratiegui
Plant Genetic Research, Cold Spring Harbor Laboratory, New York, USA
"The cell cycle and DNA damage signalling pathways regulate heterochromatin via RNAinterference"

03-03-09 Brendan Frey
University of Toronto, Canada
"The Code Underlying Tissue-Regulated Splicing"

09-02-09 John Castle
Group Leader at Rosetta Inpharmatics, LLC, Merck & Co., Inc., Seattle, USA
"Application of deep sequencing for RNA expression and DNA variation"

04-02-09 Xiang-Dong Fu
Principal Investigator, Dept. of Cellular & Molecular Medicine, University of California San Diego (UCSD), USA
"Molecular mechanism of chromosome translocation in cancer"

CELL & DEVELOPMENTAL BIOLOGY PROGRAMME

05-11-09 Louis Gervais
Principal Investigator, Professor (IBMB-CSIC), IRB, Barcelona, Spain
"In vivo cell elongation and lumen formation in a single cell"

05-11-09 Marco Milán
Development and Growth Control Laboratory, IRB, Barcelona, Spain
"Micro-RNAs and growth control"

09-09-09 Matteo Rauzi
Institut de Biologie du Développement de Marseille-Luminy, France
"Mapping Subcellular Tensile Forces Controlling Morphogenesis"

07-09-09 Ira Mellman
Vice president of Genentech, Inc in San Francisco, California, USA
"The Structural basis of EGF receptor signaling in living cells: drug discovery meets cell biology"

13-07-09 Pedro Carvalho
Postdoctoral Fellow, Department of Cell Biology, Harvard Medical School, Boston, USA
"Molecular mechanisms of ER-associated protein degradation"

03-06-09 Satyajit Mayor
National Center for Biological Sciences, Bangalore, India
"Active organization of membrane components in living cells: implications for the formation of membrane rafts"

22-05-09 Mike Rossner
Editor of Rockefeller University Press, USA
"What's in a picture? The temptation of image manipulation"

20-04-09 Ruth Kroschewski
ETH Zurich, Institute of Biochemistry, Zurich, Switzerland
"In vitro organogenesis: Mechanisms of multicellular self-organization"

19-02-09 Alexandre Vendrell
UPF, Barcelona, Spain
"SCF/CDC4 regulates Msn2 and Msn4 gene expression to counteract Hog1 induced lethality"

09-02-09 Suzanne Pfeffer
Stanford University School of Medicine, Stanford, California, USA
"Regulation of Receptor trafficking by small GTPases"

26-01-09 Verónica Becher
Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Argentina
"What do different chromosomes have in common? (An exhaustive but efficient computation)"

DIFFERENTIATION & CANCER PROGRAMME

01-12-09 Travis Stracker
IP Laboratorio de Inestabilidad Genómica y Cáncer, Institute for Research in Biomedicine (IRB), Barcelona, Spain
"The DNA damage response: preserving chromosome integrity and preventing tumorigenesis"

30-09-09 Pia Cosma
TIGEM Associate Investigator Naples, Italy
"Is cell signalling a key controlling mechanism for somatic cell reprogramming?"

20-07-09 Nadia Rosenthal
Group Leader and Senior Scientist, EMBL Monterotondo, Italy
"Enhancing mammalian regeneration"

13-07-09 Marta Flandez
Montefiore Medical Center, Department of Oncology, Bronx, New York, USA
"Wnt, KLF4 and Notch in intestinal homeostasis and the development of intestinal tumors"

29-06-09 Bogi Andersen
UC Irvine, California, USA
"Clock genes ad the hair growth cycle"

16-06-09 Urs Albrecht
Dept. of Medicine, Div. of Biochemistry. University of Fribourg, Switzerland
"Clocks, Brain Function and Dysfunction"

20-05-09 Hanane Boukarabila
Centre d'Immunologie de Marseille-Luminy (CIML), Université de la Méditerranée, Campus de Luminy, Marseille, France
"The PRC1 Polycomb Group complex interacts with PLZF/RARA to mediate leukaemic transformation"

19-05-09 Alea Mills
Cold Spring Harbor Laboratory, New York, USA
"From chromosome engineering to chromatin remodeler: CHD5 is a tumor suppressor mapping to human 1p36"

