CRG SCIENTIFIC STRUCTURE ................................................................. 04
CRG MANAGEMENT STRUCTURE ......................................................... 06
CRG SCIENTIFIC ADVISORY BOARD (SAB) ........................................ 08
CRG BUSINESS BOARD ......................................................................... 09
YEAR RETROSPECT BY THE DIRECTOR OF THE CRG: MIGUEL BEATO .......... 10

GENE REGULATION ........................................................................... 14
- Chromatin and gene expression ...................................................... 16
  - Transcriptional regulation and chromatin remodelling ................. 18
- Regulation of alternative pre-mRNA splicing during cell differentiation, development and disease ........ 21
- RNA interference and chromatin regulation .................................... 24
- RNA-protein interactions and regulation ........................................ 27
- Regulation of protein synthesis in eukaryotes ............................... 30
- Translational control of gene expression ........................................ 33

DIFFERENTIATION AND CANCER .................................................... 36
- Hematopoietic differentiation and stem cell biology ...................... 38
- Myogenesis ..................................................................................... 43
- Epigenetics events in cancer ......................................................... 46
- Epithelial homeostasis and cancer ................................................. 49

GENES AND DISEASE ..................................................................... 52
- Genetic Causes of Disease ............................................................ 54
- Gene Therapy ............................................................................... 60
- Murine Models of Disease ............................................................. 63
- Neurobehavioral Phenotyping of Mouse Models of Disease ........ 65
- Gene Function .............................................................................. 68
- Associated Core Facility: Genotyping Unit ..................................... 71
BIOINFORMATICS AND GENOMICS ........................................ 74
  ▪ Bioinformatics and genomics ........................................ 76
  ▪ Comparative bioinformatics ........................................ 81
  ▪ Genomic analysis of development and disease ..................... 84
  ▪ Associated Core Facility: Microarrays Unit ....................... 86

CELL AND DEVELOPMENTAL BIOLOGY .................................... 88
  ▪ Intracellular compartmentation .................................... 90
  ▪ Microtubule function and cell division ............................ 94
  ▪ Sensory cell biology and organogenesis ........................... 97
  ▪ Associated Core Facility: Advanced Light Microscopy Unit ...... 100

SYSTEMS BIOLOGY ......................................................... 104
  ▪ Design of biological systems ...................................... 106
  ▪ Systems analysis of development ................................. 110
  ▪ Gene Network Engineering ......................................... 114
  ▪ Metazoan Systems ................................................... 117

APPENDIX 1: ............................................................... 122
  ▪ VI CRG Annual Symposium:  
    "Genomic Regulation: Executing the Code"

APPENDIX 2: ............................................................... 124
  ▪ CRG Seminars and Programme Seminars

APPENDIX 3: ............................................................... 134
  ▪ Grants

APPENDIX 4: ............................................................... 135
  ▪ Finance and Personnel Evolution
SCIENTIFIC STRUCTURE
MANAGEMENT

STRUCTURE
ADVISORY BOARD (SAB)

DR. KAI SIMONS
Max Planck Institute of Molecular Cell Biology and Genetics, Dresden

DR. STYLIANOS EMMANUEL ANTONARAKIS
Medical Genetics, University of Geneva, Geneva

DR. MICHAEL ASHBURNER
European Bioinformatics Institute-EMBL Outstation, Department of Genetics, University of Cambridge, Cambridge

DR. PIERRE CHAMBON
Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Strasbourg

DR. IAIN MATTAJ
European Molecular Biology Laboratory (EMBL), Heidelberg

DR. JOAN MODOLELL
Center of Molecular Biology “Severo Ochoa”, CSIC & Autonomous University of Madrid, Madrid

DR. ARNOLD MUNNICH
Det. Génétique, Hôpital des Enfants Malades, Paris

DR. CHRISTIANE NÜSSELIN-VOLHARD
Abt. Genetik, Max Planck Institut für Entwicklungsbiologie, Tübingen

DR. MARC VIDAL
Dana Farber Cancer Institute, Boston, USA

DR. ERWIN WAGNER
Research Institute of Molecular Pathology (IMP), Vienna
DR. ANTONI ESTEVE
Scientific-Commercial Area Director, Esteve

DR. JOSEP PROUS, JR.
Jr., Executive Vice President, Prous Science

SR. JESÚS ACEBILLO
President, Novartis Farmacéutica, S.A.

SR. PERE BERGA
R+D Management Director, Almirall Prodesfarma, S.A.

SR. JOSEP M. TABOADA
Medical Director, Sanofi-Aventis, S.A.

SR. RAFAEL PARDO AVELLANEDA
General Manager, Fundación BBVA

SR. JORDI RAMENTOL
General Manager, Grupo Ferrer Internacional, S.A.
YEAR RETROSPECT
by the Director of the CRG: Miguel Beato

After moving to the new building of the PRBB which was nearly completed by the end of last year, 2007 has been a year of settling down in the new premises. The first months were dominated by the need to solve all kind of unanticipated technical problems related to the general services of the new building. It was only at the end of spring, after the Genotyping Facility and the Microarray Units were also installed in the PRBB building, that we started to enjoy the advantages offered by new spaces and equipment.

From the institutional point of view two important agreements were negotiated. One with the Foundation Marcelino Botín, which will finance the technology transfer of two senior groups for an initial period of 3 to 5 years. The other with la Caixa, which will sponsor a total 40 competitive fellowship/contracts, 10 each year, for the international PhD Programme of the CRG on “Genomic Regulation and Disease”. These are the two most significant agreements the CRG has reached with private foundations and I hope they represent a point of inflexion in respect to the participation of private capital in financing the functioning of the CRG.

One of the most relevant scientific events of the year was the VI CRG Symposium on “Genomic Regulation: Executing the Code”, which included excellent presentations by leaders in their fields. Genomic regulation has experienced a rapid expansion in the recent past that lead to a plethora of results difficult to assimilate by non-experts. The CRG symposium tried to cover all relevant aspects of gene regulation, from nuclear dynamics to global regulation, and non-coding RNAs. It offered the 250-300 participants an integrated vision on genomic regulation and an insight into future developments. Before the main symposium, we organized a satellite minisymposium on “Kinases Signaling to Chromatin”, a very timely and debated topic.
Another important event with participation of the CRG was “The 9th International Meeting on Human Genome Variation and Complex Genome Analysis”, organized by Xavier Estiviil, that took place in Sitges and was attended by 150 participants from all over the world.

Finally, in December the graduate students of the CRG organized the “First Young Researchers Symposium” in the PRBB auditorium. All young scientists of the CRG were invited to participate with posters, presentations or roundtable discussions. The meeting was a great success and we intend to expand the future editions to all young investigators of the PRBB.

The CRG is also increasing its activities in the fields of Science and Society and divulgation. In March Mara Dierssen of the CRG was co-coordinating the “Brain Awareness Week 2007” with a plethora of activities culminating with a Neuroparty in a large disco. In May the CRG hosted the closing act of “The one million question” project, which was attended by the winning high school class. Again Mara Dierssen designed and organized in October a scientific feast on “The illusions of the brain” with scientific and sensorial experiments and concerts. The CRG has signed a collaboration agreement with “Kutxaespacio de la Ciencia” to foster divulgation activities in the area of genomics. Till now two events took place, a participation in the II Feria de la Ciencia / Experimentalia 07 and a scientific workshop along with the Catalan Biological Society.

Among the distinctions to CRG scientists, the most significant was the selection of three young group leaders, Ben Lehner, Hernan López-Schier and Mark Isalan, for the Starting Grants of the European Research Council, and the election of Luis Serrano as member of the Spanish Royal Academy of Exact Sciences.

Without doubt the most significant experience of last year was the first external evaluation carried out in October. A panel of 12 experts formed by 7 members of the Scientific Advisory Board and 5 ad hoc reviewers evaluated during 2 days the first 5 years of work of 12 groups belonging to the oldest three programmes: Gene Regulation, Genes & Disease and
Cell Differentiation & Cancer. The three programme coordinators and the director were also evaluated. The outcome was very positive and will help in ensuring mobility, scientific excellence and the establishment of modern core facilities.

In terms of research activities all programmes have continued to consolidate their groups and facilities. In September Cedric Notredame from the CNRS in Marseille joined the Bioinformatics & Genomics programme as a senior scientist. The Light Microscopy Unit encompasses already four confocal microscopes and the Genomics Unit will incorporate two platforms for ultra-high-sequencing, the Illumina/Solexa Genome Analyzer 1G and the Roche FLX System.

A novel research activity initiated in 2007 was the transversal institutional project on "Cell Reprogramming". The project has been approved by the Scientific Advisory Board and integrates all six programmes of the CRG. The objective is to study in detail the molecular and cellular changes that accompany cell reprogramming and to use this information for a large scale screening of conditions that change the fate of adult cells. The project makes use of all core facilities and has already served to plan and initiate the implantation of a robotic facility for high throughput screening. Apart from its great scientific interest, the project is expected to yield results applicable to cell therapy and regenerative medicine.

Five years after its official inauguration, the CRG is approaching its final size. By the end of the year 288 people were working at the CRG: 25 group or unit leaders, 7 staff scientists, 79 postdocs, 75 graduate students, 61 technicians and 41 administration and support people. Some effort is still needed to complete the programmes and the core facilities but the goal to make the CRG one of the players in the champions league of biomedical research centers seems now reachable.
In 2007 the programme has consolidated its six groups and has continued the planning of the proteomic facility in collaboration with the Department of Experimental and Health Sciences (CEXS) of the UPF. There is a basic agreement for creating a common facility that will incorporate the already existing equipment and personnel from the CEXS/UPF and new equipment for quantitative proteomics and personnel hired by the CRG. The facility will be implanted in new spaces on the ground floor of the PRBB along with the Genomic and Bioinformatics facilities. After securing the resources for new equipment, including a LTQ Orbitrap machine, an ad was placed in *Nature* for the head of the facility and very good candidates have applied. Selection of candidates is planned for January 2008 and the new facility should be operative after next summer.

One important event in the fall was the first evaluation of the programme by members of the SAB and *ad hoc* reviewers. The evaluation was very positive, though it included recommendation for increased mobility to allow the incorporation of new groups. To this end the group of the Miguel Beato will leave the programme to form an independent group of the director and a new junior position will be announced early 2008.

Another important event was the organization of the VI CRG Annual Symposium on “*Genomic Regulation: Executing the Code*” by the three senior scientists of the programme. The symposium included a satellite Mini-Symposium on “*Kinases Signalling to Chromatin*” organized by Francesc Posas (UPF) and Miguel Beato, and was a great success because of the quality of the speakers and the number of active participants.

The structure of the programme during 2007 was:

- **6 Research groups:**
  - Chromatin and Gene Expression (Miguel Beato/Albert Jordán)
  - RNA Interference and Chromatin Regulation (Ramin Shiekhattar)
  - Regulation of Alternative Pre-mRNA Splicing during Cell Differentiation, Development and Disease (Juan Valcárcel)
  - RNA-Protein Interactions and Regulation (Josep Vilardeil)
  - Translational Control of Gene Expression (Raúl Méndez)
  - Regulation of Protein Synthesis in Eukaryotes (Fátima Gebauer)

- **Associated Core Facility:**
  - Proteomic Core Facility (under construction)
The group is interested in understanding how eukaryotic cells respond to external signals, in particular how different signals are integrated and transduced to the nucleus to modulate gene expression. The main experimental model is gene regulation by steroid hormones in breast and endometrial cancer cells. More specifically, attention is focused on the crosstalk of estrogen and progesterone receptors with other signalling pathways originating in the cell membrane and how this network of signalling is interpreted at the level of chromatin. The role of steroid hormones in breast and endometrial cancer cell proliferation and differentiation is another research line of the group.
RESEARCH PROJECTS


Progesterone controls proliferation and gene expression in breast cancer cells, where it activates transiently the Src/Ras/Erk pathways (Migliaccio et al EMBO J 17, 2008-18,1998) via an interaction of two domains of the progesterone receptor (PR) with the estrogen receptor alpha (ER?). Kinase activation is essential for the proliferative response of breast cancer cell lines (Ballaré et al Mol Cell Biol 23, 1994-2008, 2003) as well as for progestin induction of MMTV and other target genes (Vicent et al Mol Cell 27, 367-81,2006). In 2007 we started studying the mechanism by which Msk1, a target of Erk1/2, controls cell proliferation in response to hormones in order to develop selective strategies to interfere with the proliferative effect of estrogens and progesterone.

In collaboration with Belen Miñana, Juanjo Lozano and Lauro Sumoy from the Microarray Unit we are performing gene profiling and tiling studies in breast cancer cell lines to study the chromatin structure and 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 combined with transcriptome analysis in response to hormones (Figure 1). Using proteomics in breast cancer cells carrying tagged PR we have initiated the analysis of the PR interactome with the aim to identify the different complexes involved in hormonal regulation via various signalling pathways.

In collaboration with Christophoros Nikolau and Roderic Guigo from the Bioinformatics and Genomic programme (B&G) we have develop an algorithm for predicting nucleosome positioning in the human genome and are currently validating the predictions using tiling microarray and massive nucleosome sequencing.

We are analyzing a selection of 40 hormone responsive promoters by high resolution ChIP-on-chip to define their nucleosomal structure before and after hormone induction, and have identified numerous 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.

Using in vivo and in vitro approaches, including atomic force microscopy (AFM), we are investigating the role of histone H1 variants and their modifications in the regulation of chromatin spacing and chromatin compaction as well as in the control of transcription.

2. Regulation of MMTV transcription in the chromatin context


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 (Snf2h and Brg-1), co-regulators (CBP, PCAF, Src1) and RNA polymerase II (Vicent et al Mol Cell 27, 367-81,2006) (Figure 2). Shortly thereafter we detect the displacement of histones H2A and H2B from the promoter nucleosome containing the HREs but not from the adjacent nucleosomes (Vicent et al Mol Cell 16, 439-52, 2004).

In 2007 we have demonstrated a key role of the BAF chromatin remodelling complex in hormonal activation of MMTV. Upon hormone addition, BAF is recruited to the promoter by PR, and depleting the complex leads to
inhibition of hormonal induction. BAF mediated remodelling enables NF1 access that facilitates PR binding to the central HREs on the remodelled nucleosome. The new PR molecules bring additional BAF complexes to the promoter and lead to full activation.

We found that histone H1 enhances the activation of the MMTV promoter by PR and NF1 (Koop et al. EMBO J 22, 588-99, 2003) and are now studying the role of various H1 isoforms and their phosphorylation by Cdk2 on the remodelling and transcription of MMTV chromatin.

3. Role of steroid hormones in breast cancer, endometrial physiology and angiogenesis
C. Ballaré, V. Calvo, M. Edel

In collaboration with the Department of Pathology of the Hospital del Mar and the microarray unit of the CRG, we are studying 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 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 1 above) should permit to classify the tumours according to the perturbed signalling pathways. A correlation of this molecular description with clinical and histological data will allow 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 (Vallejo et al. Mol Endocrin 18, 3023-37, 2005). In 2007 we are studying how blood serum and the combination of estrogens and progestins induce the decidual reaction of these cells, a physiological example of trans-differentiation.

In a screen for genes that bypass a p53-mediated cell cycle arrest, we identified the small GTPase Rem2 as a suppressor of p19ARF expression that promotes endothelial cell proliferation and angiogenesis and is under steroid hormone control.

Research Subgroup:

TRANSCRIPTIONAL REGULATION AND CHROMATIN REMODELLING

1. Transcriptional regulation and chromatin remodelling of hormone responsive promoters.
A. Jordan, I. Quiles, Ll. Millán, A. Subtil

We are interested in distinguishing between direct effects of nuclear hormone receptors on transcription of target genes and those mediated by crosstalk with other signal transduction pathways. For this, we have constructed breast cancer-derived cell lines that express tagged forms of PR mutated at residues involved either in the nuclear action of the receptor (DBD and AF-2) or in its ability to interact with components of signal transduction pathways (∆ERID-I). Microarray experiments have been performed to define the subsets of genes affected on its response to hormone by the different PR defects. A majority of genes require an intact PR, able to interact with DNA and to activate kinase cascades, pinpointing the importance of cross-talk between PR modes of action.

We are also performing the characterization of the progesterone-responsive 11ß-HSD type 2 human promoter. Two distinct regions of this promoter recruit PR upon hormone addition: a proximal region where PR interacts directly with DNA, and a distal region containing STAT5A binding sites. This last region is essential for the hormone responsiveness of the promoter. Hormone treatment activates the JAK/STAT pathway and STAT5A transcription factor recruits PR, co-activators and the RNA polymerase to the distal region. Along with promoter activation, non-coding RNAs are synthesized from upstream, which nature and functionality is being investigated. In addition, we are performing chromatin immunoprecipitation (ChIP) to study histone modifications, as well as the composition of associated chromatin remodelling complexes and transcriptional complexes.
2. Role of linker histone H1 variants in chromatin and transcription.
A. Jordan, E. Diani, M. Sancho

At least six H1 variants exist in mammalian somatic cells that bind to the nucleosome core particles and linker DNA. We have used inducible RNA interference to create stable breast cancer cell lines lacking expression of each of the H1 variants. The resulting phenotype is being characterized, as the inhibition of some of the isoforms produces a proliferation defect. We are now investigating the role of each variant on global gene expression by using microarrays. In parallel, we have developed specific antibodies for H1 isoforms 1 to 5, and generated cell lines expressing HA-tagged H1 isoforms that will be used on ChIP-on-chip experiments devoted to determine the presence of each form in several promoters of interest.

3. Influence of chromatin at the integration site on the transcriptional activity of the HIV promoter.
A. Jordan, E. Gallastegui

HIV integrates at a multitude of sites without any clear preference in the human genome. The chromatin environment at the integration site influences the nucleosome structure of the viral promoter and consequently its basal and Tat-induced transcriptional activity (Jordan et al EMBO J 20, 1726, 2001), in a way that is independent of the degree of methylation of the proviral DNA (Pion et al J Virol 77, 4025, 2003). In this respect, we have shown that at low frequency integration occurs at regions of heterochromatin (i.e. pericentromeric) leading to promoter repression and to a state of viral latency that can be reactivated upon T cell activation (Jordan et al, EMBO J 22, 1868, 2003). We plan to compare the chromatin structure, histone code and protein recruitment to the proviral promoter when integrated in transcriptional-competent euchromatin or in repressed heterochromatin. In addition, by using RNA interference we are investigating the hypothetical participation of several chromatin components or histone modifying enzymes on the establishment and maintenance of the repressive promoter state, which could be candidates for the therapeutic intervention against the latent state.