31-03-09 Joan Seoane
ICREA Research Professor, Vall d'Hebron Institute of Oncology, Barcelona, Spain
"TGF-beta in cancer"

02-03-09 Chris van Oevelen
New York University School of Medicine, New York University Cancer Institute, New York, USA
"Distinct Sin3 complexes play key roles during myogenesis"

24-02-09 Timm Schroeder
Institute of Stem Cell Research, Helmholtz Center Munche, Neuherberg / Munich, Germany
"New insights into blood cell development through live imaging"

09-02-09 Alberto Martín
Department of Immunology, University of Toronto, Medical Sciences Building, Toronto, Canada
"Genetic study of Cdk2 function and regulation. Involvement of Cdk2 in tumorigenesis"

02-02-09 Regina Mayor
Institut de Medicina Predictiva i Personalitzada del Càncer (IMPPC), Badalona, Catalonia, Spain
"New insights in colorectal cancer therapy illustrated by long range epigenetic silencing"

26-01-09 Eric Kallin
Howard Hughes Medical Institute, University of North Carolina at Chapel Hill, USA / Dept. of Biochemistry and Biophysics, Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, USA
"Polycomb and associated proteins: A role for Jhdm1b and Bmi1 in gene silencing?"

09-01-09 Meritxell Alberich Jorda
Hematology Oncology, Beth Israel Deaconess Medical Center (BIDMC), Boston, USA
"C/EBPgamma in hematopoiesis and leukemia"

CORE FACILITIES

05-10-09 Mikael Kubista/ Martina Reiter
TATAA Biocenter
"qPCR advanced course"

16-09-09 Marta Amat
Izasa
"EpiTec® and PyroMark"

03-07-09 Diana Hops
Advalytix
"One single cell PCR AmpliGrid system"

16-01-09 Julien Colombelli
Advanced Digital Microscopy, Institute for Research in Biomedicine (IRB Barcelona), Barcelona, Spain
"Laser Nanosurgery reveals morphogenetic forces and mechanosensing mechanisms"

SHARED

27-11-09 Christine Voguel
Institute for Cellular and Molecular Biology, University of Texas at Austin, Texas, USA
"Proteome dynamics - system-wide studies of translation regulation and protein turnover"

26-11-09 Gian Gaetano Tartaglia
Department of Chemistry, University of Cambridge, Cambridge, UK
"Towards quantitative predictions in cell biology"

26-11-09 Marc Marti
Bioinformatics & Genomics Department, Centro de Investigación Príncipe Felipe, Valencia, Spain
"Towards a third dimension in gene regulation"

26-11-09 Stephan Ossowski
Max Planck Institute for Developmental Biology, Tübingen, Germany
"SHORE and GenomeMapper: Analysis Tools for the Arabidopsis thaliana 1001 Genomes Project"

26-11-09 Sven Bermann
Department de Génétique Médicale, Swiss Institute of Bioinformatics, Lausanne, Switzerland
"Computational analysis of biological systems - from developmental patterning to genome-wide association studies"

25-11-09 Albert Tenesa
Colorectal Cancer Genetics Group, School of Molecular and Clinical Medicine, University of Edinburgh, Edinburgh, UK
"Genetics of colorectal cancer"

25-11-09 Frank Johannes
Groningen Bioinformatics Centre (GBiC), University of Groningen, Groningen, The Netherlands
"Dissecting the epigenetic architecture of complex traits"

25-11-09 Marc Rehmsmeier
Gregor Mendel Institute of Molecular Plant Biology, Vienna, Austria
"Computational Biology of RNA and DNA cis-regulatory elements"

25-11-09 Mónica Campillos
Dr. Peer Laboratory, EMBL, Heidelberg, Germany
"From human phenotypes to drug targets"

25-07-09 Gustavo Stolovitzky
IBM Computational Biology Center, Yorktown Heights, New York, USA
"Modeling and simulation of small and large scale biological networks"

16-02-09 Kirill Degtyarenko
EMBL-EBI and European Patent Office
"The fine art of chemical drawing and good annotation practice for chemical data"