PUBLICATIONS


Bierings R, Beato M, Edel MJ. An endothelial cell genetic screen identifies the GTPase Rem2 as a suppressor of p19ARF expression that promotes endothelial cell proliferation and angiogenesis. J Biol Chem, [Epub ahead of print]


BOOK CHAPTERS


Figure 1. Crosstalk between PR and ER in the inner side of the cell membrane. In addition to the direct activation of PR followed by homodimerization and binding to HREs in chromatin, progesterone can induce the activation of a preformed ER-PR complex attached to the cell membrane via paloitylation, leading to interaction of ER with c-Src and activation of the mitogenic Ras > Raf > Mek > Erk1/2 cascade. In the nucleus Erk1/2 can phosphorylate transcription factors, cell cycle proteins and chromatin proteins, directly or via activation of downstream kinases. One of those, Msk1, phosphorylates histone H3 at S10 and initiates activation of some target promoters. PR can also activate c-Src directly, and c-Src can activate the PI3K/Akt pathway and the Jak/Stat5 pathway leading to regulation of other target genes. Blocking the different signalling pathways with the indicated inhibitors followed by expression profiling should identify the pathways used by hormones to regulate different gene networks.

Figure 2. Steps in the activation of the MMTV promoter. Progesterone interacts with nuclear PR and also with a cytoplasmic PR/ER complex, which activates Erk1/2 via interaction of ER with c-Src. Activated pErk phosphorylates PR and Msk1, forming a ternary complex pPR/pErk/pMsk1 that binds to the exposed HRE1 on the MMTV nucleosome and phosphorylates H3 at S10. This leads to displacement of a repressive complex containing HP1g, as a prerequisite for the recruitment of ATP-dependent remodelling complexes and the ejection of H2A/H2B dimers. This remodelling enables binding of NF1 that stabilizes the open conformation and facilitates binding of further PR molecules to the internal HREs, and further recruitment of SWI/SNF-like remodelling complex. Subsequently, PR recruits transcriptional co-activators and the transcription initiation complex.
Alternative pathways of RNA processing lead to the synthesis of different mRNAs from a single gene, thus expanding the coding potential of the genome. Our group investigates how different cell types, or the same cells under different conditions, decide which mRNA to produce from a single mRNA precursor. This involves selective removal of some sequences—known as introns—and splicing together of the remaining sequences—exons—to generate mature mRNAs. During 2007 we have made progress in understanding the biological impact of alternative splicing in a variety of normal and pathological processes, from sex determination in Drosophila to muscular dystrophies or tumor progression. We have also identified regulatory factors that control their activities through gene networks and learned about the diversity of mechanisms used to control alternative splicing of genes important for cell proliferation and programmed cell death.
RESEARCH PROJECTS

1. Biological impact of alternative splicing

_Drosophila_ sex determination offers a textbook example for how gene expression can be regulated posttranscriptionally. Two key splicing regulators, expressed in a sex-specific manner, induce sex-specific changes in splicing of a handful of target genes which, in turn, trigger transcriptional programs that orchestrate all aspects of sexual dimorphism and behavior. Using splicing-sensitive microarrays (collaboration with Don Rio, Univ. California at Berkeley), we have identified hundreds of additional sex-specific changes in alternative splicing which can deeply influence the function of the corresponding gene products and contribute to sex-specific phenotypes.

Alternative splicing changes can also occur under pathological situations. We have identified such changes in samples from patients suffering from myotonic dystrophy. These changes can affect programs of gene regulation as well as the contractile properties of the muscle, potentially contributing to the disease pathology. Changes associated with tumor progression have been detected in regulatory regions of mRNAs with important functions in cell cycle control.

2. Splice site recognition

The proteins TIA-1 and TIAR regulate 5' splice site recognition by binding to uridine-rich sequences downstream from weak splice sites and facilitating recruitment of U1 snRNP. We have identified a genetic network that controls the activities of these factors. Previous work had shown that TIA-1 and TIAR can regulate their own splicing to control the levels of their protein-coding transcripts (Le Guiner et al, _JBC_ 276, 40638). We have now found that TIAR regulates alternative splicing of TIA-1 to generate protein isoforms with distinct regulatory activities. This circuit can serve to buffer variations in the levels of these proteins (Figure 1).

Additional sequences and factors that regulate 5' splice site recognition have been identified through a combination of bioinformatic and experimental approaches (collaboration with Robert castelo and Roderic Guigó). The picture emerging from these studies is that a variety of auxiliary sequences contribute to splice site recognition, offering possibilities for regulation of alternative splicing.

3. Alternative splicing of the Fas receptor

The exon 6 of Fas pre-mRNA can be alternatively spliced to generate mRNAs encoding proteins with antagonistic functions in the control of programmed cell death. We have found that activation of Fas signalling can influence alternative splicing of the receptor by inducing phosphorylation of TIA-1, which in turn promotes inclusion of Fas exon 6 and accumulation of the pro-apoptotic form of the receptor.

RBM5, the product of a putative tumor suppressor gene, and SPF45, a protein frequently overexpressed in tumors -to which confers resistance to antitumor drugs- antagonize the effects of TIA-1. We have learned that RBM5 regulates Fas splicing by blocking late steps in spliceosome assembly and acting as a selector of splice site pairing. In contrast, we have found, in collaboration with the group of Michael Sattler (Technical Univ., Munich) that multiple interactions of the UHM domain of SPF45 with factors involved in early steps in 3' splice site recognition is necessary for promoting exon 6 skipping (Figure 2). These findings illustrate how regulatory factors exploit the diversity of molecular interactions underlying the splicing process to influence the choice between alternative splice sites.

PUBLICATIONS


Figure 1. A posttranscriptional gene network that controls the activity of TIA-1 / TIAR proteins. Alternative splicing of TIA-1 exon 5 leads to the production of two protein isoforms, TIA-1a and TIA-1b, of which TIA-1b is more active as a splicing regulator. The TIA-1-related protein TIAR regulates this alternative splicing event by promoting exon 5 inclusion. Decreased levels of TIAR result in exon 5 skipping and therefore increased levels of the more active isoform of TIA-1. TIAR itself is regulated by TIA-1 (Le Guiner et al, JBC 276, 40638), thus establishing a regulatory gene network that modulates TIA-1/TIAR function.

Figure 2. Mechanism of Fas alternative splicing regulation by the splicing regulator SPF45. Initial 3' splice site recognition is achieved through cooperative interactions between SF1/BBP and the two subunits of U2AF, which recognize different sequence elements in the 3' end of the Intron (A). This is followed by recruitment of U2 snRNP, which is stabilized through base pairing interactions between U2 snRNA and the intron as well as through interaction between SAP155 and U2AF65 (B). SPF45 modulates alternative splicing and this function requires binding to SF1 and SAP155, which may disrupt the interactions leading to 3' splice site definition described in A and B. Both the U2AF65-SF1, U2AF65-SAP155 and SPF45 interactions with SF1 and SAP155 are mediated by U2AF Homology Motifs (UHM) present in both U2AF65 and SPF45.
GENE REGULATION
RNA Interference and Chromatin Regulation

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.

GROUP STRUCTURE

Group Leader: Ramin Shiekhattar
Staff Scientist: Glenda Harris
Postdoctoral Researchers: Iidem Akerman, David Baillat, Malte Beringer, Clement Carre, Klaus Fortschegger, Fan Lai, Frederic Tort, Laure Weill
Technician: Albane Lamy de la Chapelle, Leonor Avila
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 remodeling 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.

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.

PUBLICATIONS

Smith ER, et al. Drosophila UTX is a histone H3 Lys27 demethylase that colocalizes with the elongating form of RNA polymerase II. Mol Cel Biol, Epub 2007 Nov 26


(*) All these publications are the result of the work of Dr. Ramin Shiehattar at the Wistar Institute, Philadelphia, USA.
GENE REGULATION
RNA-Protein Interactions and Regulation

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 further characterize the extent in which splicing is regulated in *Saccharomyces*.

*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.

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
   M. Bragulat, Gaja Lesnicar

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 behave in the opposite way, 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 indicates that the observed effect of Cbp80 could be unrelated to its role on U1 snRNP binding, providing data on a new function of this important factor.

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.
   S. Macias

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.

3. Proteomic and Structural analyses of the Inhibited Complex
   John LaCava

We have the possibility to explore, analytically and structurally, a regulated splicing complex. We can purify the L30-stalled pre-spliceosome, or inhibited complex, in analytical amounts, either from in vitro approaches or from specially designed strains, in vivo. Our aim is to identify all the components present in this complex. Their interactions will be further analyzed by cross-linking experiments; and finally, if technically feasible, by cryo-EM, in collaboration with other leader groups in the field (Reinhard Luhrmann).

Figure 1. Regulation of RPL30 splicing by L30. (A) Secondary structure of the RNA element required for L30 binding. Nucleotides 17-50 are not involved. Mutation C9 to U (blue) abolishes regulation, probably by weakening L30 binding. The 5' splice site and nucleotides complementary to U1 snRNA are indicated. (B) L30 blocks spliceosome assembly at an early step, generating a new stable complex, the “inhibited” complex.
4. Other *S. cerevisiae* genes with regulated splicing

*Markus Meyer*

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 5' 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).

**PUBLICATIONS**

The control of mRNA translation is an important mode of regulation used in a wide variety of biological situations. We are interested in regulation by RNA-binding proteins and microRNAs. We study these mechanisms under three different biological contexts: X-chromosome dosage compensation, early embryonic patterning and cell differentiation.
RESEARCH PROJECTS

1. Translational control of dosage compensation

X-chromosome dosage compensation in Drosophila is achieved by the binding of the dosage compensation complex (DCC) to hundreds of sites on the male X-chromosome, which promotes a ~2 fold hypertranscription. This process is inhibited in female flies because the critical DCC subunit MSL2 is not expressed. The female-specific RNA-binding protein Sex-lethal (SXL) binds to uridine stretches present in the 5' and 3' UTRs of msl-2 mRNA and inhibits its translation. Translational repression requires an additional factor that is recruited by SXL to the 3' UTR of msl-2. We have identified this factor as the Drosophila homolog of mammalian Upstream of N-ras (UNR). UNR is an ubiquitous cytoplasmic protein present in both male and female flies (Figure 1). Analysis of the domains of UNR involved in msl-2 regulation suggests that UNR contacts additional factors for translational repression, which we are currently trying to identify (Figure 2). Our results indicate that SXL provides a sex-specific function to UNR, and suggest that UNR is a novel regulator of dosage compensation in Drosophila. We are testing the role of UNR in fly development and dosage compensation, as well as identifying other mRNA targets of UNR.

2. Translational regulation of early embryonic patterning

The antero-posterior and dorso-ventral patterning system in Drosophila heavily depends on translational control. The 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 mostly unknown. We have used a cell-free cytoplasmic polyadenylation/translation system to study the translational regulation of toll mRNA and have identified novel cytoplasmic polyadenylation elements. We are currently trying to isolate the machinery binding to these elements by RNA-affinity chromatography.
3. Regulation of p27<sup>kip1</sup> mRNA translation

p27<sup>kip1</sup> is a cyclin-dependent kinase (cdk) inhibitor that blocks the mammalian cell cycle in G1. Proper modulation of p27<sup>kip1</sup> levels is essential for cell proliferation and differentiation. One of the mechanisms that dictate the levels of p27<sup>kip1</sup> is the translational regulation of its mRNA. We have identified regulatory sequences for translation in p27<sup>kip1</sup> mRNA by using cell transfection approaches, and have identified microRNAs that recognize these regulatory sequences. We are testing the relevance of these miRNAs in cell cycle progression and differentiation.

PUBLICATIONS

Abaza I and Gebauer F. Functional domains of Drosophila UNR in translational control. RNA, in press.

Abaza I and Gebauer F. Trading translation with RNA binding proteins. RNA, in press.

BOOK CHAPTERS


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. Determination of the 3'-UTR features that define the timing of cytoplasmic polyadenylation and the silencing of an mRNA.

The analysis of the polyadenylation state of the endogenous cyclin mRNAs during meiosis, the capability of numerous 3'-UTR variants to direct translational repression and subsequent cytoplasmic polyadenylation and translational activation, as well as the analysis of the trans-acting factors assembled on specific cis-acting elements, allowed us to define a combinatorial code that can be used to predict the translational behavior of CPE-containing mRNAs (Figure 1, Piqué et al.). We have translated these rules into regular expressions, to define qualitative and quantitatively whether a given mRNA could be a target for cytoplasmatic polyadenylation control, and performed a computational analysis, identifying hundreds of mRNAs potentially regulated by CPEB, mainly related to cell cycle and cell differentiation, but also to other biological events such as chromosome segregation, synaptic stimulation, embryonic polarity or even implicated in angiogenesis and tumor development.

We are performing a proteomic analysis of the trans-acting factors recruited by the different arrangements of cis-acting elements, defined in our combinatorial code for CPE-mediated translational control, to determine the molecular mechanism that define the phase-specific translational regulation of the different mRNA subpopulations.

2. Sequential waves of polyadenylation and deadenylation drive meiotic progression.

We have found an additional mechanism, overlaying the above-described rules, controlling the phase-specific deadenylation of a subset of CPE-regulated mRNAs (Figure 1, Belloc and Méndez). Our results show that an “early” wave of cytoplasmic polyadenylation activates a negative feedback loop, which opposes CPEB activity on mRNAs containing both CPEs and AREs, by activating the synthesis of C3H-4, which in turn recruits the CCR4-deadenylase complex to the ARE-containing mRNAs. These sequential waves of polyadenylation and deadenylation define a circuit of mRNA specific translational regulation that drives meiotic progression. We are following our research on the biological function of C3H-4 in two main directions. First, determining the regulation of C3H-4 by post-translational modifications. Second, performing a genome-wide screening for C3H-4 and CPEB-regulated mRNAs and define their meiotic functions.

3. Cytoplasmic polyadenylation role in the mitotic spindle formation and chromosome segregation during cell division

To determine the relevance of localized CPE-mediated translational control during meiotic progression, we have shown that spindle-localized translational activation, by cytoplasmic polyadenylation, is essential to complete the first meiotic division and also for chromosome segregation in Xenopus oocytes (Figure 2, Eliscovich et al.). We are performing functional screenings to identify in a comprehensive manner the mRNAs locally translated by CPEB during spindle formation/maintenance in oocytes and somatic cells.

4. Meiotic translational control by other members of the CPEB family of proteins:

Other than the above mentioned CPEB1, the CPEB family of proteins is composed of other three members. We are working on the mechanisms of translational regulation by other members of the CPEB family of proteins, identifying their target mRNAs and their biological functions in cell cycle control.

5. Translational control of mitotic cell cycle and cancer:

Using the knowledge acquired in Xenopus oocytes, we are trying to determine whether cytoplasmic polyadenylation also regulates cell cycle progression in somatic cells. We are also studying its regulation in normal and tumoral tissues and the physiological relevance of CPEB(s) overexpression in tumor development.
PUBLICATIONS


DIFFERENTIATION AND CANCER
Coordinator: Thomas Graf

Research within the programme concentrates on adult stem cell biology, mechanisms of cell fate instruction and cancer. More specifically, it covers the areas of:

- Transcription factors and signaling pathways in the development, regeneration and function of muscle cells (Pura Muñoz-Canoves)
- Epigenetic events in PML-RAR induced leukemia (Luciano Di Croce)
- Hematopoietic cell differentiation and reprogramming (Thomas Graf)
- Epithelial stem cells in the skin and cancer (Salvador Aznar-Benítez)
- Associated services: Histopathology and FACS Units

All four groups of the programme work with mammalian cell cultures as well as with mice. They share their expertise in epigenetic analyses, flow cyometry, fluorescence microscopy, histopathology and mouse genetics. The group leaders meet regularly to discuss programme issues. In addition, there are weekly data seminars where members of the programme present their work. The programme has presently two group leader openings, and searches are ongoing to fill these positions.

There were two highlights during the year. The first was the publication of a cover article of Luciano DiCroce’s group in Molecular Cell on the role of polycomb proteins in PML-RAR induced leukemia. This article also received a commentary in the same issue. The second was an article by Pura Muñoz’s group in Cell Metabolism, which described an essential role of IL-6 for the recruitment of muscle satellite stem cells during hypertrophic muscle growth. This article, which was highlighted by Cell Metabolism, received a great deal of attention in the Spanish press and media.
THOMAS GRAF HAS A SENIOR ICREA GROUP LEADER POSITION.

GROUP STRUCTURE:

Group Leader: Thomas Graf (ICREA, CRG)

Postdoctoral Fellows: Florencio Varas (FIS, CRG)
Alexis Schubert (Marie Curie)
Ester Sanchez Tillo (CRG)
Francisca Rubio Moscardo (HEROIC)

PhD Students: Lars Bussmann (CRG)
Alessandro DiTullio (Leonardo, CRG)

Technicians: Luisa Irene de Andrés (CRG)
Vanessa Chigancas (Plan Nacional)
Mattia Schiaulini (CRG)

DIFFERENTIATION
AND CANCER

Hematopoietic Differentiation and Stem Cell Biology

The lab focuses on transcription factor-induced trans-differentiation of blood cells, the role of transcription factors in blood cell differentiation and hematopoietic stem cell biology.
1. Reprogramming B lineage cells into macrophages: A model system to study the molecular mechanism of transdifferentiation

During blood cell formation, the earliest multipotent progenitors branch into a common lymphoid and a common myeloid precursor. However, it is largely unknown which transcription factors determine the branching between the two compartments, and how these two lineages become established and maintained. In attempts to answer this question, we have chosen the approach developed in our earlier work with chicken cell lines, consisting in the perturbation of transcription factor networks by enforced transcription factor expression in committed hematopoietic cells.

Previously we found that the bZip type transcription factor C/EBPalpha, which is expressed in macrophages but not in B cell precursors, effectively induces a switch of B cell precursors towards functional macrophages. The activation of myeloid genes requires the collaboration between C/EBPalpha and the transcription factor PU.1, which is expressed in both B cell precursors and immunoglobulin positive B cells. In contrast, the extinction by C/EBPalpha of the late B cell marker CD19 is PU.1 independent, and is caused by an inhibition of the CD19 regulator, Pax5. That the induced changes are due to a true trans-differentiation and not to the selection of inadvertent myeloid contaminants in the cells examined could be shown by an in vivo lineage tracing experiments. In addition, the macrophages generated exhibited immunoglobulin rearrangements, again unambiguously showing their B cell origin. Using this system we are asking the following questions:

A) How complete is the C/EBP alpha induced trans-differentiation?

We have now developed an inducible system that consists of a fusion between C/EBPalpha and the estrogen hormone binding domain. Primary B cell precursors expressing this protein can be induced to synchronously reprogram into macrophage-like cells within a few days after addition of estradiol. Affymetrix gene expression arrays showed that 1000 probesets (equivalent to approximately 500 genes) are upregulated within 3 hours after induction and a similar number are downregulated. After 5 days, these numbers increase to about 5000 probesets that go either up or down. A comparison of the genes that have changed in day 5 induced cells with the gene expression pattern in normal macrophages reveals that there is an overlap of 75% to 80% of the gene expression patterns. However, many of the gene expression differences seen are not due to differences in differentiation but to the selective downregulation in the artificial macrophages of cell cycle, DNA repair and DNA replication genes, possibly reflecting the fact that C/EBP alpha induces cell cycle arrest.