Appendix 3

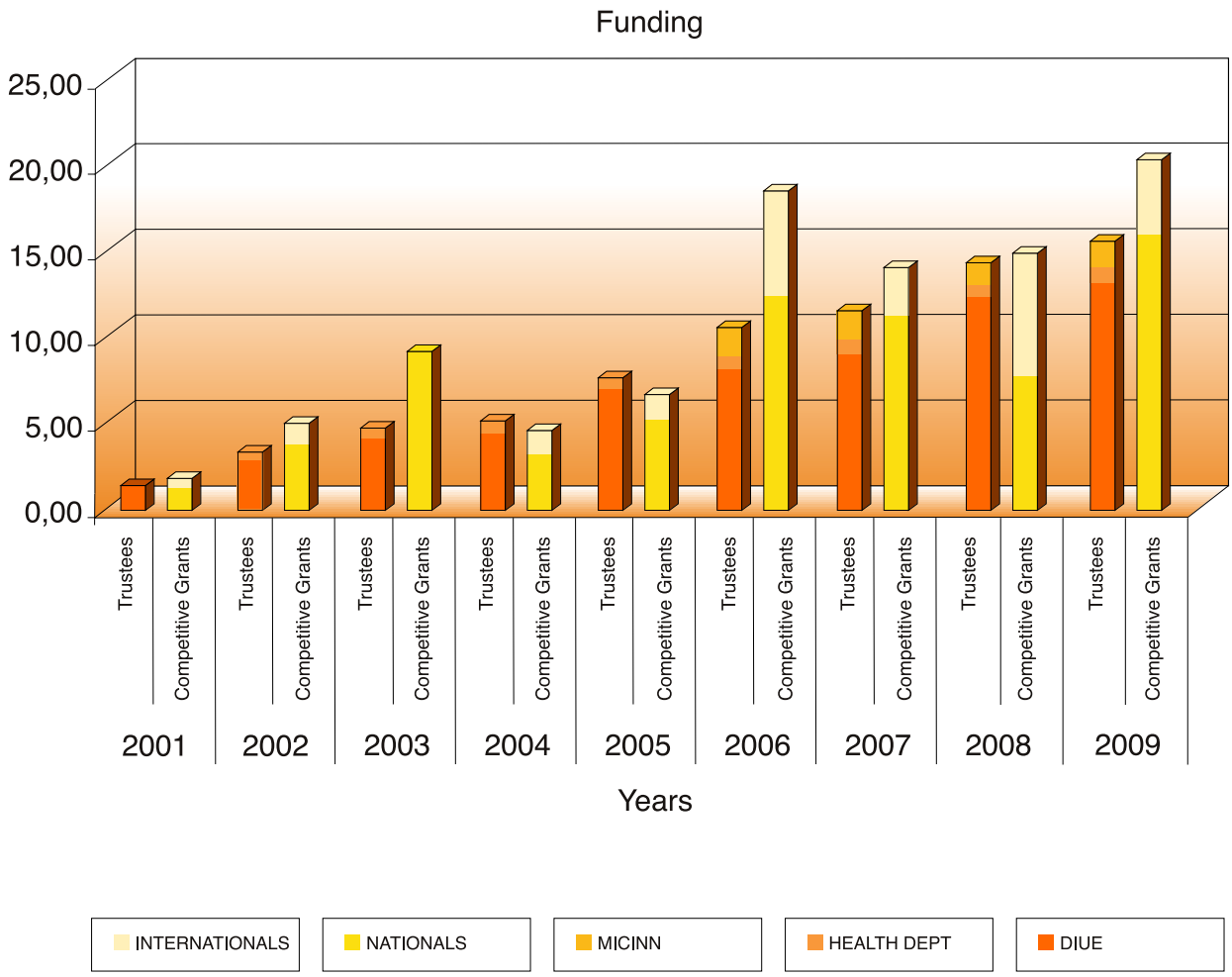
GRANTS

Competitive funding, contracts and sponsorship obtained from 1st January to 31st December 2009 by the CRG.

ORGANISM	AMOUNT (€)
MINISTERIO DE CIENCIA E INNOVACION	7.248.987 €
EUROPEAN COMMISSION	5.592.950 €
GENERALITAT DE CATALUNYA- FEDER	4.102.389 €
AGENCIA GESTIO D'AJUTS UNIVERSITARIS	959.818 €
FUND.CLINIC PER LA RECERCA BIOMEDICA	386.100 €
SWISS NATIONAL SCIENCE FOUNDATION	290.406 €
FUNDAÇAO PARAR A CIÊNCIA E A TECNOLOGIA	240.720 €
ASSOCIATION FOR INTERNATIONAL CANCER RESEARCH	204.288 €
CENTRE D'INNOVACIÓ I DESENVOLUPAMENT EMPRESARIAL	170.414 €
FUNDACION DESARROLLO INVESTIGACION GENOMICA Y PROTEOMICA	155.624 €
EUROPEAN MOLECUOLAR BIOLOGY ORGANISATION	121.514 €
CANADIAN INSTITUTES OF HEALTH RESEARCH	93.870 €
FONDATION JEROME LEJEUNE	80.000 €
HUMAN FRONTIER SCIENCE PROGRAM	63.600 €
EUROPEAN SCIENCE FOUNDATION	47.000 €
FUNDACIÓN ESPAÑOLA PARA LA CIENCIA Y TECNOLOGIA	30.000 €
FUNDACION CIEN	25.000 €
OTHER	201.913 €
TOTAL	**20.014.592 €**