B) Do the cells retro-differentiate before they go forward?

In principle there are several possibilities how trans-differentiation could occur. The lymphoid cells might trans-differentiate directly, generating the expression profile of macrophages by linearly upregulating macrophage specific genes and downregulating B cell specific genes. Or they might go back in differentiation before going forward. If so, the cells might temporarily acquire an expression program characteristic of the closest monopotent progenitor, the nearest lymphoid/ myeloid progenitor, hematopoietic stem cells or even to ES cells. To address this question we have obtained Affymetrix gene expression data sets of hematopoietic stem cells as well as various intermediate hematopoietic progenitors (collaboration with R Manson and SE Jacobsen) and compared their gene expression profiles with that of pre B cells, macrophages and cells at various stages of transdifferentiation from pre B cells to macrophages. Our studies so far suggest that during C/EBP alpha induced B cell to macrophage conversion a substantial number of the genes that are induced to change temporarily recapitulate the expression pattern of CLPs and LMPPs, progenitors that represent lymphoid and lymphoid/myeloid committed progenitors, respectively.
2. A novel pre B cell line system to study trans-differentiation

A major limitation of the B to macrophage trans-differentiation model developed is that only limiting numbers of preB cells can be obtained from animals, precluding large-scale biochemical and epigenetic studies. We therefore developed a cell line that can be induced to trans-differentiate into macrophage-like cells by introducing a fusion of by C/EBP alpha and the estrogen receptor hormone binding domain and treating the cells with beta estradiol. These cells downregulate B cell restricted transcription factors and genes of the B cell receptor complex within hours after addition of the hormone. They also upregulate macrophage specific genes to levels similar to those seen in control macrophages. The trans-differentiated cells exhibit several properties of functional macrophages, such as large size, phagocytic capacity and substrate adherence. Induced differentiation becomes irreversible after the inducible pre B cell line is exposed for 20 to 36 hours to beta estradiol followed by hormone withdrawal. Finally, experiments with antisense constructs showed that this system can be used to perform loss of function studies.

3. Transcription factors with blood cell instructive capacity can reprogram fibroblasts into macrophage like cells

It is widely assumed that the only truly totipotent cells of the body are embryonic stem (ES) cells whose chromatin becomes irreversibly modified as they progress through differentiation. However, the fact that mammals can be cloned by transfer of nuclei of differentiated cells into oocytes shows that even specialized somatic cells can be epigenetically reprogrammed to acquire totipotency. Therapeutic cloning, in which ES cells are first developed from the nuclei of somatic cells which can then be used to generate desired cell types, is based on these insights. Recent experiments suggest that it is even possible to reprogram somatic cells in culture: Expression of a specific combination of four transcription factors in fibroblasts induced the acquisition of an ES cell phenotype, including the capacity of the cells to generate all three germ layers. However, as in the nuclear transfer experiments, the induced reprogramming was very inefficient. At present the possibility that only a small percentage of somatic cells retain a chromatin configuration susceptible to remodeling cannot be distinguished from other possibilities, such as the activation, by retroviral integration, of genes that contribute to the process (see below).

An explanation for the apparent ease with which committed hematopoietic cells can be converted into one another is that they are closely developmentally related, thus expressing overlapping sets of transcription factors and sharing similar chromatin configurations. Implied in this assumption is that hematopoietic transcription factors should not be able to reprogram more distantly related cells. We therefore determined whether expression of the myeloid lineage instructive transcription factors PU.1 and C/EBP alpha can alter the phenotype of fibroblasts to express macrophage markers. Fibroblasts are non-hematopoietic cells derived from mesenchymal stem cells, as are adipocytes, myocytes, osteoblasts and chondrocytes. We found that the combination of an attenuated form of PU.1 together with either C/EBP alpha or C/EBP beta induced the upregulation of a large number of macrophage/ hematopoietic markers in NIH-3T3 fibroblasts. The combinations of these factors also upregulated macrophage markers in fibroblasts derived from mouse embryos and adult skin. Based on changes in cell morphology (Fig.1), activation of macrophage-associated genes and extinction of fibroblast-associated genes, cell lines containing PU.1 and C/EBP alpha resembled macrophages. The lines also displayed macrophage functions: they phagocytozed small particles and bacteria, mounted a partial inflammatory response and exhibited strict CSF-1 dependence for growth. The myeloid conversion is primarily induced by PU.1, with C/EBP alpha acting as a modulator of macrophage specific gene expression. Our data suggest that it might become possible to directly induce the trans-differentiation of skin-derived fibroblasts into cell types desirable for tissue regeneration without first generating ES cells.
4. Absence of common retroviral integration sites in fibroblast-derived pluripotent cells

As mentioned above, several laboratories have reported the induced reprogramming of mouse and human fibroblasts into pluripotent ES-like cells, using retroviruses carrying the Oct4, Sox2, KLF7 and Myc transcription factor genes. In these experiments the frequency of reprogramming was less than 1% of the transfected cells, raising the possibility that retroviral integrations activate an endogenous gene(s) that is additionally required to induce reprogramming. We have therefore determined the retroviral integration sites in 6 mouse fibroblast derived induced pluripotent stem (iPS) cell clones, using LM PCR. Fifty four integration sites were identified that mapped to gene transcription units or that were within gene reach areas. When genes were included whose transcription start sites are within 200 kb from the retroviral insertion sites, thus becoming potentially activated by the integrations, the number of candidate genes increased to 249. However, none of these genes were found to be targeted in more than one iPS clone and bioinformatic analyses revealed no enrichment within a common network or pathway. Our observations therefore support the notion that the 4 transcription factors commonly used for fibroblast reprogramming are sufficient for the conversion into iPS cells and that the observed low conversion frequencies must have alternative explanations.

5. Lineage priming in hematopoietic stem cells (HSCs), is transcriptionally regulated and reveals heterogeneity within the HSC compartment

Studies on the gene expression of multipotent hematopoietic progenitors revealed that they express, surprisingly, markers of both myeloid and erythroid cells, a phenomenon termed lineage 'priming' In previous work we extended this concept to bona fide HSCs, using lineage tracing with a mouse that expresses the Cre recombinase gene under the control of the lysozyme gene (a macrophage marker). We could also show more directly the expression of lysozyme in a subset of HSCs using another knock in model, in which GFP had been
inserted into the lysozyme locus. However, how these lineage restricted genes are regulated in HSCs remained unexplored. We have now studied this question first by showing that HSCs express various cell surface myeloid markers, such as Mac-1 and Gr-1, besides lysozyme. Then we examined the effect of reducing the concentration of key hematopoietic regulators on the expression of lysozyme and myeloid surface antigens. Our studies showed that in animals heterozygous for PU.1 there is indeed a significant reduction in the expression of all myeloid markers tested in HSCs. In contrast, reduction of the levels of C/EBP alpha expression led to an decrease in some markers while others were increased. These results indicate that C/EBP alpha acts both as an activator and repressor of myeloid expression in HSCs. Finally, no significant effect was seen in mice heterozygous for C/EBP beta. We also examine the differentiation potential of lysozyme positive and negative HSCs and found that surprisingly, the latter yielded a higher frequency of lymphoid versus myeloid differentiation. These observations suggest that myeloid transcription factors are the mediators of lineage priming and that additional epigenetic mechanisms must exist that translate this early gene expression into a biased lineage readout, thus also implying heterogeneity within the HSC compartment.

**PUBLICATIONS**


(*) All these publications are the result of the work of Dr. Thomas Graf at Albert Einstein College of Medicine, New York, USA.
The main interest of our group is to elucidate the mechanisms controlling myogenesis in vitro and in vivo, with an emphasis in skeletal muscle regeneration and inherited myopathies. Myogenesis is largely controlled by the basic helix-loop-helix (bHLH) family of muscle regulatory transcription factors (MRFs), including MyoD, Myf5, myogenin and MRF4, and by the myocyte-enhancer factor-2 (MEF2) family of proteins, which regulate the expression of muscle-specific genes. The p38 MAPK activity is induced during myogenic differentiation, being this activation required for myoblast fusion and differentiation in vitro. We are interested in analyzing the molecular mechanisms responsible for the promyogenic effect of p38. Moreover, we aim to investigate the role of the distinct isoforms of the p38 MAPK in vivo, using mice deficient in each isoform. Based on our earlier work, a strong emphasis is also devoted in our laboratory to the analysis of the role of the plasminogen activation (PA) system components in skeletal muscle regeneration (after injury or in inherited myopathies). Finally, Antonio Serrano (a Ramón y Cajal investigator in our laboratory) is analyzing the mechanisms involved in the regulation of muscle fiber type and size in vivo. In summary, our laboratory is pursuing three main lines of research:

- Control of muscle-specific gene transcription by p38 MAP kinases.
- Role of the plasminogen system in skeletal muscle regeneration in vivo.
- Molecular mechanisms regulating adult skeletal myofiber growth and type.
RESEARCH PROJECTS

1. Control of muscle-specific gene transcription by p38 MAP kinases

Myogenesis is largely controlled by the basic helix-loop-helix (bHLH) family of muscle regulatory factors (MRFs), including MyoD, Myf5, myogenin and MRF4. The MRFs exert their function by promoting muscle-specific gene transcription through a specific DNA sequence, the E-box. Selective and productive recognition of chromatin targets requires heterodimerization of MRFs with the ubiquitous E proteins, E12 and E47. Thus, formation of the functional MRF/E47 heterodimer is pivotal in controlling muscle gene expression. We have investigated the potential mechanisms underlying the role of p38 MAP kinases in this process. Our results so far have been the following:

- p38 MAPK activity regulates MyoD/E47 association in vitro and in vivo by phosphorylation of the E protein at serine 140, thereby promoting muscle-specific gene transcription

- p38alpha (but not p38beta, p38gamma or p38delta) is necessary for myogenesis. p38alpha controls myoblast proliferation both in vitro as in neonatal muscle. Indeed, myoblasts derived from p38alpha-deficient mice show persistent proliferation, and a subsequent block in myoblast differentiation. This result uncovers a novel mechanism explaining the fundamental role of p38 in myogenesis, and constitutes the first dissection of the relative contribution of the four p38 MAP kinases to this process.

2. Role of the fibrinolytic system in skeletal muscle regeneration in vivo.

We had previously observed a muscle regeneration defect in uPA (urokinase plasminogen activator)-deficient mice after injury, which correlated with fibrin deposition and a decreased recruitment of blood-derived monocytes and lymphocytes to the damaged muscle. Our results are the following:

- uPA deficiency exacerbates dystrophy in mdx mice, an animal model of Duchenne Muscular Dystrophy (DMD), and injury-induced muscle regeneration, via two mechanisms: fibrin degradation and mobilization of bone marrow cells.

- Fibrin accumulates in the muscle of mdx mice, exacerbating dystrophy progression, and have analyzed the underlying mechanisms mediating the deleterious role of fibrin.

- Preliminary results show that the muscle size of PAI-1-deficient mice is larger than that of wild type mice, suggesting that PAI-1 may be negatively regulating muscle growth. The underlying mechanisms are being analyzed at present.

From Perdiguero E, Ruiz-Bonilla V, Serrano AL, Muñoz-Cánoves P. Cell cycle 2007

Skeletal muscles adapt to increasing workload by augmenting their fiber size, through yet poorly understood mechanisms. We identified the cytokine IL-6 as an essential regulator of satellite cell-mediated hypertrophic muscle growth. IL-6 is locally and transiently produced by growing myofibers and associated satellite cells (muscle stem cells), and genetic loss of IL-6 blunted muscle hypertrophy in vivo. IL-6 deficiency abrogated satellite cell proliferation and myonuclei accretion to the preexisting myofiber, by impairing STAT3 activation and expression of its target gene cyclin D1. We plan to continue investigating in further detail the mechanisms controlling growth of adult skeletal muscle.

PUBLICATIONS


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.

Our research investigation is focused on understanding the role of several protein complexes that are involved in chromatin dynamics and metabolism, which when altered could participate in the establishment and maintenance of the aberrant silencing of tumor suppressor genes during transformation. Our results suggested that the Nucleosome Remodelling and Deacetylase complex (NuRD), Polycomb group of proteins (PcG) and the histone variant macroH2A are - with different timing and kinetics - involved in setting up an altered chromatin structure with aberrant gene silencing in acute promyelocytic leukemia (APL). We believe that the results of our research will lead to the identification of new biological tools with a potential impact on cancer therapeutic intervention.
RESEARCH PROJECTS

1. Polycomb and cancer

Polycomb repressive complex 2 has been strongly implicated in cancer development, but to date mechanistic insight into the function of PRC2 in cancer cells is lacking. In addition, in mammalian cells, it is not well understood how PRC2 is targeted to promoter regions. Using as paradigm the oncogenic transcription factor PML-RARa, we have investigated the role of Polycomb group proteins in the establishment and maintenance of the aberrant silencing of tumor suppressor genes during transformation induced by the leukemia-associated PML-RARa fusion protein. We show that in leukemic cells knockdown of SUZ12, a key component of Polycomb repressive complex 2 (PRC2), reverts not only histone modification but also induces DNA de-methylation of PML-RARa target genes. This results in promoter reactivation and granulocytic differentiation. Importantly, the epigenetic alterations caused by PML-RARa can be reverted by retinoic acid treatment of primary blasts from leukemic patients. Our results demonstrate that the direct targeting of Polycomb group proteins by an oncogene plays a key role during carcinogenesis.

2. Biochemical link between epigenetic marks and chromatin remodelling

In plants, as in mammals, mutations in SNF2-like DNA helicases/ATPases were shown to affect not only chromatin structure but also global methylation patterns, suggesting a potential functional link between chromatin structure and epigenetic marks. The SNF2-like ATPase containing nucleosome remodelling and deacetylase corepressor complex (NuRD) is involved in gene transcriptional repression and chromatin remodelling. As mentioned above PML-RARa represses target genes through recruitment of DNMTs and Polycomb complex. We have recently investigated a direct role of the NuRD complex in aberrant gene repression and transmission of epigenetic repressive marks in acute promyelocytic leukemia (APL). Our results indicate that PML-RARa binds and recruits NuRD to target genes, including to the tumor-suppressor gene RARß2. In turn, the NuRD complex facilitates Polycomb binding and histone methylation at lysine 27. Retinoic acid treatment, which is often used for patients at the early phase of the disease, reduced the promoter occupancy of the NuRD complex. Knock-down of the NuRD complex in leukemic cells not only prevented histone deacylation and chromatin compaction, but also impaired DNA and histone methylation as well as stable silencing, thus favouring cellular differentiation. These results unveil an important role for NuRD in the establishment of altered epigenetic marks in APL, demonstrating an essential link between chromatin structure and epigenetics in leukemogenesis.

3. 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 ChIP-on-chip analysis, using Agilent microarray, shows that macroH2A is deposited in the regulatory regions of tumor suppressor genes and genes regulated by Polycomb group in an isoform-specific fashion. We are studying the role of macroH2A in transcriptional regulation of these genes during cell fate decisions and cancer development.

**PUBLICATIONS**


We are interested in studying pathways involved in self-renewal and homeostasis of adult epithelia and how these contribute to the progression and aggressiveness of human tumours.
RESEARCH PROJECTS

1. Impact of Rac1/PAK2/Myc pathway in epidermal homeostasis and cancer in vivo.

Adult epithelia are in constant need of renewal. A population of adult stem/progenitor cells (SCs) ensure maintenance of the undamaged tissue but also integrity in response to external stimuli. In our lab, we primarily use the skin as a model of epithelium with a high rate of turn-over and with well defined somatic epSCs populations. Homing and exit of these from their niche are tightly regulated processes that integrate extracellular cues and cell autonomous genetic programmes. We aim to identify some of the molecular mechanisms that modulate the behaviour of adult stem cells and how the tight control of these signals is lost in tumours.

Our recent work has shown that the interplay between Rac1 GTPase and the proto-oncogene Myc is essential for balancing epidermal renewal and differentiation (Benitah et al., 2005; Benitah and Watt, 2007). The family of Rho GTPases is involved in changes in cell morphology, adhesion, invasion, polarization and proliferation, among others (Benitah and del Pulgar et al, 2005). Moreover, Rho GTPases are of clinical interest since deregulation of several members of its family is a common feature in human cancers that correlates with aggressive tumour behaviour (Benitah and del Pulgar et al, 2005).

We have deleted Rac1 in the undifferentiated compartment of adult epidermis, hair follicle and sebaceous glands in an inducible manner (Keratin-14-CreER/Rac1flox/flox). In these mice activity of Cre recombinase is dependent on administration of tamoxifen. Deletion of Rac1 in K14CreER/Rac1KO mice causes a rapid proliferation and irreversible mobilization of the epidermal and hair follicle SCs (epSCs) from their niche (Figure 1). Depletion of epSCs upon Rac1 deletion ultimately results in alopecia, due to loss of hair follicle cycling, and failure of epidermal renewal. A similar effect was observed upon embryonic deletion of Rac1 in the epidermis (Keratin-5-Cre/Rac1flox/flox mice; Benitah and Watt, 2007). Accordingly, Rac1, together with the hematopoietic specific Rac2, is essential for haematopoietic stem cell maintenance and proper haematopoiesis (Cancelas et al, 2005).

At the molecular level, Rac1 maintains epSC self-renewal by modulating the activity of the transcription factor Myc (Benitah et al, 2005). Myc is one of the first and best characterized human oncogenes found to date, and much effort is being invested to elucidate its precise functions in normal tissues and in tumours. In steady-state conditions, Myc induces exit of epSCs from their niche.

Figure 1 - Inducible deletion of Rac1 in adult epidermis of K14CreER/Rac1flox mice. Upper Panel: Loss of Rac1 causes an early hyperproliferative phase (Early) followed by irreversible terminal differentiation (Middle). Ultimately, the epidermis looses its self-renewal capacity and only remnants of the cornified stratum remain (Late). Lower Panel: Hair follicle stem cells, depicted as Label Retaining Cells (LRCs), mobilize out of the niche and proliferate upon epidermal KO of Rac1. Wholemounts of mouse tail epidermis show that upon KO of Rac1, LRCs egress out of the hair follicle stem cell niche (KO Early) and proliferate (Ki67 positive; KO Early and Middle). Ultimately, mobilization out of the niche coupled to proliferation results in loss of LRCs (KO Late).
to permit their subsequent differentiation (Waikel et al, 2001; Arnold et al, 2001; Frye et al, 2003). However, sustained overexpression of Myc in the epidermis causes a decline of epSCs number and a hyperproliferative state that predisposes skin to the development of squamous tumours (Waikel et al, 2001). The effect of Myc on epSCs takes place through a bifunctional mechanism: induction of proliferation and de-adhesion from the stem cell niche (Gebhardt and Frye et al, 2006).