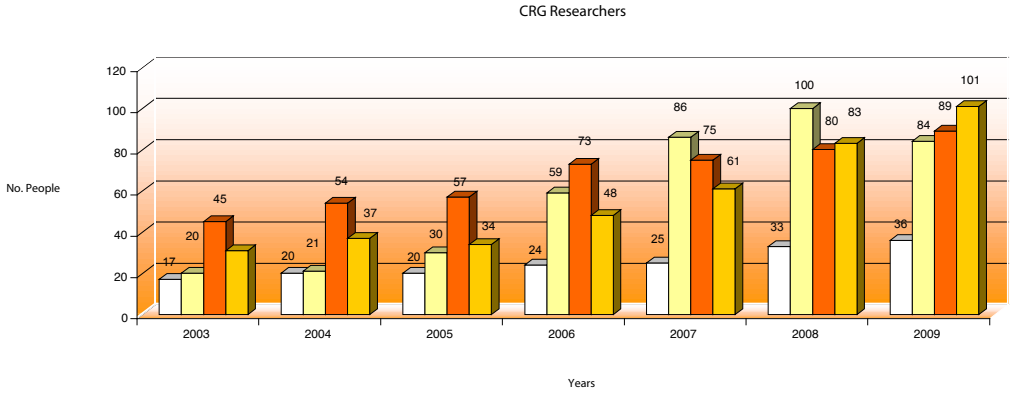
Appendix 4

FINANCE & PERSONNEL EVOLUTION AT THE CRG



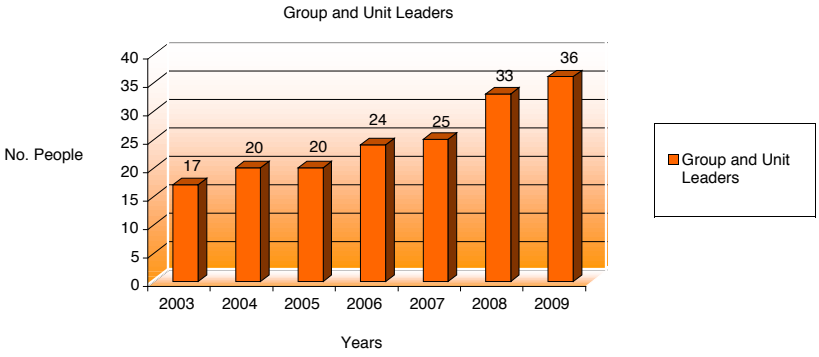
CRG Researchers

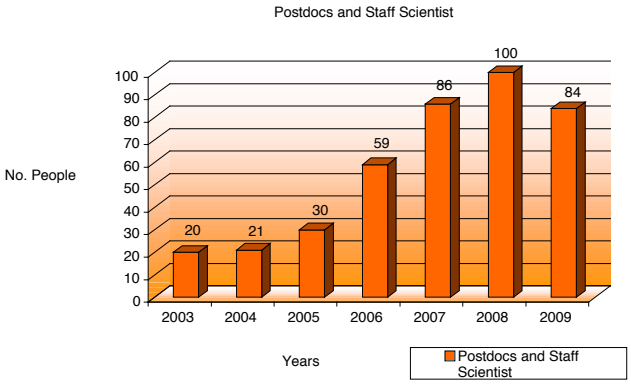
	Group and Unit Leaders	Postdocs and Staff Scientist	PhD Students	Technicians
2003	17	20	45	31
2004	20	21	54	37
2005	20	30	57	34
2006	24	59	73	48
2007	25	86	75	61
2008	33	100	80	83
2009	36	84	89	101



CRG Personnel

	Group and Unit Leaders
2003	17
2004	20
2005	20
2006	24
2007	25
2008	33
2009	36



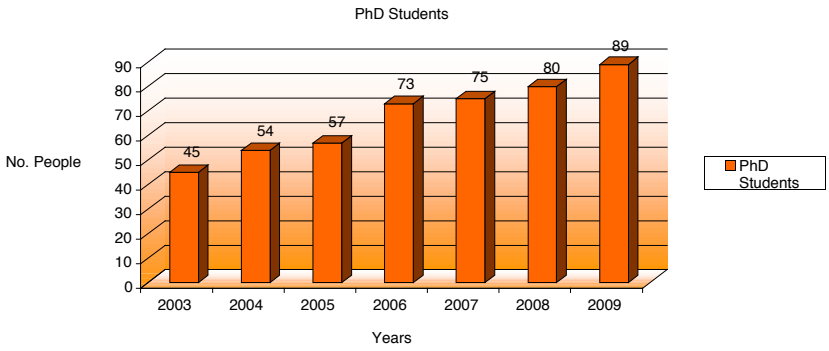


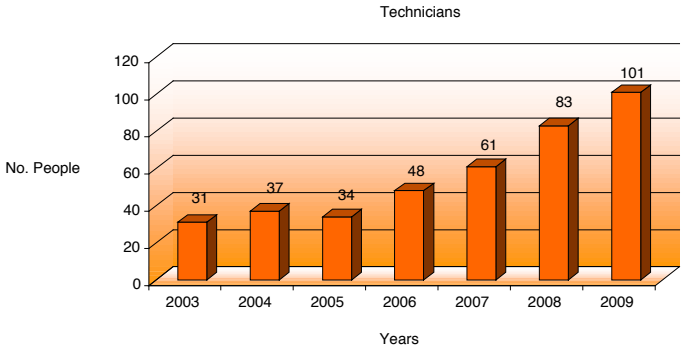
CRG Personnel

	Postdocs and Staff Scientist
2003	20
2004	21
2005	30
2006	59
2007	86
2008	100
2009	84