We have described that the serine/threonine kinase PAK2 colocalizes to, and is activated by, Rac1 in the epidermis. Once active, PAK2 phosphorylates Myc at three C-terminal aminoacids, Thr358, Ser373 and Thr400 (Benitah et al, 2005; Huang et al, 2004). Phosphorylation of Myc at these residues modifies its transcriptional activity, preventing its function over epSCs (i.e exit from the stem cell niche). We are further characterizing the impact of the Rac1/PAK2/Myc pathway on the homeostasis of the skin in vivo using novel epidermal mouse models.

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

We have performed Affymetrix microarray analysis of primary human keratinocytes kept in a stem cell state or a transit amplifying state. Since the Rac pathway is important for balancing epidermal homeostasis, in addition, we are using a proteomic approach to identify proteins that interact with members of this signalling cascade. 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.

3. 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 cavity. 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.

PUBLICATIONS


(*) This publication is the result of the work of Dr. Salvador Aznar-Benitah at the London Research Institute, Cancer Research UK, London, UK.
The Genes and Disease Programme of the CRG aims to use genomic variability to identify the molecular basis of disease, to investigate the function of genes with potential roles in common human disorders, and to develop therapeutic approaches to correct some diseases. The Programme combines large-scale experimental approaches with advanced genetic strategies to elucidate some of the biological determinants of 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 Down syndrome phenotypic traits, by using models that overexpress dosage-sensitive genes. Programme researchers use murine models to develop therapeutic approaches, that could correct features involved in mental retardation, anxiety disorders and cancer. The Programme Group Leaders are part of the Biomedical Research Networks on Rare Disorders and on Epidemiology and Public Health. Both are CIBERS (“Centros de Investigación Biomédica en Red”) of the “Instituto de Salud Carlos III” (Spanish Ministry of Health). These CIBERS fund research in rare genetic disorders and in genomic and molecular epidemiology. The whole Programme is also supported as a consolidated group, from the Catalan Government. Programme researchers have participated in several networks supported by the “Instituto de Salud Carlos III – Fondo de Investigación Sanitaria” (ISCIII-FIS), including Clinical Genetics, Neuroscience, Hearing Impairment, Psychiatric Genetics, and Cancer, and in research networks supported by the Department of Education and Universities of the "Generalitat de Catalunya", such as the Gene Therapy Network, and the Murine Models Network. All groups of the Programme are supported by National and European grants.

Current Structure of the Programme:

- **Research Groups:**
  - Murine Models of Disease (Mariona Arbonés)
  - Neurobehavioral Analysis (Mara Dierssen)
  - Genetic Causes of Disease (Xavier Estivill)
  - Gene Therapy (Cristina Fillat)
  - Gene Function (Susana de la Luna)

- **Associated Core Facility:**
  - Genotyping Unit (Mònica Bayés)
GENES AND DISEASE
Genetic Causes of Disease

The group focuses on the study of human genome variability at the nucleotide and genomic levels and uses this information to evaluate the conferred predisposition or resistance to various diseases. Large-scale segmental duplications (SDs) and copy number variants (CNVs) have been studied by members of the ENCODE (Encyclopedia of DNA Elements) project to examine in detail a targeted 50 Mb of the human genome. This analysis has provided the basis for further large-scale investigations to characterize smaller CNVs, which could be important for studies of disease and evolution. The group is studying several CNVs in relation to common complex disorders, and is specifically searching for CNVs associated with psoriasis, multiple sclerosis, migraine, and psychiatric disorders. We have also tested the hypothesis that functional genetic variants may confer susceptibility to several related psychiatric disorders (e.g., substance-abuse, anxiety, and eating, psychotic and mood disorders) for genes involved in neurotransmission and/or neurodevelopment. We have found that multiple SNPs are strongly associated with several disorders, and may represent related disease susceptibility factors. We have analyzed the neurotrophin family of regulatory factors in patients affected by psychiatric disorders and detected multiple positive associations, thus indicating the important role that neurotrophin pathways play in mental disorders. The group has evaluated the contribution of RCAN1 (Regulator of Calcineurin) in neuronal dysfunction and death. We have revealed that RCAN1 contributes to increased neuronal susceptibility to oxidative stress, suggesting that RCAN1 is a dosage-sensitive gene in the context of neurodegeneration. In addition, the group is also exploring the contribution of non-coding miRNAs to the susceptibility to complex diseases. We have observed that selective sets of miRNAs are linked to neuronal survival and death responses, thus modulating neurodegenerative processes.
RESEARCH PROJECTS

1. Copy number variants, segmental duplications and genome structure

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 (Figure 1). Many of these SDs coincide with structural variations or copy number variants (CNVs) of the human genome. In collaboration with members of the International Consortium of Copy Number Variants, our group has contributed to the development of the first whole genome map of structural variants of the human genome. Over 1,400 CNVs were detected using SNP and BAC arrays in an analysis of the 270 HapMap samples. The group applied this information to evaluate the content in SDs and CNVs of the genomic regions studied by the ENCODE (Encyclopaedia of DNA Elements) Project. ENCODE involves 35 groups that have provided more than 200 experimental and computational datasets that examined in detail a targeted 30 Mb of the human genome. This analysis has provided the basis for further large-scale investigations to characterize smaller CNVs, which could be important for disease and evolution studies.

2. Copy number variants associated with complex disorders

The group is now studying several CNVs in relation to common complex disorders and are specifically searching for CNVs associated with psoriasis, multiple sclerosis, migraine, and psychiatric disorders. We performed comparative genomic hybridization (CGH) experiments between samples of patients and controls and used multiple technologies to further analyse SDs and CNVs at the genome scale. We developed and validated clone-based arrays for CGH (arrayCGH). We also used Agilent and Illumina platforms to implement the latest oligonucleotide-based scanning approaches, as well as multiplex ligation probe amplification (MLPA) and multiple amplification probe hybridization (MAPH) assays for rapid validation of primary genome scanning data.

We have made specific progress with psoriasis which is a chronic disorder of the skin affecting most ethnic groups, but with the highest prevalence (3%) in northern European populations. We identified several CNVs that are associated with psoriasis. One of them is within the epidermal differentiation complex, in a region that was previously found to be linked to psoriasis. We defined a region of chromosome 1q21, span-
ning genes from this region, varying in copy number between patients and controls. The link between this CNV and psoriasis was further confirmed by SNPs tagging the CNV. Replication of these findings in other cohorts is ongoing. Genes affected by this CNV could be a target for pharmacological and biological treatment and prevention of psoriasis.

3. Genetic variants associated to psychiatric disorders

Epidemiological genetic studies in the psychiatric field have also indicated that genetic risk factors are not specific for individual psychiatric diagnoses but influence liability for a range of disorders. We tested the hypothesis that functional genetic variants may confer susceptibility to several related psychiatric disorders. We analyzed five main psychiatric diagnostic categories (substance-abuse, anxiety, eating, psychotic and mood disorders) and two control groups, for potential functional SNPs, in genes involved in neurotransmission and/or neurodevelopment. We found that multiple SNPs are strongly associated with several disorders, and may represent individual and related disease susceptibilidad factors (Figure 2). A functional SNP in the bradykinin receptor B2 (BDKRB2) gene was associated with panic disorder, substance abuse and bipolar disorder, and two additional BDKRB2 SNPs with obsessive-compulsive disorder and major depression. The association of BDKRB2 with several psychiatric diagnoses supports the view that common genetic variants could confer susceptibility to clinically related phenotypes, and defines a new functional indicator in the pathophysiology of psychiatric diseases.

The neurotrophin family of regulatory factors promote neuronal proliferation, regeneration and connectivity during development, and participates in the plasticity and maintenance of neurons throughout adulthood. We selected several SNPs located in the genomic regions of NGF, NTRK1, BDNF, NTRK2, NTF3, NTRK3, CNTF, CNTFR, NT4-5 and p75 for case-control or family-based association studies in patients with anxiety disorders (panic disorder and obsessive-compulsive disorder), schizophrenia, eating disorders (anorexia nervosa and bulimia nervosa), affective disorder (bipolar disorder and major depression) and substance abuse disorder (opiate dependence). Multiple positive associations were detected for several disorders, indicating the impor-
tant role that neurotrophin pathways play in human behaviour and mental disorders. Some of these associations have been replicated in patients from different European countries.

4. Functional genomics of neurological disorders

Oxidative stress is a common hallmark of many neurodegenerative disorders involved in neuronal dysfunction and/or neuronal death. DSCR1 (Down Syndrome Candidate Region 1, now renamed RCAN1, Regulator of Calcineurin) is overexpressed in Down syndrome and Alzheimer disease patients and possibly contributes to their neuropathology. The group has revealed that RCAN1 is a deleterious factor that contributes to increased neuronal susceptibility to oxidative stress, suggesting that RCAN1 is a dosage-sensitive gene in the context of neurodegeneration.

The group is exploring the contribution of non-coding miRNAs, to the susceptibility to complex diseases. miRNAs are important modulators of gene expression in a number of physiological and pathological conditions (Figure 3). Our recent results suggest that selective sets of miRNAs are linked to neuronal survival and death responses. Functional genomics approaches include miRNA profiling and high throughput sequencing to identify small non-coding RNAs linked to the progress of the disease. Global transcriptomic and proteomic analysis linked to miRNA deregulation will lead to the identification of specific miRNA targets relevant in the context of neuronal dysfunction. Neurodegeneration-linked common processes, such as impaired metabolic pathways, oxidative stress, necrosis and apoptosis are being analyzed. The main aim is to gain information about gene regulatory networks modulating neurodegenerative processes, which might be suitable for diagnosis and identification of putative therapeutic targets.

PUBLICATIONS


Llorens F, Gil V, Iraola S, Carím-Todd L, Martí E, Estivill X, Soriano E, del Río JA, Sumoy L. Developmental analysis of Lingo-
1/Lern1 protein expression in the mouse brain. Interaction of its intracellular domain with Myt1L. Dev. Neurobiol, in press.


Gene therapy is an emerging field that holds the promise of treating a wide variety of diseases. However, before this can be achieved, successful vector systems must be developed to deliver therapeutic genes and successful preclinical studies in animal models need to be carried out. Moreover, a broad understanding of the disease pathology is required to be able to design candidate gene transfer approaches. The group is interested in understanding the pathophysiology and developing optimal gene therapy approaches for pancreatic cancer and Down syndrome. Gene delivery vectors that can selectively and efficiently target the cells of interest are being developed and their gene transfer efficiency is being evaluated in living animals by the use of non-invasive molecular imaging techniques. Pharmacokinetic and pharmacodynamic studies are being conducted to evaluate the therapeutic response in preclinical mouse models. Disease amelioration by partial phenotypic correction, as consequence of particular genetic interventions, could become a key determinant towards the understanding of complex phenotypes.

**GROUP STRUCTURE**

**Group Leader:** Cristina Fillat

**Postdoctoral Fellows:** Xavier Altafaj

**PhD Students:** Meritxell Huch
Jon Ortiz
Laura García
Daniel Abate Daga
Anabel Jose

**Masters:** Alex Bío de Rojas

**Technicians:** Núria Andreu
RESEARCH PROJECTS

1. Pancreatic cancer

One of the most devastating diseases that our society is presently facing is cancer. Particularly pancreatic cancer is the fifth cause of cancer deaths in industrialized countries. This neoplasia has a very bad prognosis mainly due to the late diagnosis and inefficient current therapies. Gene therapy emerges as a candidate approach for the treatment of this type of cancer. The group has been involved in the past few years in exploring the feasibility of gene directed enzyme prodrug therapy in pancreatic cancer, showing some but limited effects. Based on those observations we are presently interested in the development of more potent and selective agents. To be able to increase their potency we are working with different systems that may facilitate the spreading of the cytotoxic compound into the tumor mass, such as by the use of Protein Transduction Domains (PTDs) fused to suicide genes or the modulation of intercellular gap junctions. We have also explored the feasibility of arming replicative adenovirus, with the suicide Herpes Simplex Thymidine Kinase (TK) gene and shown that effective pancreatic tumour regression can be achieved with this system when combined with ganciclovir. However, in order for the virus to prove of therapeutic value, a degree of selectivity is required. With a view to the latter, we evaluated the possibility of controlling replication through the tumor selective expression of E1A. We extensively studied the potential selectivity of candidate promoters, and found evidence of tumour specificity. Moreover the group is also interested in the understanding of the basic mechanisms of tumor cell killing induced by particular suicide systems and in the definition of the genetic signatures that can sensitize pancreatic tumors to these particular therapies.

2. Down syndrome

Down syndrome is the most common cause of mental retardation. It is a multi-system disorder with a wide range of physical features, health and development problems. As we begin to understand the role that specific genes could have in Down syndrome, we perform partial gene therapy approaches for specific defects. Alternatively, the use
of gene therapy approaches can be a very useful tool to contribute to understand the role of particular genes to the phenotype. From the results of our group and others there are strong evidences indicating that DYRK1A is one of the genes that play a relevant role in Down syndrome. Indeed, transgenic mice that overexpress DYRK1A present neurodevelopmental delay, motor alterations and cognitive deficits, similar to those described in Down syndrome patients. We are currently testing the rescue of defined Down syndrome phenotypes in the TgDyrk1A and in the Ts65Dn models by reducing DYRK1A overexpression in the brain by RNA interference technology.

**PUBLICATIONS**


The overall goal of the laboratory is to investigate the in vivo function of genes involved in diseases affecting the central nervous system (CNS) using the mouse as a model system. Research projects of the laboratory focus on the analysis of human chromosome 21 genes that are dosage sensitive and are predicted to contribute to the morphological and cognitive deficits associated to both, monosomy and trisomy 21. Since aneuploidies are developmental disorders, we are studying genes that are important in the development of the brain. Previous work of the group showed that DYRK1A plays a fundamental role in the development of the brain by regulating the total number of neurons that are produced during neurogenesis. We are currently studying the role of DYRK1A in the CNS trying to identify, at the cellular and molecular levels, functions that are regulated by this protein kinase. In addition, we have contributed to the functional analysis of RCAN1 (DSCR1), another chromosome 21 gene, that regulates calcineurin activity and is involved in Down syndrome neurodegeneration and in Alzheimer’s disease.
RESEARCH PROJECTS

1. Role of the protein kinase DYRK1A in retina development

To provide some insights about the function of DYRK1A in CNS development we have analyzed the impact of DYRK1A gene dosage variations in the retina, a CNS structure that has been extensively used as a model system to study neurogenesis. Similarly to what has been previously found in the brain, mice with only one functional copy of Dyrk1A (Dyrk1A+/- mice) have retinas that are thinner than normal with decreased numbers of specific types of cells. Conversely, mice with three copies of the gene (tgYAC152f7) have larger retinas with more cells. Analysis of the embryonic and postnatal neurogenesis of these two strains of mice showed that changes in the amounts of DYRK1A protein do not alter the proliferation and the specification of retina precursor cells. Unexpectedly, we found that DYRK1A tightly regulates the number of cells that die through activation of the intrinsic cell death pathway during differentiation.

2. Effect of DYRK1A on adult brain neural stem/progenitor cells

The production of new neurons is a process that in mammals last for the entire life. To investigate if DYRK1A has any effect in adult brain neurogenesis, we have performed an in vivo analysis of the neurogenic niche of the subventricular zone (SEZ) of the adult brain in Dyrk1A+/- mice and control littermates. We have found that Dyrk1A+/- mice have less GFAP-positive stem cells in their SEZs than wild-type mice, and that DYRK1A is essential to maintain the proliferative activity of these cells in vivo. In accordance with this result, we showed that neuronal stem/progenitor cells (NSCs) isolated from the SEZ of Dyrk1A+/- mice are less responsive to mitogenic stimulation in a sphere-forming assay than control cells. Sphere cells seeded at clonal densities showed that EGF-dependent self-renewal capacity of Dyrk1A+/- NSCs is significantly impaired. Our results point to DYRK1A as a new regulator of the neurogenic niches in the adult brain.

PUBLICATIONS


The overall goal of our research is to understand the role of putative candidate genes for human complex genetic diseases that affect the structural elements connecting the neurons with consequences on brain circuits that underlie cognitive systems. Genetically modified mouse models are a critical resource in the characterization of genes of biological importance, and in the dissection of the pathogenesis of neuropsychiatric and neurological disorders. A mouse model of human disease or gene function is, however, of limited value unless properly characterized and accurate phenotype assessment is a core issue of genetic manipulation. This will lead to a better knowledge of the genetic substrates regulating the expression of complex behavioral traits.
RESEARCH PROJECTS

1. Down syndrome

The neuropathological processes underlying Down syndrome (DS) mental retardation and their genetic dependence remain an open question. Our laboratory is investigating specific links between cognitive impairments and memory disorders in patients with DS and behavioral deficits in mouse models of this disease. We are also currently working on candidate genes involved in dendrites/spine dysmorphology and altered neural plasticity in learning and memory brain circuits. The aims of the project are: a/ to identify the physiological role of gene products and the dosage-dependent effects on neurodevelopment, learning and memory, and neurodegenerative processes in DS; and b/ to identify cellular and molecular substrates that regulate the emergence of different forms of learning and memory. We are coordinating the work package for mouse models in a EU Integrated Project with the aim to make use of partial trisomy and monosomy mouse models (Y. Herault, CNRS). The second phase of this project is oriented to determine how the interaction between these gene products and the environment contribute to the expression of learned behaviors in vivo neuroplasticity models.

2. Pathogenetic and molecular mechanisms involved in anxiety disorders

Patients with neuropsychiatric disorders, such as panic/anxiety disorders, have an altered emotionality profile and abnormal social behaviors. A second research line in our group is aimed at identifying genetic causative and vulnerability factors underlying anxiety-related behavior and that could predict the onset of panic disorder. To this aim we use genetically modified animal models that help to elucidate mechanisms that may be acting in humans. Our current project is focused on the biochemical changes that occur in the CNS during development that are determinant of emotionality-related behaviors and to the development of therapeutic strategies that may overcome and/or prevent the brain alterations leading to panic attacks. Our interest is also aimed at elucidate the deficits of specific neurotransmitter systems that possibly underlie the inability of persons with anxiety disorders to correctly identify the fear-related information and the possible common neurobiological pathways responsible for co-morbid processes. We are interested in candidate genes that participate in the dysfunction of brain circuits involved in fear-related memories and in mouse behavioral traits relevant to panic and to anxiety. We have demonstrated that NTRK3 (TrkC) can exert a major role in anxiety disorders based on the observation that its overdosage leads to an increased anxiety-like behavior and panic reaction, possibly due to the trophic effect attained on the catecholaminergic nuclei. We now explore specific factors affecting the course of the disease in particular those that are associated with the regulatory capacity of the stress system. In collaboration with Dr. R. Maldonado (Pompeu Fabra University) we have initiate a project for studying the implication of NTRK3 in processes co-morbid to panic disorder, such as predisposition to substance abuse. Also we will analyze the predisposition to stress (collaboration with Dr. A. Armario, Autonomous University of Barcelona). The role of other gene products like different nicotinic receptor subunits in behavior is also considered. Mice overexpressing alpha 7, alpha 5, beta 2, or beta 4 subunits, will be used to study the contribution of different nicotinic receptor subunits to (a) the expression of normal behaviors, (b) the sensitivity to the behavioral effects of panicogenic/panicolitic agents, and (c) the development of dependence and tolerance.