CRG Personnel

	PhD Students
2003	45
2004	54
2005	57
2006	73
2007	75
2008	80
2009	89



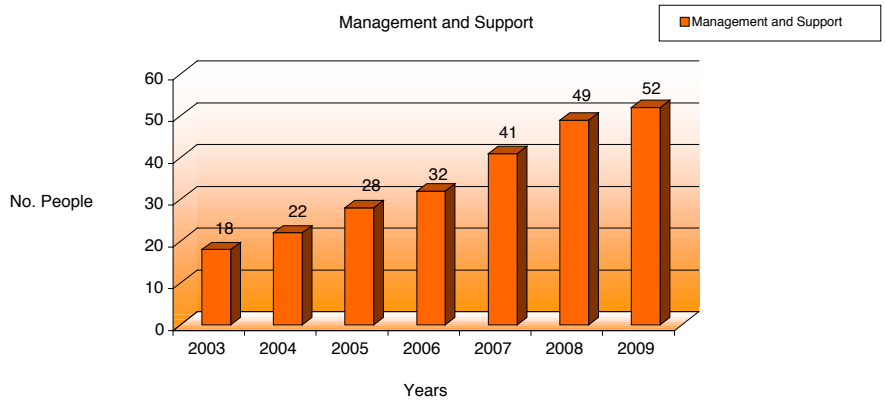


CRG Personnel

	Technicians
2003	31
2004	37
2005	34
2006	48
2007	61
2008	83
2009	101

CRG Personnel

	Management and Support
2003	18
2004	22
2005	28
2006	32
2007	41
2008	49
2009	52



Appendix 5

PRESS CLIPPING

Científicos españoles identifican dos genes implicados en la psoriasis

BARCELONA Redacción

Investigadores del Centro de Regulación Genómica de Barcelona (CRG) liderados por Xavier Estivill, coordinador del programa Genes y Enfermedad del CRG, han identificado un importante factor genético responsable de la psoriasis. Después de analizar varias regiones del genoma de 2.500 personas de España, Holanda, Italia y Estados Unidos, el equipo ha llegado a la conclusión de que las personas a las que le faltan las dos copias de los genes conocidos como LCE3B y LCE3C tienen un 20% más de riesgo de desarrollar la enfermedad. Estos dos genes desempeñan un papel importante en la formación de una epidermis adecuada. Los pacientes con psoriasis, sin embargo, carecen, con una alta frecuencia, de estos genes.

La psoriasis es un enfermedad inflamatoria y crónica de la piel que afecta a más de un millón de personas en España y que suele aparecer entre los 15 y los 40 años. Su origen está en la combinación de

Xavier Estivill en el CRG

LOS PIPES ESPAÑOLES

MARIANO BARBACID

MANEL BEATS

THOMAS GRAF

Los galardonados con los Premios Ciutat de Barcelona, anoche en el Auditori al finalizar el acto de entrega.

Barcelona entrega sus premios en el

EL PAÍS, Barcelona

El año 2008 pasará a la historia asociado al sustantivo *crisis*. Lo

en un Auditori cuyo escenario había sido transmutado en el aula de un colegio de la época republicana.

Ivorra, Jona, un joven detenido en 2006 tras participar en una manifestación contra la especulación urbanística y pendiente de juicio, para lanzar consign reivindicativas y desplegar u pancarta en el segundo anfite tro del Auditori durante la entr ga del premio correspondient arquitectura y urbanismo. En pancarta se podía leer "Prou peculación. Absolució Jona".

Tenden

Ciencia de primera

La investigación de Cataluña obtiene un creciente reconocimiento internacional

Más cienc

ESPAÑA padece un déficit secular en inv gación científica. Desde antes del provoc vo "que inventen ellos", que lanzó Una no a principios del siglo XX, la ciencia e novido endémicamente limitada al pl personal, muchas veces con rasgos de in actividades desarrolladas, en la sug en el exterior. Apenas ha habido pla tados por los gobiernos, las univers

el Periódico

20 minutos

EL PUNT BARCELONA

avenços en cèl·lules mare, ció celular i càncer al CRG

erts mundials en el camp de la diferenciació cel·lular presentaran, avui treballs en la vuitena edició del Simposi Anual del Centre de Regulació que se celebrarà a l'auditori del Parc de Recerca Biomèdica de Barcelo ntiation and Cancer" està organitzat per Thomas Graf, Luciano Di Croc es, investigadors del programa de Diferenciació i Càncer del CRG.

People & Ideas

Vivek Malhotra: Gaga for the Golgi

AVUI+

Your source for the latest research news

First-Ever Blueprint of 'Minima Complex Than Expected

ScienceDaily (Nov. 27, 2009) — What are the bare essentials of life, the indispensable ingredients required to produce a cell that can survive on its own? Can we describe the molecular anatomy of a cell, and understand how an entire organism functions as a system? These are just some of the questions that scientists in a partnership between the European Molecular Biology Laboratory (EMBL) in Heidelberg, Germany, and the Centre de Regulacio Genòmica (CRG) in Barcelona, Spain, set out to address.

In three papers published back-to-back in *Science*, they provide the first comprehensive picture of a minimal cell, based on an extensive quantitative study of the biology of the bacterium that causes atypical pneumonia, *Mycoplasma pneumoniae*. The study uncovers fascinating novelties relevant to bacterial biology and shows that even the simplest of cells is more complex than expected.

Mycoplasma pneumoniae is a small, single-cell bacterium that causes atypical pneumonia in humans. It is also one of the smallest prokaryotes -- organisms whose cells have no nucleus -- that don't depend on a host's cellular machinery to reproduce. This is why the six research groups which set out to characterize a minimal cell in a project headed by scientists Peer Bork, Anne-Claude Gavin and Luis Serrano chose *M. pneumoniae* as a model: it is complex enough to survive on its own, but small and, theoretically, simple enough to represent a minimal cell -- and to enable a global analysis.

A network of research groups at EMBL's Structural and Computational Biology Unit and CRG's EMBL-CRG Systems Biology Partnership Unit approached the bacterium at three

a ciencia catalana alcanza l rimera división de la genóm

EL PAIS

La vida es más compleja de lo que se

Investigadores europeos definen los requisitos de funcionami autosuficiente

EL PAÍS - Madrid - 26/11/2009

Un nuevo mapa de todas las proteínas en una de las bacterias más pequeñas. Mye los científicos a determinar la mínima maquinaria celular requerida para la vida. esta bacteria es sorprendentemente compleja.

Luis SERRAN

CON EL 66% DE LA PRODUCCIÓN CIENTÍFICA DEL ESTADO

Catalunya impulsa la investigación biomédica

La Generalitat dota de infraestructuras punteras para que el ámbito de la biomedicina investigue con vistas a una mejora de la atención sanitaria

Migration feeds culture of growth

The cluster is making progress despite having powerful rivals, says Victor Mallet


LA SEGUNDA
DIRECTOR
Telefónica
sobre el Forum
Vacas en los pastos de una explotació
Descifrado el genom
para buscar mejoras
14 expertos mundiales sobre células madre y cáncer participan el VIII simposio del CRG
Un total de 14 expertos mundiales en el campo de las células madre y el cáncer participarán entre mañana y el viernes en el VIII simposio anual del Centro de Regulación Genómica (CRG) de Barcelona, una cita qu
Tendencias
Más allá del genoma
Los últimos avances allanan la vía hacia la creación de organismos sintéticos
'Hat-trick' científico
ia y de calidad
científica en Catalunya en los últimos diez años, que ha duplicado el número de trabajos en una década. Un período durante el cual han aumentado en 13 puntos porcentuales los trabajos firmados en Catalunya en co
science & society
Science
The New York Times
LA VANGUARDIA


La Voz de Galicia
.es
adn.es
Cell
Down de Madrid otorgará el VI
Jaime Blanco' a Mara Dierssen
id (FSDM), en colaboración con la Dirección Genera
Discapacidad del Ministerio de Educación, Política y
ación en síndrome de Down 'Jaime Blanco' al trabaj
Els secrets de la vaca
rios | +0 -0 (0 votos)
DIARIO MEDICO.COM
CRITICA LA COMERCIALIZACIÓN
Roderic Guigó: "Secuenciar el genoma individual no es útil"
Secuenciar el genoma humano parecía el inicio de una revolución, pero ocho años después sólo sirve para que una empresa comercialice un servicio de nula utilidad, según ha admitido el bioinformático Roderic Guigó.
Carmen Fernández. Barcelona - Martes, 26 de Julio de 2006 - Actualizado a las 00:00h
Roser Pinyol: «A menudo, la sociedad ve a los como cajas negras»
"LA CUESTIÓN ES SER CAPACES DE TRAER CEREBROS, DE RUSIA, UGANDA O EE UU"

CRG
Centre
for Genomic
Regulation

Centre for Genomic Regulation

PRBB Building
Dr. Aiguader, 88
08003 Barcelona, Spain

Tel.: +34 93 316 01 00
Fax: +34 93 316 00 99

comunicacio@crg.es
http://www.crg.es

Members of the Board of Trustees:

Generalitat de Catalunya

GOBIERNO DE ESPAÑA
MINISTERIO DE CIENCIA E INNOVACIÓN

Institut adscrit a la
UNIVERSITAT POMPEU FABRA