3. Phenotype Ontology Project

Phenotype analysis of mice has tended to be qualitative rather than quantitative in nature, but new more sophisticated tests of locomotor and cognitive function to mice are emerging. In the context of an Integrated European Project, our group is the Phenotypic Unit responsible for the second stage specialized phenotypic analysis of an in vivo library of partial trisomies/monosomies of MMU16. The results of the phenotypation will be standardized and categorized; to ensure that behavioral, functional and morphological
characterization methods should be directly comparable between different groups to build a useful body of data. Traditional phenotypic descriptions are captures as free text but the information retrieval based on free text is extremely limited because of the inherent lack of accuracy and specificity. It is thus necessary to develop phenotypic descriptors that allow the creation of phenotypic annotation, so that one or more phenotypic descriptors will provide the phenotype of a particular model.

PUBLICATIONS


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. The group studies the functional roles of several HSA21 genes.
DYRK1A is one of the HSA21 genes for which changes in gene doses result in neuropathological alterations. It encodes for a protein kinase of the DYRK family of kinases. DYRK kinases (DYRK and HIPK subfamilies) constitute one of the families that belong to the CMGC group of protein kinases, formed by the CDKs, MAPKs, GSKs, CLKs and SRPKs. There are several reasons for choosing DYRK1A as one of our target molecules for investigation. Firstly, we regard the phenotypes shown by transgenic mice in which the gene is either overexpressed or has been deleted to be very interesting, and secondly, we believe that DYRK1A might act as a crossroads for different signalling pathways since its substrates list consists of a variety of both cytosolic and nuclear proteins, transcription factors included.

Although DYRKs phosphorylate their substrates on serine and threonine, they autophosphorylate their activation loop on an essential tyrosine. This event is the result of an intramolecular phosphorylation reaction coupled to DYRKs translation that renders enzymes fully active. Given that an activating kinase appears not to be necessary, one might think that there is no room for activity regulation. However, DYRK1A seems to be extremely sensitive to gene dosage, and thus it is plausible to hypothesise that minimal changes in its activity would give rise to profound effects on the pathways it might control. With this in mind, we are interested in finding mechanisms that could possibly regulate the activity of DYRK1A. We have been able to show that DYRK1A autophosphorylates, via an intramolecular mechanism, on Ser-520. Phosphorylation of this residue, which seems to be subjected to dynamic changes in vivo, mediates the interaction of DYRK1A with 14-3-3β. Our results suggest a model in which the catalytic activity of DYRK1A is regulated by autophosphorylation and binding to 14-3-3 protein (Figure 1). Thus, our work argues against the current view of DYRK1A as a constitutive kinase that acquires full competency during its translation, and provides the first clear evidence of a mechanism of regulation for DYRK1A activity.

Other putative DYRK1A interacting-proteins are being explored. These include not only substrates for DYRK1A (downstream targets), but also molecules that can act as effectors and thus, represent upstream modulators of DYRK1A in signalling cascades. For instance, we have identified Sprouty2, a well-known antagonist of receptor tyrosine kinases (RTK) signaling, as an interacting partner of DYRK1A. DYRK1A and Sprouty2 are present in protein complexes in mouse brain, they co-purify with the synaptic plasma membrane fraction, and their expression overlaps in several structures in this tissue. We are currently studying the functional outcomes of this novel interaction.

![Figure 1. 14-3-3 binds to DYRK1A via a phospho-serine 520. Binding results in an increase in the catalytic of DYRK1A towards exogenous substrates. Phosphorylation levels of serine residue 520 are modulated by fibroblast growth factor (FGF) stimulation.](image-url)
2. RCAN1: a HSA21 gene involved in the regulation of the phosphatase calcineurin
Lali Genescà, Àlicia Raya

A growing number of substrates and regulatory proteins have been shown to bind to the catalytic domains of kinases or phosphatases through docking groove interactions. This type of interaction not only has a tethering role but is also a means of regulating the activity of these signaling molecules. A prominent example of this is the MAPK cascade, in which docking interactions are responsible for the recruitment of not only substrates but also positive and negative regulators. Recent reports have suggested that the phosphatase calcineurin binds one group of substrates, members of the NFAT family of transcription factors, through a similar docking interaction. We are currently characterizing the motifs involved in the interaction of calcineurin with one group of modulators, the RCAN protein family. Members of this family, which is conserved from yeast to humans, bind to and inhibit calcineurin-mediated activities in vitro. The human RCAN family is composed of three members, RCAN1, RCAN2 and RCAN3, which have a high amino acid identity in the central and C-terminal region. RCAN1, originally named DSCR1 for its location in the Down Syndrome Critical Region on human chromosome 21, is the most-studied RCAN in mammals mainly because of its potential roles in Down syndrome and Alzheimer’s disease. Our results suggest that the same docking domain in calcineurin that interacts with NFAT is used by RCAN family members. We expect to decipher the molecular mechanism underlying the negative action of RCAN proteins on the CN/NFAT pathway.

PUBLICATIONS


Single nucleotide polymorphisms (SNPs) are the most common type of genetic variation in the human genome. A small 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, hypertension, diabetes, asthma and cancer. SNP genotyping can be useful for genetic mapping, disease association studies, population genetics, and for other fields of research, including research in model organisms and trait selection for agricultural, cattle farming or aquaculture applications. Some of the genotyping technologies also enable copy number variants (CNVs) to be identified, loss of heterozygosity (LOH) to be accurately characterized and DNA methylation status to be assessed. As a core facility we are committed to providing medium and high throughput genotyping services and analysis support at established rates to scientists from the CRG, PRBB (UPF and IMIM) and external public and private institutions. Our research activity, in some cases arising from the involvement of the Genotyping Unit in collaborative projects, is focused on the identification of genes that contribute to genetically complex disorders.

Group Leader: Xavier Estivill
Unit Responsible: Mònica Bayés
Postdoctoral Fellows: Magda Montfort
Technicians: Carles Arribas, Anna Brunet, Cecilia García, Kristín Kristjansdottir, Silvia Carbonell, Josiane Wyniger, Anna Puig

GENES AND DISEASE
Associated Core Facility: Genotyping Unit
SERVICES

The Genotyping Unit, supported by “Genoma España”, through the National Genotyping Centre (CeGen, www.cegen.org) provides support to scientists for genotyping projects in every aspect of research, from planning, DNA extraction, genotyping, data interpretation, through to statistical analysis. The Unit offers custom, cost effective and flexible solutions for projects of any scale to internal users and external users from public or private institutions.

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

- Automated DNA extraction from blood or other tissues (Chemagen)
- DNA quantification using Picogreen (Molecular Probes)
- Whole Genome Amplification using GenomiPhi (Amersham)
- Custom Genotyping by SNPlex (Applied Biosystems): genotyping of 24-48 SNPs selected by customer.
- Custom Genotyping with GoldenGate technology (Illumina): genotyping of 96-1,536 SNPs selected by the customer
- Focused-content SNP Genotyping with GoldenGate technology (Illumina):
  - LinkageIVb Panel: 5,861 SNPs distributed evenly across the human genome
  - Mouse LD Linkage Panel: 377 SNPs optimized for N2 and F2 mouse genetics crosses
  - Mouse MD Linkage Panel: 1,449 SNPs distributed evenly across the mouse genome
  - MHC Mapping Panel: 1,239 SNPs distributed evenly across the MHC region
  - Cancer SNP Panel: 1,421 SNPs in cancer genes
  - DNA Test Panel: 360 SNPs that may be used as genomic controls.
  - Methylation Cancer Panel I (NEW SERVICE): 1,505 CpG loci selected from 807 genes

Whole-genome genotyping with Infinium technology (Illumina):
- HumanHap650Y BeadChip: 555,000 tag SNP markers derived from the International HapMap Project (phase II) plus 100,000 Yoruba-specific tag SNPs
- Human 1M BeadChip (NEW SERVICE): over one million markers for SNPs and CNV throughout the genome

All services are integrated with robust software tools for experimental design, management of data and analyses. Extensive quality control measures (both human and computational) let us further refine the quality of data.

During 2007, the Unit has extracted approximately 4,600 DNA samples from blood, and has produced more than 2 million genotypes with SNPlex, 0.6 million genotypes through GoldenGate (Illumina) and 70 million genotypes through Infinium (Illumina).

COLLABORATIVE RESEARCH PROJECTS

Our research is focused on the identification of susceptibility genes and the interacting environmental exposures that contribute to genetically complex disorders/trait such as postpartum depression, anorexia, attention deficit/hyperactivity disorder, asthma, psoriasis, and behavioural and cognitive development in early childhood. Methodologies include traditional linkage analysis, sib-pair and affected-pedigree-member methods, case-control or family-based association studies, genome-wide scans and candidate gene analysis.

Topics under investigation through collaborations with other institutions are:

- Identification of the genetic basis of the variability in the response to methadone maintenance treatment (M Torrens, IMIM)
- Search for genetic variants involved in smoking cessation (JM Argimon, Fundació Gol i Gorina)
- Study of genes involved in the circadian rhythm in major depression (M Urretavizcaya, CSUB)
- Study of genes involved in environmental-asthma interactions in a European population (M Kogevinas, CREAL)
Health impacts of long-term exposure to disinfection by-products in drinking water (Mark J Nieuwenhuijsen, CREAL)

Genetic basis of attention deficit/hyperactivity disorder (Miquel Casas, Hospital Universitari Vall d’Hebron, Amaia Hervàs, Mútua de Terrassa and Bru Cormand, Universitat de Barcelona)

**PUBLICATIONS**


The Bioinformatics and Genomics programme includes a number of research groups, which use computational analysis to address relevant questions in genome research. Currently, the programme has two active groups in Genome Bioinformatics, lead by Roderic Guigó and Comparative Bioinformatics, lead by Cedric Notredame. The programme is expected to grow with two additional groups during year 2008.

The two groups have been very active during the past year. A number of solid collaborations have been established between these groups and several experimental groups from other CRG programmes. The program is also taken active leadership in the delineation of the newly established CRG genomic unit, which will provide service in microarrays, genotyping, sequencing and bioinformatics, as well as in the design and implementation of the computing scientific network.

Current structure of the programme:

- 2 Research Groups:
  - Genome Bioinformatics (Roderic Guigó)

- Associated Core Facility: Microarray Unit
Research in the Genome Bioinformatics group focuses on the development and application of methods to identify functional domains in genomic sequences, with emphasis on 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 is involved in NIH-funded ENCODE project.
RESEARCH PROJECTS

1. Gene Prediction

We are working in the development of geneid, an “ab initio” gene prediction program, and sgp a comparative gene finder. Geneid and sgp have been used in the annotation of many eukaryotic genomes. We have been collaborating with Genoscope, and with the Broad Institute, among other institutions, in the annotation of many eukaryotic genomes, and have participated in the analysis and annotation of the genomes of the 12 drosophila species recently sequenced. We have also been collaborating with the group of Stylianos Antonarakis, from the University of Geneva, in the analysis of the first sequences obtained in the heterochromatic fraction of the human genome.

2. Prediction of Selenoproteins

Particularly difficult in eukaryotic genomes is the prediction of selenoprotein genes, because selenocysteine is specified by the UGA codon, normally an stop codon. Since year 2000 we have been developing computational methods for selenoprotein prediction. During the last years we have successfully used this methods to characterize mammalian selenoproteins (Castellano et al., 2001, Kryukov et al., 2003, Castellano et al., 2004). Recently, using comparative genomics methods we have discovered a novel selenoprotein families, whose phylogenetic distribution is challenging long standing assumptions about the taxonomic distribution of eukaryotic selenoproteins (Castellano et al., 2006, see also figure 1). Through the comparative analysis of the 12 drosophila genomes, we have been able to identify the first animals lacking selenoprotein genes.

3. Splicing

In strong collaboration with the group of Juan Valcárcel, from the CRG’s Gene Regulation programs we are investigating the mechanisms by means of which splice signals are recognized and processed. We are developing new methods to infer sequences that may play a role in the regulation of alternative splicing, and have been investigating the dynamics of the evolution of U12 introns. We have developed a database of U12 introns (Alioto, 2007).

4. Recognition of Promoter Regions

In collaboration with Xavier Messeguer from the Universitat Politècnica de Catalunya, we have initiated a research line on algorithmics for promoter recognition; in particular addressing the problem of comparing and characterizing the promoter regions of genes with similar expression patterns. This remains a challenging problem in sequence analysis, because often the promoter regions of co-expressed genes do not show discernible sequence conservation. In our approach, thus, we have not directly compared the nucleotide sequence of promot-
ers. Instead, we have obtained predictions of transcription factor binding sites, annotated the predicted sites with the labels of the corresponding binding factors, and aligned the resulting sequences of labels—to which we refer here as transcription factor maps (TF-maps, see Figure 2). To obtain the global pairwise alignment of two TF-maps, we have adapted an algorithm initially developed to align restriction enzyme maps. We have optimized the parameters of the algorithm in a small, but well-curated, collection of human–mouse orthologous gene pairs. Results in this dataset, as well as in an independent much larger dataset from the CISRED database, indicate that TF-map alignments are able to uncover conserved regulatory elements, which cannot be detected by the typical sequence alignments (Blanco et al., 2006).

5. ENCODE Project

The National Human Genome Research Institute (NHGRI) launched a public research consortium named ENCODE, the Encyclopedia Of DNA Elements, in September 2003, to carry out a project to identify all functional elements in the human genome sequence (The ENCODE consortium, 2004). In its pilot phase the project is aiming to characterize all functional elements in 1% of the human genome. Within ENCODE, we are leading the GENCODE consortium with the goal of identifying all protein coding genes in the ENCODE regions. A first version of this annotation was released in early 2005, and we organized EGASP a community experiment to assess the quality of this annotation. An special issue of Genome Biology has been published during year 2006 devoted specifically to EGASP (Reese and Guigó, eds).

Within the framework of the ENCODE project, we have been collaborating with the groups
of Stylianos Antonarakis, from the University of Geneva, and Tom Gingeras, from Affymetrix, to exhaustively characterize the transcript diversity of protein coding loci. Towards that end, we have developed the RACEarray strategy. In such strategy, RACE products originated from primers anchored in exons from annotated protein coding genes are hybridized into high density genome tiling arrays, and sites of transcription specifically linked to the index exon are in this way uncovered. Such experiments are revealing as many novel exons as annotated ones (Denoeud et al., 2007). We have further used this strategy to “normalize” the relative abundance of alternative splicing transcripts prior to random clone selection for sequencing (see Figure 2).

6. Nucleosome positioning and Chromatin Structure

The search of sequence prerequisites for nucleosome positioning has been a longstanding 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.

PUBLICATIONS


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. All the applications related to our work are provided to the community through an international network of web servers that can be accessed from www.tcoffee.org
RESEARCH PROJECTS

1. Development of the multiple sequence alignment package T-Coffee

T-Coffee is a popular multiple sequence alignment package able to accurately align biological sequences (proteins and nucleic acids). The main specificity of T-Coffee is its ability to combine the output of alternative programs into one unique alignment. T-Coffee has been consistently shown to be among the most accurate multiple sequence alignment packages for protein sequence alignments. In collaboration with the group of Des Higgins, in Dublin University College, we have recently introduced and validated a novel mode of T-Coffee named M-Coffee and able to combine the output of the main packages in order to generate multiple sequence alignments that are among the most accurate currently available (Moretti, Armougom et al. 2007).

2. Development of template based multiple sequence alignments

Over the last years we have pioneered the development of a novel kind of multiple sequence alignments known as template based sequence alignments. This concept makes a heavy usage of the underlying framework of T-Coffee, its principle is relatively new and marks 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 database search that enriches the information content of the sequence. For instance, one can associate the provided sequences with homologous structures, thus making it possible to produce a structure based alignment. In a similar way it is possible to combine available sequences with some profile information. This approach makes it possible to couple all sort of sequence comparison techniques and unify them within a unique multiple sequence framework. We have recently shown that the resulting method is one of the most accurate of its kind, especially when using structural information.

3. Comparative Genomics

In collaboration the group of Phillip Bucher, we have initiated a project for the comparative analysis of Highly Conserved Non Coding Elements (HCNEs) in insects and vertebrates. Through these analyses we have shown that the rates of evolutions were different in these two groups. This projects forms a primer for further study of the comparative evolution of genomic features, it also constitutes a first step toward the development of multiple genome comparisons(Retelska, Beaudog et al. 2007).

4. Multiple Comparison of Life trajectories

Methods developed in bioinformatics for sequence analysis can in theory be applied to any type of strings exhibiting properties similar to those encountered in biological sequences. These may include strings resulting from experimental observation acquired through continuous monitoring of a phenomenon developing across time through a limited number of alternative states. In order to test this hypothesis we have collaborated with a group of sociologist from Lausanne University specialized in the compilation of life trajectories through longitudinal studies. Under this protocol, trajectories are coded as sequences made of pre-defined alphabet indicating various
situations. We have used these sequences to train substitution matrices similar to those used in bioinformatics (Gauthier, Widmer et al. 2007). The purpose was to quantify empirically the level of similarity between various social situations. We expect this methodology to be very general and suitable to any kind of longitudinal studies, including the analysis of longitudinal phenotypes in animal model systems.

5. Multiple Sequence Alignment Server

Thanks to a partnership with the Swiss Institute of Bioinformatics, we are currently running a very powerful multiple alignment server that provides access to the most sophisticated features of T-Coffee. The server is mirrored on several sites, including the CNRS, the European Bioinformatics Institute, Cornell University and the Spanish EMBNet Node. Altogether all the mirrors receive close to 20,000 hits a month originating from an average of 90 countries.

PUBLICATIONS


Retelska D, Beaudoing E et al. Vertebrate conserved non coding DNA regions have a high persistence length and a short persistence time. BMC Genomics, 8(1): 398 (2007). (*)

(*) All these publications are the result of the work of Dr. Cedric Notredame at the Institut de Biologie Structurale et Microbiologie (IBSM), in Marseille, France.
The group is interested in the use of global genomic analysis tools to discover target and co-regulated genes affected under specific conditions, to understand the function of these newly characterized genes through inference from gene expression profile data, to correlate these with changes in transcription factor occupancy at gene promoter and chromatin modifications, to understand large scale copy number variation through comparative genomic hybridization, and more recently, to study the expression of small non-coding RNA, and of miRNAs in particular, and to develop novel experimental strategies to characterize their mRNA targets.

The main focus of our research is on technological development and validation of microarray based assays both at the experimental as well as the bioinformatic analysis level, with the goal to maintain a state of the art core facility, the Microarray Unit. Part of our activities are centred on comparison across platforms. With the advent of new methodologies such as next generation high throughput sequencing, we are exploring the performance of this new technology in expression profiling of mRNA and miRNA compared to array based methods.

GROUP STRUCTURE:

Group Leader: Lauro Sumoy
Postdoctoral Fellow: Franc Llorens
RESEARCH PROJECTS

1. Microarray benchmarking, process quality control, standardization and meta-analysis

One of the main current challenges in microarray research is the comparison between different data sets. We intend to advance in the development of methods that allow inference of conserved patterns of gene co-regulation using meta-analysis methodologies. We plan to apply these to understand the signalling pathways affected by growth factors using cell culture models.

In addition, we plan to apply these novel developments to the comparison and mining of microarray datasets in public databases. Through participation in several large scale projects involved in the study of processes such as hormone response, chromatin remodelling, or diseases such as cancer or genomic disorders, we are developing new approaches to the study of gene expression profiles by use of cross-platform standardization, meta-analysis and multivariate methods. This should allow us to integrate already published datasets and information derived from our own microarray experiments.

2. Microarray technology development

Many of our technological research activities are focused on implementing new experimental applications. Many of these arise from the involvement of the microarray core facility in many different collaborative projects. We are committed to developing and optimizing procedures for the design, fabrication, hybridization, processing and analysis of data generated from DNA microarrays.

Our immediate plans are to explore three new technologies:
- miRNA expression profiling
- array based targeted sequence enrichment for ultrasequencing applications
- identification of miRNA targets

3. Microarray bioinformatics

We have set up automated image data acquisition, pre-processing, filtering, normalization and quality control Web based software for analysis of microarray experiments. We are implementing an Oracle based database for direct data browsing by service users. This includes capabilities for analysis of two colour (cDNA and BAC spotted arrays, Agilent, Exiqon) and single channel (Affymetrix, Illumina) microarray data, including pre-processing, normalization, statistical analysis, functional analysis, chromosome position visualization and linking to genome browsers.

PUBLICATIONS


As a core facility, the laboratory is mainly responsible for the experimental and bioinformatics aspects of different research projects that use microarrays. In addition, it provides microarray methodologies as a service at established rates to scientists from the CRG, PRBB (UPF, IMIM, CREAL, CMRB, CAT) and other external public and private institutions.

UNIT STRUCTURE:
- Senior Technician: Anna Ferrer
- Technician: Heidi Mattlin
- Bioinformatician: Xavier Pastor

Guest or associated members:
- Bioinformatician: Pau Rué
  (Consolider, M. Beato - CRG)
SERVICES

Services offered include: microarray probe selection and design, microarray probe preparation, microarray fabrication through contact spotting, RNA purification, quality control and amplification, RNA and DNA sample labelling, hybridization of microarrays and data processing and analysis. We have already used microarrays to study gene expression in different species (yeast, human, rat, mouse, maize, cork and Mycobacterium) and customized targeted applications (pancreas, breast cancer and neural specific expression arrays, splicing arrays, antibody arrays for cytokine and signal transduction protein detection). We also have used custom BAC arrays and oligonucleotide arrays for comparative genomic hybridisation on specific chromosomes, copy number variation regions and whole genomes. More recently we have tested promoter arrays for chromatin immunoprecipitation on microarrays, including transcription factor, histone modification and promoter activity profiling through RNApol-II as well as micrococcal nuclease based nucleosome positioning assays. The facility is set up for optimal processing of in situ synthesized long oligonucleotide arrays (Agilent, Illumina and Exiqon).

The Microarray Unit offers applications including:

- mRNA expression profiling on spotted cDNA and Agilent microarrays, and Illumina bead arrays.
- miRNA expression profiling on Exiqon and Agilent microarrays.
- array based comparative genomic hybridization (aCGH) on spotted BAC and Agilent microarrays.
- chromatin immunoprecipitation on array (CHIP on chip) on spotted promoter, CpG island and Agilent microarrays.

PUBLICATIONS BY CORE FACILITY USERS


The programme of Cell and Developmental Biology is rapidly growing to its full capacity. This year I left my position as professor of Biology at UC San Diego, California and relocated to CRG as the coordinator of cell and developmental biology programme. We have succeeded in recruiting a new group leader, Manuel Mendoza from ETH, Zurich, who will arrive to the CRG in 2008. Manuel is specifically interested in the mechanism by which defects in spindle midzone activate a checkpoint called NoCut. NoCut insures cytokinesis occurs only after migration of all chromosomes to the poles. The arrival of Manuel will strengthen our overall interest in regulation of cell compartmentation and division. Two additional group leaders will be recruited to compliment the existing expertise in the program during 2008.

Timo Zimmermann arrived from the EMBL, Heidelberg to head the Advanced Light Microscopy Unit (ALMU) at CRG in March 2007. ALMU provides a range of instruments for high-end light microscopy for CRG researchers and their diverse research fields, as well as training and support in imaging applications. Two staff members organize training classes and maintain equipment. Two confocal microscopes and a completely automated wide field microscope with microinjection capability are now up and running. Additional instruments will be set up in 2008 to cover the increasing need for imaging resources in the growing institute.

Hernan-Lopez Schier was awarded the prestigious ERC grant to work on organogenesis of sensory systems.

Structure of the Programme:

Coordinator: Vivek Malhotra
Senior Group: Isabelle Vernos
Junior Group: Hernan Lopez-Schier
Advanced Light Microscopy Facility: Timo Zimmermann
We are interested in two very basic questions. One, the formation of transport carriers with specific emphasis on the process of membrane fission, which is required to separate cargo filled transport carriers from the maternal compartment. Two, partitioning of Golgi membranes into daughter cells during cell division and the regulation of mitotic entry by Golgi organization.
RESEARCH PROJECTS

1. Membrane fission
Yuichi Kawana, Laetitia Casano, 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 \( \beta \gamma \). This was the first clear documentation of the involvement of trimeric G proteins in membrane fission (Jamora et al., Cell. 1997).

Soon thereafter we found that \( \beta \gamma \) 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). DAG was known to be

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.
required for protein secretion in yeast but
our findings placed its requirement specifi-
cally in the events leading to membrane fis-
sion. Soon thereafter, we realized that
there were three different isoforms of PKD
in the mammalian cells and surprisingly, all
were required for Golgi to cell surface
transport (Yeaman et al., Nature Cell
Biology. 2004). Moreover, all three forms
were found to be specific for the trafficking
of only those proteins that were destined to
the basolateral cell surface. Apically tar-
geted proteins did 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) and the formation of the
dimers do not require DAG (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 through a 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
Pfizenmeir and colleagues, Germany). Ceramide can be used to generate DAG and
we suggest that this event occurs once the
transport carriers are formed to prevent
further production of DAG. This might be
necessary to prevent uncontrolled produc-
tion of vesicles (Bard and Malhotra, Ann Rev
Cell and Dev Biol.2006). We have identified
a novel protein called Yusukin (after a post-
doc who identified it), which is a substrate
of PKD and required for Golgi to cell surface
transport. Clearly, a number of PKD sub-
strates are now in hand but the only way
forward to a mechanicistic understanding of
membrane fission is to reconstitute this
process with pure components. Purified
Golgi membranes are being used to reconsti-
tute the formation of cargo filled transport
carriers in vitro with purified recombinant
proteins describe above. The aim is to mon-
itor the assembly of proteins mentioned
above on Golgi membranes, identify hither-
to unknown components and measure
changes in lipid composition in the events
leading to membrane fission. In sum, our
relatively unorthodox approach to use a
chemical to vesiculate Golgi membranes is
revealing the mechanism of membrane fis-
sion during protein transport.

2. New transport components identified
Kota Saito

A genome wide screen was carried out to
identify new components required for pro-
tein secretion. From the 22,000 genes test-
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 gene 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.

3. Unconventional secretion.
Juan Duran

Some of the interleukins, fibroblast growth factors (FGF), 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 discoidium 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). Mammalian cells contain 2 GRASPs, GRASP55 and GRASP65. We are testing the involvement of these proteins in unconventional secretion in mammalian cells.

4. Inheritance of Golgi membranes during cell division.
Gianni Guizzunti

How are Golgi membranes partitioned into daughter cells during cell division? 3 major findings from our laboratory on this issue of fundamental importance follow. 1, The Golgi membranes retain their identity throughout the cell cycle and do not relocate to the ER their delivery into daughter cells (Pecot and Malhotra, Cell 2004 and 2006). 2, Golgi membranes undergo a change in their overall organization, which is mediated by a specialized MAPK pathway (Acharya et al., Cell. 1998) and polo-like kinase (Sutterlin et al., PNAS. 2001). 3, A change in golgi organization is required for entry of cell into mitosis (Sutterlin et al., Cell. 2004). In vitro assay reconstituting Golgi fragmentation by mitotic cytosol and system wide siRNA are being used to identify new components involved in regulating Golgi organization. We are particularly interested in components involved in connecting Golgi membranes with the centrioles, the substrates of MAPK and polo-like kinase on the Golgi membranes and the mechanism regulating the shape and size of Golgi cisternae.

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


* All these publications are the result of the work of Dr. Vivek Malhotra at the University of California in San Diego, La Jolla, USA.
ISABELLE VERNOS HAS AN SENIOR ICREA GROUP LEADER POSITION.

GROUP STRUCTURE:

Group Leader: Isabelle Vernos

Postdoctoral Fellows: Teresa Sardon Sylvain Meunier Roser Pinyol (since April 2007)

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

Technician: Luis Bejarano María Sanz

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 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. Pathways for microtubule assembly during M-phase

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. Concomitantly after nuclear envelope breakdown, a centrosome independent pathway relying on a RanGTP gradient triggers MT nucleation in the vicinity of the condensed chromatin. We want to understand the molecular mechanisms underlying these two pathways and how they are regulated during the cell cycle.

- In the context of a collaborative project involving several research groups (Centrosome 3D) we have started functional and structural studies on the centrosome.
- Relatively little is known on the molecular mechanism underlying the RanGTP dependent pathway for MT nucleation and stabilization. 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). To unravel the molecular mechanism underlying the RanGTP dependent pathway for MT assembly in M-phase, we continue our studies on TPX2 and its interaction partners and we aim at identifying the other RanGTP regulated factors involved in the same pathway.
- We have previously shown that the Xenopus TACC family member, Maskin, plays an essential role for microtubule growth from the centrosomes during M-phase and that both its localization and function are regulated by phosphorylation by the Aurora A kinase. We also obtained some evidence indicating that Maskin works in concert with XMAP215 to oppose the destabilizing activity of

![Figure 1: Immunofluorescence images of Xenopus tissue culture cells showing Aurora A (in yellow) strongly localized to the centrosome in different cell cycle stages: prophase (left), pro-metaphase (middle) and metaphase (right). Microtubules are in red and DNA in blue.](image)
We are studying the putative role of Maskin in the RanGTP dependent microtubule assembly pathway and the functional implications of its interaction with the Aurora A kinase.

2. Regulation of spindle assembly by the kinase Aurora A

The Aurora A kinase has been implicated in several important processes including centrosome maturation during G2, mitotic entry, centrosome separation and bipolar spindle assembly. However its precise role is still relatively unclear. We have previously shown that Aurora A interacts with TPX2 in a RanGTP dependent manner resulting in kinase activation and that it interacts with Maskin and regulates its function through phosphorylation. Using the egg extract system we have studied the function of Aurora A during spindle assembly by dissecting the different pathways involved in MT nucleation and stabilization. We have found that Aurora A work through different mechanisms to regulate MT assembly ensuring bipolar spindle formation.

3. Role of molecular motors during M-phase: functional studies on the chromosome-associated microtubule dependent motors Xkid

As the mitotic spindle starts to form, chromosomes establish dynamic interactions with the microtubules. These interactions play an active role in spindle formation and are responsible for the movement of chromosomes that lead to their alignment on the metaphase plate and their segregation during anaphase. Some of these interactions are mediated by chromokinesins, kinesin-like proteins that localize to the chromosome arms during M-phase. We have previously identified two of them in the Xenopus system: Xklp1 (Vernos et al., 1995; Walczak et al., 1998) and Xkid. Xkid is required for chromosome alignment on the metaphase plate (Figure 2) (Antonio et al., 2000) and plays a role in cell cycle progression during Xenopus oocyte meiotic maturation (Perez et al, 2002). We are currently performing experiments in oocytes and egg extract to try to understand the role of this molecular motor in meiotic cell cycle progression and its regulation by phosphorylation.

Figure 2: Xkid localizes along chromosome arms during M-phase. In Xenopus egg extracts chromosomes do not align on the metaphase plate in the absence of Xkid.
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 underlying the acquisition and maintenance of tissue architecture, and its relationship to the function of sensory organs. We use the mechanosensory lateral line of the zebrafish (Danio rerio) (Figure 1) 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.
RESEARCH PROJECTS

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

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 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.

2. Sensory organ growth and regeneration, with an emphasis on epithelial remodelling and 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. Extensive genetic and molecular studies in Drosophila have shown that the establishment of planar polarity relies on the concert ed activity of many proteins. The cellular responses to polarity cues, especially in...
remodelling tissues are not understood. We are trying to define 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.

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 (Figure 3).

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.

3. Sensory perception and integration, and sensory dysfunction

The zebrafish has relatively acute hearing, which the animal uses to school, capture prey and evade predators. We have identified a mutation that leads to profound deafness in the zebrafish. Hair cells develop normally and survive in adult homozygous mutant fish, but they are unable to transduce mechanical stimuli. This mutation appears to affect a very specific component of the mechanotransduction machinery in hair cells. We have recently identified the affected gene and begun to decipher its biological function in hearing and balance (unpublished results). We are also attempting to understand the mechanisms that govern the communication between sensory organs and the central nervous system in vertebrates using genetic, molecular and cell biological approaches. The knowledge gathered with these initial studies should permit the rational design of tests to evaluate the contributions of the relevant molecules, and will ultimately allow further progress into understanding how an animal responds to complex environmental stimuli.
CELL AND DEVELOPMENTAL BIOLOGY

Associated Core Facility: Advanced Light Microscopy Unit

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 automated screening over 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 capabilities of advanced light microscopy at the organismic, cellular and molecular level. Methods available in the facility include optical sectioning, spectral imaging, in-vivo timelapse imaging, Fluorescence Resonance Energy Transfer (FRET) detection, Fluorescence Recovery After Photobleaching (FRAP) as well as image analysis, particle tracking and 3D rendering.

Within its first year of operation, the ALMU has already become strongly used by researchers from all CRG programs and further instruments will be acquired to complement the existing ones and to meet the still increasing need for imaging.
Based on the preparations by the CDB acting coordinator Isabelle Vernos in the previous years, the Advanced Light Microscopy Unit started operation in March 2007 after an initial phase of refurbishments and installations during January and February. The microscopy room was completely refurbished to accommodate the microscope systems. All necessary supply systems (electricity, ventilation, CO$_2$ and pressurized air) were installed and the room was divided into six independent microscopy cubicles. In the entrance area, a bench space with fridge, freezer, centrifuge and incubator was set up for sample handling.

In 2007, two confocal microscopes were installed, one (Leica TCS SPE) for standard immunofluorescence applications, one (Leica TCS SP5) for high-end applications like Fluorescence Recovery after Photobleaching (FRAP) or Fluorescence Resonance Energy Transfer (FRET). The SP5 system was additionally fitted with a custom-made microscope incubation system for full environmental control for in-vivo imaging.

A fully automatized high-speed fluorescence microscope (Zeiss Cell Observer HS) was set up and equipped with an Eppendorf microinjection system, environmental control and an additional EM-CCD camera for low-light imaging applications.

All systems were co-financed through FEDER equipment grants from the Spanish government.

Additionally small equipment like special objectives, a beampath inverter and a unit for fast temperature changes were acquired and can be used on several of the systems.

All systems are used regularly by CRG researchers after a mandatory training session on the instrument. System booking is done through the online reservation system.

Usage times and the number of users have continued to increase throughout the year. From the usage data of 2007, a yearly usage of 10,000 hours can be projected for the three initial microscope systems, which translates into an average daily use of more than nine hours per system. By December 2007, the facility had 45 users from CRG research groups and additionally researchers from two external groups in Barcelona and three visiting scientists from

**Figure 1:** Current microscopy equipment of the Advanced Light Microscopy Unit (ALMU), covering confocal and widefield fluorescence imaging as well as sample preparation and image processing.
France and Germany. A significant proportion of the researchers uses not only one but several of the available systems for their experiments. The in-house users are evenly distributed over all CRG programs with the exception of the Bioinformatics Program. Usage hours vary between programs according to the need for time-intensive in-vivo timelapse experiments.

In addition to the installed systems, several microscope systems (e.g. for screening microscopy and for fast in-vivo confocal imaging) have been evaluated in-house in the last year and grant applications for additional systems have been approved or are currently under way.

The evaluation and installation of systems will continue into 2008 and will be completed at the end of the year.

Figure 2: Distance dependence of Fluorescence Resonance Energy Transfer (FRET) interactions. Suitable Donor and Acceptor molecules have to be within nanometers of each other for FRET to occur. Due to this distance specificity FRET can be used to study molecular interactions.
PUBLICATIONS


(*) The contribution to this publication is the result of the work of Dr. Timo Zimmermann at the EMBL, Heidelberg, Germany.
In 2006, we recruited our first three group leaders: James Sharpe, Ben Lehner and Mark Isalan. Together they cover modelization, experimental analysis of developmental systems and the engineering and design of new gene networks (see below). In 2007, we hired Matthieu Louis who incorporated to the CRG in December 2007. Finally, in October 2007, we interviewed new candidates for the position of Junior Group Leader and in January 2008, we have made an offer to one of them. In March 2008, if the hiring is successful, the programme will be full.

All the members of the programme have been very successful in securing external funding and in publishing their work in top journals. Significant advances have taken place in the research lines of the different groups. Thus, we will define this year as a year of consolidation.

Structure of the Programme:

Coordinator: Luis Serrano
Senior Group: James Sharpe, Ben Lehner, Mark Isalan
Junior Groups: Matthieu Louis (starting in January 2008)

Systems Manager: Yann Dublanche
Responsible of Equipment: Raul Gomez
Grant Manager: Michela Bertero
The group of **Luis Serrano** is aiming at a quantitative understanding of biological systems to an extent that would enable prediction of systemic features and with the hope to reach rational design and modify their behaviour. This applies to any system comprising biological components that is more than the mere sum of its components, or, in other words, the addition of the individual components results in systemic properties that could not be predicted by considering the components individually. By achieving this objective the group aims at new global understanding and treatment of human diseases in which the target will not be a single molecule but a network. For this purpose the group on one hand develops new software and theoretical approximations to understand complex systems and on the other performs experiments to validate the predictions.

**GROUP STRUCTURE:**

<table>
<thead>
<tr>
<th><strong>Group Leader:</strong></th>
<th><strong>Luis Serrano</strong></th>
</tr>
</thead>
</table>
| **Staff Scientist:** | **Christina Kiel**  
|                  | **François Stricher** |
| **Postdoctoral Fellows:** | **Tobias Maier**  
|                   | **Raik Grünberg**  
|                   | **Eva Yus**  
|                   | **Almer van der Sloot**  
|                   | **Emanuele Raineri**  
|                   | **Patrick Herde**  
|                   | **Paolo Ribeca**  
|                   | **Vicente Tur**  
|                   | **Alejandro Nadra**  
|                   | **Andreu Alibés** |
| **Students:** | **Marc Güell**  
|               | **Ronan Burgeois** |
| **Technician:** | **Justine Leigh**  
|                | **Sira Martinez** |
RESEARCH PROJECTS

1. Quantitative understanding of M. pneumoniae

This proposal has two main objectives: a) the first complete global understanding of a living organism (M. pneumoniae) and b) its engineering to create a shuttle vector that could be modified InVitro and then transfected in human cells incorporating itself as a new cell organelle. Once inside the cell, the living vector should couple its division to that of the cell host and be able to deliver proteins in the cell to correct human pathologies. M. pneumoniae is one of the smallest free-living bacteria that exist (~680 genes), does not have cell wall and can be introduced inside human cells. The project will benefit from a structural genomics project aimed at determining the structures of all M. pneumoniae proteins (strgen.org/proteome/), from existing extensive previous work (www.zmbh.uni-heidelberg.de/M_pneumoniae/genome/) and from the participation in the context of the CRG/EMBL partnerhips of a consortium involving the EMBL structural and computational biology programme. Currently all proteins of the organism are been tagged for pull down to identify all the protein complexes, electron microscopy is done on these complexes to obtain 3D shapes which will be fitted inside EM 3D tomograms of the bacteria at 40-50 A resolution.

In our group during 2007 we have done: a) analyzed the whole proteome in collaboration with Mathias Mann (MPI-Munich). b) designed and used DNA arrays to analyze the expression of different genes under different conditions. c) Finished a metabolic map of Mycoplasma and based on it design a defined medium to grow the bug d) Developed homologous recombination e) Developed a plasmid.

While we compile all this information we have started to modify the organism in order to transfect human cell lines, engineer a regulation system of its division that will couple it to that of the host cells and develop a system able to secrete proteins inside the host.

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.

During 2007, we have designed new variants of TRAIL specific for DR4, which are now under patenting. In addition, we have started to create new specific variants of other members of the TNF family. .

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. 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. Here, we want to design proteins and peptide variants, which have a strong selectivity for one protein or receptor.

We will select different signalling pathways, focusing on surface receptor mediated signalling cascades, like the EGF/Ras pathway and TNF signalling. Based on this, specific signal transduction pathways can be studied in great detail. Further, it opens the possibility to study the cellular effect of changing biophysical properties, like association and dissociation rate constants. Designing new small peptide molecules will allow on one hand to specifically activate one particular route in a pathway (i.e. Ras-Raf), but also could offer new drugs based on peptidomimetics and modified proteins (i.e. TRAIL variants). While so far approaches in drug design mainly aim to specifically inhibit a protein interaction, which in many cases also inhibits other protein complexes, we here want to design protein and peptide variants, which specifically and selectively bind to one target protein/receptor.

During 2007, we have analyzed signal transduction pathways in vision in collaboration with the group of Marius Ueffing and Gianni Cesaretti, using structural information to dissect it. In addition, we have studied the importance of kinetics in signal transduction pathways using RAF mutants designed to have the same Kd but with different Kon and Koff values.

**PUBLICATIONS**


Pey AL, Stricher F, Serrano L and Martinez, A. Predicted effects of missense mutations


(*) All these publications are the result of the work of Dr. Luis Serrano at the EMBL, Heidelberg, Germany

**PATENTS**

<table>
<thead>
<tr>
<th>Application Nr.</th>
<th>GB 0723059.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>UK priority filing</td>
</tr>
<tr>
<td>Title</td>
<td>Improved cytokine design</td>
</tr>
<tr>
<td>Priority/Filing date</td>
<td>23/11/2007</td>
</tr>
<tr>
<td>Applicants</td>
<td>NUIG / RuG / CRG / Triskel Therapeutics</td>
</tr>
<tr>
<td>CRG Inventors</td>
<td>Luis Serrano / Almer van der Sloot / Vicente Tur</td>
</tr>
</tbody>
</table>
Our lab has 2 primary goals: (1) To further our understanding of developmental biology 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 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.
RESEARCH PROJECTS

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

Since inventing OPT has already proven itself in a number of research fields, it is still a “new” technology and probably has not yet achieved its maximal potential. We are therefore continuing to develop OPT both in terms of hardware and software. Additionally, since the interests of the lab focus on small biological specimens (a few millimetres in diameter, sometimes labelled as “mesoscopic” imaging) we are exploring and developing other 3D imaging principles that could provide extra information for understanding embryo development.

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 is the ability, for the first time, to quantify the number of Islets of Langerhans in an intact adult mouse pancreas in a single scan (see figure above). 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, and at the beginning of 2007 we published a report (in collaboration with the lab of Dr. Ulf Ahlgren in Umea, Sweden) on this project in *Nature Methods* (4:31-33, 2007).

3. 4D Time-Lapse Optical Projection Tomography

Optical projection tomography has proven to be a powerful tool for developmental biologists. But so far OPT has been performed almost exclusively on ex-vivo specimens which have been optically cleared to increase the quality of 3D images obtained. Building on the success of this new technology we are now exploring various improvements to allow the 4D imaging of the developing mouse limb bud in culture. This technique now provides us with some of the data on tissue movements that is essential for understanding embryo development.

![Figure 1: Top](image1.png) We have worked closely with the MRC to design a commercial version of the scanner. **Bottom row:** OPT scans of a whole adult mouse pancreas (left) which has been fluorescently-labelled with antibodies against insulin to highlight the 3D distribution of Islets of Langerhans (centre). This data can be morphometrically quantified to provide statistics on the volumetric size distribution (right) - *Nature Methods* (2007) 4:31-33.
subsequent computer modelling. In the near future it should also provide data on dynamic gene expression patterns.

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 aim to explore 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.

Figure 2: The finite-element model used for simulating limb development. The 3D shape is represented by a tetrahedral mesh (left and centre), and different proliferation rates therefore cause a global movement of the tissue (right).
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 ex-vivo analysis. However, the spatial patterns of certain genes (e.g., Sox9) are extremely dynamic – apparently changing hour-by-hour. Knowing the age of a given limb bud should therefore be very important when comparing results, but the field has not had a convenient but accurate method for determining the stage of a limb.

Our lab has 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. Researchers with jpg images of their limb buds will be able to log-on to our web-site and determine an accurate age of the specimen within a few minutes.

PUBLICATIONS


(*) This publication is the result of the work of Dr. James Sharpe at the MRC Human Genetics Unit, Edinburgh, UK.
We are interested in engineering synthetic gene networks to control gene expression in cells and to construct self-organising patterns, analogous to those used by organisms in morphogenesis and development. By transfecting cell populations with various gene networks, we hope to find the ‘design principles’ underlying why certain networks form particular structures or functions. We are exploiting this information to deliver genetic programs into cells to make them differentiate in the ways we desire.

The Group is divided into two subgroups, one of which is dedicated to making artificial DNA-binding domains and the other which employs these technologies, and related ones, to synthesise artificial gene networks.
RESEARCH PROJECTS

Mireia Garriga, Phil Sanders, Marco Constante and Andreia Carvalho

We are looking at methods of designing self-organising patterns using mammalian cells as an engineering scaffold. To do this we are employing a technique that we recently developed: magnetic beads, coated with PCR products, can be used to transfect cells with gene network constructs, with spatio-temporal control (Nature Methods 2: 113-118 (2005); Nature Protocols 1, 526 - 531 (2006); Fig 1).

This approach of reconstituting systems in order to test our understanding of them should be generally applicable to the study of any biological network with a spatial component.

We are carrying out several pattern-engineering projects in parallel, corresponding

![Figure 1: Our system for magnetically-defined transfection. (a) A GFP-expressing PCR product is directed to cells using the paramagnetic bead as a scaffold. Note the multiple beads adjacent to transfected cells. This technique is being employed to engineer pattern-forming gene networks in eukaryotic cells. The spatial control of transfection can be applied either to individual cells or to many cells in a culture. (b) Spatial transfection of a region of an MDCK cyst with GFP. (c) We have discovered that it is possible to induce single tubules from MDCK cysts using HGF-coated magnetic beads. This is the first step towards engineering a functional spatial pattern-forming gene network in this system.](image)

![Figure 2: Mechanism of homologous gene repair by zinc finger nucleases. In our laboratory, we are designing and synthesising a number of different zinc finger proteins to target unique disease-related genes. By linking the fingers to FokI nuclease, a cut is made in the genome. A gene repair matrix (blue) then recombines into the genome to repair the mutant gene (red).](image)
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.

2. Synthesising zinc fingers for gene therapy and gene repair

Frank Herrmann and Emmanuel Fajardo

As part of two EU-funded projects (Netsensor, EC Contract No. 012948 http://netsensor.embl.de/ and Integra, EC Contract No. FP6 - 29025) 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 will 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 2007, we synthesised a number of functional zinc fingers to recognise the p53 gene, two of which have activity as site-specific nucleases. Each nuclease recognises and cuts different 24 bp regions in the p53 gene sequence (Fig 3). The next steps will involve testing the constructs for gene repair activity in vivo. This would be a useful tool to study cancer because >50% of all human cancers have mutations in p53. The zinc finger technology will allow us to repair or mutate these regions in different cell lines at will.
We are using the methods of systems biology to study two problems:

how cells are reliably programmed to develop different fates during embryogenesis, 
how the genotype of an organism is predictive of its phenotype.

For experimental work our favourite model is the simple multicellular animal C. elegans in which we can use high-throughput experimental approaches as well as real-time quantitative in vivo measurements at the resolution of individual cells. For computational work we use datasets from any organisms that can be used to address the problems we are interested in.
RESEARCH PROJECTS

1. Programming and re-programming cell fates

We are using systematic experimental and computational approaches to understand how the expression of regulatory genes is used to program alternative cell fates during embryogenesis, how these cell states are defined, and how cell fates can be experimentally re-programmed in vivo.

We are also using precise in vivo measurements to understand how cell states are determined robustly (without errors) in the context of noise and varying environmental conditions (e.g. Lehner, 2008), and how regulatory networks that define cell states evolve between species (e.g. Vavouri et al., 2007).

2. Interpreting your personal genome

The ability to cheaply sequence entire genomes is opening up a new era for human genetics. To capitalise on this wonderful opportunity we want to develop predictive methods that can be used to relate how changes in DNA sequence correspond to changes in organism phenotype. For loss-of-function phenotypes this is possible using integrated gene networks (e.g. Lee et al., 2008). However most human genetic variation alters gene expression rather than coding sequence, and we want to understand the effects of these more subtle perturbations (e.g. Semple et al. 2008), with the goal of developing predictive methods. A further important issue is that each of our genomes contains many different mutations, and we need to understand how mutations in multiple genes can combine to produce novel (synthetic) phenotypes (Lehner, 2007; Tischler et al., 2008).

PUBLICATIONS


Semple JI, Vavouri T, Lehner B. A simple principle concerning the robustness of protein complex function to changes in gene expression. BMC Systems Biology, in press.


The VI CRG Symposium was held on 8 and 9 November 2007, at the PRBB Auditorium in Barcelona. This edition was titled “GENOMIC REGULATION: EXECUTING THE CODE”. The ambitious and main purpose of this symposium was covering all relevant aspects of gene regulation including presentations of 18 worldwide experts grouped in 6 sessions. Sessions went from nuclear dynamics analysed through screening techniques, to global gene regulation, chromatin, RNA processing and translational regulation, to end up with small RNAs regulators. This integration strategy offered the attendants the opportunity to get a complete picture to understand what was the situation of genomic regulation at that moment and where the most promising investigation lines were directed.

Together with this general approach and within the framework of a satellite mini-symposium, titled “Kinases to Chromatin”, the objective was going more deeply into a hot topic nowadays, like the function of kinases in gene regulation through its effects on chromatin structure and dynamics. In this emergent field, we were lucky to have renowned experts within the PRBB. All this allowed us to get together the most relevant scientists in this area, who also talked or participated in the general symposium. In all, 25 foreign scientists participated actively in both events, fact that generated a high level discussion.

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 scientific community in the area of Barcelona, as well as for all the attendants coming from the rest of Spain and Europe. This fact shows the dissemination of the symposium was considerably successful, since scientists from different Spanish and European institutes were attracted to attend. On the other hand, it is very important to highlight that contacts amongst several of the speakers’ groups were established, which in some cases have crystallized in collaboration projects.

Sessions were structured by topics, in order to gather similar research groups. The symposium included excellent presentations, as for example the one by Tom Gingeras, from Affymetrix, Inc., in the States; Tony Kouzarides, from The Gurdon Institute, in the United Kingdom; Rick Young, from the Whitehead Institute, in the States; and James Manley, from the Columbia University, in the States.

The symposium was announced at the CRG website and was disseminated to amongst many scientific institutions around the world. Moreover, thanks to the symposium, some of the speakers appeared in the media. Therefore, we do believe the symposium contributed to increase the visibility of the CRG and Barcelona in this field and allowed offering a global panorama of this innovative research area. The high number of attendants (around 200), 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.
Genomic Regulation: Executing the Code

Organizers: Miguel Beato, Ramin Shiёkhattar, Juan Valcárcel

Invited speakers and chairmen:

SHELLEY BERGER
Philadelphia

MVLES BROWN
Boston

PIERRE CHAMBON
Illinois

ROBERT DARNELL
New York

GIDEON DREYFUSS
Philadelphia

ANNE EPHRUSI
Heidelberg

PETER FRASER
Cambridge

TOM GINGEAS
Santa Clara

SNV GREWAL
Bedford

TARIO KOUZARIDES
Cambridge

ANGUS LAMOND
Dundee

JAMES MANLEY
New York

ROBERT MARLIESSSEN
Cold Spring Harbor

IAIN MATTAJ
Heidelberg

DISHRAV PATEL
New York

DANNY REINBERG
New York

JOEL RICHTER
Worcester

RAMIN SHIEKHATTAR
Barcelona

ALI SHLATIFARO
St. Louis

ROB H. SINGHER
New York

VARUNI SONENBERG
Montreal

JUAN VALECAER
Barcelona

RICK YOUNG
Cambridge

Satellite Mini-Symposium
“Kinases signalling to chromatin”

Organizers: Francesc Pousa & Miguel Beato
8 November 2007 (afternoon)

Invited speakers:

GUSTAV AMMERER
Vienna

SHELLEY BERGER
Philadelphia

JAMES R. BAIE
Winnipeg

RAFAEL E. HERRERA
Houston

LOUIS MAHDAVEAN
Oxford

RICK YOUNG
Cambridge

On-line registration at www.crge.es
Centre for Genomic Regulation (CRG), PRBB Building, Dr. Agustador 85, 08033 Barcelona (Spain) • Ph: +34 93 318 01 00
PRBB-CRG SEMINARS 2007

APPENDIX 2

PRBB-CRG SEMINARS

28-09-07
JONATHAN WIDOM
Dept. Biochemistry, Molecular Biology & Cell Biology, Northwestern University, Evanston, Illinois, USA
"The Genomic Code for Nucleosome Positioning"

27-09-07
MARC TIMMERS (PRBB-CRG)
Dept. of Physiological Chemistry, Division Medical Genetics, University Medical Centre - Utrecht, The Netherlands
"TFIID: dynamics and histone methylation"

18-09-07
CHARLES CANTOR (PRBB-CRG)
Dept. of Biomedical Engineering, Boston University, USA // Chief Scientific Officer, Sequenom Inc., San Diego, USA
"Medical Applications of DNA Mass Spectrometry"

17-09-07
THOMAS A. COOPER (PRBB-CRG)
Baylor College of Medicine, Houston, Texas, USA
"Coordinated regulation of alternative splicing during development and its disruption in disease"

14-09-07
TIm BLISS (PRBB-CRG)
Division of Neurophysiology, National Institute for Medical Research, London, UK
"LTP and memory: what next?"

25-07-07
SRIDHAR HANNEHALLI (PRBB-CRG)
Assistant Professor, Penn Center of Bioinformatics, Dept. of Genetics, Philadelphia, USA
"Computational Analysis of Transcriptional Regulation - Signals, Interactions and Networks"

13-07-07
RANDOLF MENZEL (PRBB-CRG)
Freie Universitat Berlin, Institut fur Biologie - Neurobiologie, Berlin, Germany
"Memory traces in the mushroom body of the honeybee"

6-07-07
PETER F. STADLER (PRBB-CRG)
Lehrstuhl fur Bioinformatik, Institut fur Informatik, University of Leipzig, Germany
"The Non-Coding Majority — Bioinformatics of Non-Protein-Coding "Genes"
6-07-07
STEFANO BIFFO (PRBB-CRG)
Associate Prof. of Cell Biology, Univ. of Eastern Piedmont & Lab of Molecular Histology, Milan, Italy
"eIF6 is a rate-limiting controller of ribosome Biogenesis and translation and a mediator of tumorigenesis".

29-06-07
ARTURO C. VERROTTI (PRBB-CRG)
Dipartimento di Biochimica e Biotecnologie Mediche, Universita di Napoli “Federico II” and CEINGE-Biotecnologie Avanzate, Napoli, Italy
"Molecular mechanisms of translational regulation in Drosophila"

20-06-07
STEPHEN J. CHANOCK
Pediatric Oncology Branch, National Cancer Institute
"Genetic susceptibility to breast and prostate cancer: the devil is in the details"

15-06-07
KENNETH D. POSS
Assistant Professor, Cell Biology, Duke University Program in Genetics, Cell & Molecular Biology Program, Duke University Medical Center, Durham, USA
"Mechanisms of organ regeneration in zebrafish"

8-06-07
ALLAN JACOBSON
UMASS Medical School, USA
"Genetic nonsense, from bench to bedside"

1-06-07
RON HAY
School of Life Sciences, University of Dundee, Scotland, UK
"The role of SUMO modification in transcriptional regulation"

30-05-07
LIOR PACHTER
Associate Professor of Mathematics & Computer Science, Dept. of Mathematics, University of California at Berkeley, USA
"From Drosophila and Transposable Elements to the Neighbor-Net Algorithm and Phylogenetic Networks"

25-05-07
JONATHON PINES
Gurdon Institute, Cambridge, United Kingdom
"Getting in and out of mitosis"

18-05-07
MIGUEL ÁNGEL VIDAL
Developmental & Cell Biology, Centro de Investigaciones Biologicas (CSIC), Madrid, Spain
"Polycomb Ring1B and cell proliferation control"

4-05-07
JOAN SEOANE
Head of Gene Expression & Cancer Research Group, Institut Recerca Hospital Universitari Vall d’Hebron, Barcelona, Spain
"The dark side of TGF-beta: The oncogenic role of TGF-beta in glioma"

27-04-07
MIGUEL A. PIRIS
Programa de Patologia Celular, CNIO, Madrid, Spain
"Lymphoma molecular machinery"

20-04-07
ALAN HINNEBUSCH
National Institute of Child Health and Human Development, NIH, Bethesda, USA
"Functions of translation initiation factors inl scanning and AUG recognition"

13-04-07
ASIFA AKHTAR
EMBL, Heidelberg, Germany
"Dosage compensation, old complex new facts"

16-03-07
REBECCA HELAD
Associate Professor of Cell & Developmental Biology, Molecular & Cell Biology, Berkeley, University of California, USA
"Mechanisms of mitotic spindle assembly and function"

9-03-07
ALFRED WITTINGHOFER
Professor, Max-Planck-Institut fur Molekulare Physiologie, Dortmund, Germany
"Mechanisms of signaling via G proteins"
2-03-07
AXEL ULLRICH
Dept. of Molecular Biology, Max Planck Institute for Biochemistry, Martinsried, Germany
"OncoGenome Analysis as Basis for Cancer Drug Development"

22-02-07
DAVID RON
Skirball Institute, New York, USA
"Cellular adaptations to unfolded protein load"

16-02-07
ERIC WESTHOFF
Institut de biologie moleculaire et cellulaire, Strasbourg, France
"RNA Motifs and RNA folding"

14-02-07
ANTHONY J. BROOKES
Professor, Department of Genetics, University of Leicester, UK
"Abundant short-range 'hidden' structural variation in the human genome"

9-02-07
LUIS GARCIA
Genethon & CNRS UMR 8115, Evry, France
"Dystrophin rescue by using exon skipping: a therapeutic avenue for DMD"

2-02-07
JUAN MENDEZ
CN10, Madrid, Spain
"Control of human DNA replication"

17-01-07
NICHOLAS LUSCOMBE
EMBL-European Bioinformatic Institute Wellcome, Trust Genome Campus, Cambridge, UK
"Transcription regulation: from parts list to genomic network"

PROGRAMME SEMINARS 2007

GENE REGULATION

30-11-07
MARIA PAOLA PARONETTO
Institute of Experimental Neuroscience CERC Fondazione Santa Lucia, Rome, Italy
"The role of the RNA binding protein Sam68 in germ cell"

16-10-07
LUIZ OTAVIO PENALVA, PH.D
Children’s Cancer Research Institute - UTH-SCSA, Department of Cellular and Structural Biology, San Antonio, Texas, USA
"Finding connections between post-transcriptional regulation and cancer"

19-09-07
CHIARA GAMBERI
McGill University, Department of Biology, Montréal, Canada
"BICAUDAL C REGULATES nos EXPRESSION DURING DROSOPHILA OOGENESIS"

14-06-07
BARTOSZ TARKOWSKI
Biotechnology student, Warsaw University, Poland
"Investigation of interactions between E4orf4 and proteins containing RNA-recognition motives"

31-05-07
STEVE LIEBHABER AND NANCY COOK
University of Pennsylvania, USA
"Novel pathways governing mRNA synthesis and stability"

3-05-07
RUTH SIMON
Department of Pathology, Mount Sinai School of Medicine, New York, USA
"The homeobox gene Sax2 and energy homeostasis"

19-04-07
CELIA JERÓNIMO
Institut de Recherches Cliniques de Montréal, Canada
"Systematic analysis of the protein interaction network for the human transcription machinery"
Unlocking the door of small RNAs: LNA as a tool in microRNA research

TopBP1 mediated chromatin modification and the DNA damage response

COUP-TFI orphan receptor in breast cancer cells: Influence on Estrogen receptor alpha activity and estrogen responsiveness

Structural and functional implications of chromatin containing core histone variants

ATP-dependent remodelling of H1 containing nucleosomal arrays

Mechanisms of Epigenetics: facts, clues, mysteries

The role of ubiquitin in the regulation of Drosophila NFkB signaling

BCL6: from epigenetics to therapeutic targeting of diffuse B cell leukemia

Pleiotrophic roles for syndecan-4 in muscle satellite cells: activation, proliferation, migration, and differentiation
17-05-07  
ANTONIO RUIZ-VELA  
CNIO, Madrid, Spain  
"RNA interference screen for cell death regulators: identification of MCL-1 as a critical factor that modulates gene expression"

7-05-07  
MIRNA PEREZ-MORENO  
Laboratory of Mammalian Cell Biology, The Rockefeller University, New York, USA  
"Links between Cell adhesion, Inflammation and Cancer"

7-05-07  
JUAN JOSE VENTURA  
Molecular Oncology, Spanish National Cancer Center (CNIO), Madrid, Spain  
"Essential role of p38δ MAP kinase in lung stem/progenitor cell proliferation and differentiation: implications for tumorigenesis"

7-05-07  
ERNESTO GUCCIONE  
European Institute of Oncology, Dept. of Experimental Oncology, Milano, Italy  
"Myc and chromatin: a mutual relationship".

7-05-07  
CRISTIANO SIMONE  
Division of Medical Genetics, Department of Biomedicine in Childhood, University of Bari, Italy  
"Signal-dependent control of autophagy and cell death in colorectal cancer: the role of the p38 pathway"

23-04-07  
ALEX MEISSNER  
Whitehead Institute for Biomedical Research, Cambridge, USA  
"Epigenetic Reprogramming and induced pluripotency"

26-02-07  
J.C. ZUNIGA-PFLUCKER  
Profesor Department of Immunology, University of Toronto, Senior Scientist Sunnybrook Research Institute, Canada  
"Reconstructing T cell development in vitro"

**GENES & DISEASE**

5-09-07  
FRANCISCO M. DE LA VEGA  
Computational Genetics, Applied Biosystems  
"Pursuing association signals of complex disease by next-generation, high-throughput sequencing"

26-07-07  
EVA RIVEIRA  
Université Catholique de Louvain, Brussels, Belgium  
"Genotype-phenotype correlations provide new insights in the pathogenesis of Gitelman’s syndrome"

17-07-07  
FLORIAN GRAEDLER & JEAN CLAUDE GERARD  
Illumina Inc.  
"Introducción al ‘Illumina Genome Analyzer’, última generación de sistemas de secuenciación de ADN".

17-07-07  
MARCUS DROEGE  
Global Marketing Director, Genome Sequencing.  
‘GS-FLX technology in DNA sequencing: a revolution that speeds up the discovery in genomics and transcriptomics’

13-06-07  
CARLOS AVENDANO  
Depto. de Anatomía, Histología y Neurociencia Fac. de Medicina, Universidad Autónoma de Madrid, Spain  
"Stereology: A fast lane to quantitative morphology... with road traps".

13-02-07  
THOMAS SCHULZE  
Central Institute of Mental Health, Germany  
"Exact phenotype definition for complex genetic traits: Novel strategies to establish valid diagnostic entities in psychiatric genetics".

29-01-07  
FRANCESCO TIMOLATI  
University Hospital, Switzerland  
"Neuregulin 1 beta modulates oxidative stress, contractile function and gene expression in cardiomyocytes subjected to cancer therapy"


**BIOINFORMATICS & GENOMICS**

26-11-07
IONAS ERB
Biozentrum, University of Basel, and Swiss Institute of Bioinformatics, Switzerland
"Toward a Grammar of Yeast’s Transcriptional Regulatory Code"

23-11-07
MATTHIAS ZYTNICKI
French National Research Institute for Agronomics of Toulouse, France
"Localizing non-coding RNAs with DARN!"

15-11-07
GABRIEL VALIENTE
UPC, Barcelona, Spain
"Alignment of Phylogenetic Networks"

8-11-07
DEEPAK TAKKAR
Silicon Graphics Inc
"SGI in BioSciences - Accelerating the pace of research"

30-10-07
GÜNTER MERTES
Business Manager Gene Expression, QIAGEN
"Fast, Cost Effective and flexible solutions in: real time PCR, effective Gene Silencing and miRNA quantification"

28-09-07
THOMAS DERRIEN
CNRS, Genetique et Developpement, Faculte de Medecine, France
"Comparative genomics and dog genome annotation"

18-09-07
ALVARO RADA
Inst. for genetics and pathology, Rudbeck Laboratory, Uppsala University, Sweden
"Affymetrix Chip-on-Chip workshop - Whole-genome maps of USF1 and USF2 binding and histone H3 acetylation reveal new aspects of promoter structure"

5-09-07
SUBHAJYOTI DE
MRC Laboratory of Molecular Biology, Cambridge, United Kingdom
"Change in genomic neighborhood of genes shapes expression divergence during human evolution"

31-07-07
ELODIE PORTALES-CASAMAR
University of British Columbia, Centre for Molecular Medicine and Therapeutics, Child and Family Research Institute, Vancouver, Canada
"PAZAR: a framework for collection and dissemination of cis-regulatory sequence annotation"

19-07-07
DARRELL CONKLIN
Department of Computing, City University, London, UK
"A putative Drosophila helical cytokine functioning in innate immune signalling"

28-06-07
MANUEL IRMIA
Department of Genetics, University of Barcelona, Spain
"When did Alternative splicing evolve?"

21-06-07
DAVID TALAVERA
Molecular Modeling and Bioinformatics Unit, Barcelona Scientific Parc & Protein Structure and Modelling Node, Genome Spain
"Alternative Splicing: a Bioinformatics View"

24-05-07
NOBURU JO SAKABE
Ludwig Institute for Cancer Research, Sao Paulo, Brazil
"Sequence features responsible for intron retention in human"

17-05-07
GENÍS PARRA
Genome Center, UC Davis, USA
"Assessment of the completeness of eukaryotic genomes"

10-05-07
GABRIEL VALIENTE
UPC, Barcelona, Spain
"Feasible Biochemical Reactions in Metabolic Pathways"
3-05-07
CARLOS QUIJANO
Computacionals Genomics, Barcelona
Supercoputer Centre, Spain
"Preferential retention of developmental regulators as tandems in insects"

26-04-07
BORJA CALVO
Intelligent Systems Group, Department of Computer Science and Artificial Intelligence, University of the Basque Country - UPV/EHU, Spain
"Learning with Positive and Unlabeled Examples"

30-03-07
JEFF ROBSON
Biotools B&M Labs. S.A.
Demo Teórico-Práctica de RotorGene 6000 de Corbett

9-03-07
MIKAEL KUBISTA
Professor at Chalmers University of Technology and Founder and Chairman of TATAA Biocenter AB
"Quantification of gene expression by reverse transcription real-time PCR"

9-03-07
MIKAEL KUBISTA
Professor at Chalmers University of Technology and Founder and Chairman of TATAA Biocenter AB
"Real-time PCR gene expression profiling"

9-03-07
MÓNICA MONTE
Eppendorf Ibérica
"Eppendorf and Real Time PCR"

15-02-07
XAVIER CRISTINA
EUROPEAN GENOTYPING BUSINESS DEVELOPMENT SPECIALIST, Applied Biosystems
"Applied Biosystems 7900 Fast Real Time PCR. Technical Introduction and Applications"

5-02-07
BART DEPLANCKE
Program of Gene Expression and Function, University of Massachusetts Medical School, Worcester, Massachusetts, USA
"Deciphering metazoan transcription regulatory networks using a gene-centered protein-DNA interaction mapping approach."

29-01-07
VINCENT LACROIX
Laboratoire de Biométrie et Biologie Évolutive, Université Lyon 1, France
"Motif Search in Metabolic Networks"

16-01-07
VICTOR SOLOVYEV
Professor, Department of Computer Science, Royal Holloway, University of London, UK
"Computational identification genes and promoters in eukaryotic and bacterial genomes"

16-01-07
FRANCISCO AZUAJE
IEEE Senior Member, University of Ulster, School of Computing and Mathematics, Northern Ireland, UK
"Computational intelligence and the study of complex biological systems."

16-01-07
CEDRIC NOTREDAME
Structural and Genetic Information, CNRS UPR 2589, Marseille, France
"Revealing Evolutionary Traces through Multiple (Sequence) Comparisons"

15-01-07
ERAN HALPERIN
Senior Research Scientist, International Computer Science Institute, California, USA
"Computational Methods for Disease Association Studies"

15-01-07
GABRIEL VALIENTE
Algorithms, Bioinformatics, Complexity and Formal Methods Research Group, UPC, Barcelona, Spain
"Comparative Analysis of Biochemical Pathways and Networks"

15-01-07
HECTOR GARCÍA MARTÍN
Postdoctoral Fellow, Microbial Ecology Program, DOE Joint Genome Institute (Lawrence Berkeley Lab), California, USA
"Wastewater Sludge Metagenomics and Beyond"

15-01-07
PAVEL SUMAZIN
Assistant Professor at Portland State University, Department of Computer Science, Oregon, USA
"Reverse engineering regulatory circuits from sequence-element anchors"
"Repression of an R8 targeting program is essential for R7 synaptic specificity"

"Study of the role of microtubules and actomyosin in migrating cortical interneurons"

"Proteoglycan function in early Xenopus development"

"Coordination of cytokinesis with chromosome segregation"

"Excitable dynamics in a gene-regulatory network"

"Evolutionary and Developmental Dynamics of the Gap Gene Network"

"The tumour suppressor PTEN in zebrafish development"

"Bringing integrins to order, alphaVbeta3 and alpha5beta1 characterization in fibroblasts"
9-07-07
CRISTINA VOGEL
Institute for Cellular and Molecular Biology, University of Texas at Austin, Austin, USA
"From evolution and expression - different perspectives on the protein repertoire defining organismal properties"

20-06-07
GUILLERMO PRADOS
National Instruments
National Instruments Software presentation

29-05-07
TECAN
TECAN
Tecan Robots DEMO: Evo 150, Evo 100, TeStack.

28-05-07
MARK HENKELMAN
Mouse Imaging Centre (MICe), Department of Medical Biophysics, University of Toronto, Toronto, Canada
"Imaging for mouse phenotyping"

23-05-07
MARTA SALES
Northwestern Institute on Complex Systems & Dept. Chem. & Bio. Eng., Northwestern University, Evanston, USA
"Topology and dynamics of complex biological systems"

23-05-07
JULIEN COLOMBELLI
EMBL, Heidelberg, Germany
"Mechanical and Molecular relaxation of the cytoskeleton after laser nanosurgery"

21-05-07
GIDEON SCHREIBER
Department of Biological Chemistry, Weizmann Institute of Science, Rehovot, Israel
"The feelings of proteins towards each other, and what can we do about it"

18-04-07
CÉCILE LESTURGEOUN
Application Specialist, Caliper Life Sciences (Vertex)
Caliper Life Sciences Liquid Handling workstations and Drug Discovery Applications: Genomics, Proteomics & Screening

22-02-07
RUBEN ABAGYAN
The Scripps Research Institute, Department of Molecular Biology, La Jolla, California, USA
Predicting molecular interactions

22-02-07
PATRICE VACHETTE
IBBMC, UMR 8619 CNRS-Université Paris-sud, Orsay, France
Small-angle X-ray Scattering : a useful addition to the structural biologist's toolkit

22-02-07
DR. IMRE BERGER
Institute of Molecular Biology and Biophysics, Swiss Federal Institute of Technology, Zürich, Switzerland
Multiprotein expression platform for structural genomics of eukaryotic complexes

22-02-07
DAVID EISENBERG
UCLA-DOE Institute for Genomics and Proteomics, Depts of Chemistry & Biochemistry and Biological Chemistry, Molecular Biology Institute, LA, USA
Protein interactions in health and disease

22-02-07
CHERYL ARROWSMITH
Division of Cancer Genomics and Proteomics, Ontario Cancer Institute (OCI), Toronto, Canada
Structural & Chemical Genomics of Human Gene Families

6-02-07
MARTA IBAÑES
Estructura i Constituents de la Matèria, Facultad de Física, Universitat de Barcelona, Barcelona, Spain
"Cell differentiation: from individual to collective behaviour"

6-02-07
JUAN POYATOS
Structural and Computational Biology Programme, Spanish National Cancer Centre (CNIO), Madrid, Spain
"Understanding a wise old bird with a few simple words: Pragmatic Systems Biology"
6-02-07
JAVIER BUCETA
Centre Especial de Recerca en Química Teórica, Parc Cientific de Barcelona, Barcelona, Spain
"Boundary Formation in the Drosophila Wing: a Systems Biology Approach"

6-02-07
DENIS THIEFFRY
Université de la Méditerranée, Marseille, France
"Logical Modelling of genetic regulatory networks"

5-02-07
ATTILA CSIKASZ-NAGY
Budapest University of Technology and Economics, Molecular Network Dynamics Research Group, Hungary
"Models of cell cycle regulation"

5-02-07
DENIS DUPUY
Dana-Farber Cancer Institute, Harvard Medical School, USA
"High-throughput quantitative analysis of in vivo gene expression in C.elegans"

5-02-07
MATTHIEU LOUIS
Laboratory of Neurogenetics and Behavior, The Rockefeller University, New York, USA
"Chemotaxis in Drosophila melanogaster: Making sense of graded olfactory stimuli"

5-02-07
TIMOTHY RAVASI
Jacobs School of Engineering, Department of Bioengineering, University of California, San Diego, USA
"System approaches to infer transcriptional regulatory networks in mammals"

11-01-07
PETER PHILIPPSEN
Professor of Microbiology, Biozentrum University of Basel, Switzerland
"S. cerevisiae and A. gossypii: Evolution of different life styles starting from the same set of genes"
### APPENDIX 3

## GRANTS

The grants that the CRG obtained from 1st January to 31st December 2007 are the following:

<table>
<thead>
<tr>
<th>Organism</th>
<th>Amount (euros)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MINISTERIO DE EDUCACION Y CIENCIA.</td>
<td>7,177,496,59</td>
</tr>
<tr>
<td>EUROPEAN COMMISSION</td>
<td>4,264,983,70</td>
</tr>
<tr>
<td>FUNDACION MARCELINO BOTIN</td>
<td>1,320,000,00</td>
</tr>
<tr>
<td>NATIONAL INSTITUTE OF HEALTH NIH</td>
<td>904,477,96</td>
</tr>
<tr>
<td>CONSORCI PARC DE RECERCA BIOMEDICA DE BARCELONA</td>
<td>902,961,93</td>
</tr>
<tr>
<td>MINISTERIO DE SANIDAD Y CONSUMO-FIS</td>
<td>732,611,06</td>
</tr>
<tr>
<td>FUNDACION DESARROLLO INVESTIGACION GENOMICA Y PROT</td>
<td>591,174,14</td>
</tr>
<tr>
<td>FUNDACIO MARATO TV3</td>
<td>350,835,50</td>
</tr>
<tr>
<td>FUNDAÇAO PARA A CIÊNCIA E A TECNOLOGIA</td>
<td>253,800,00</td>
</tr>
<tr>
<td>MUSCULAR DYSTROPHY ASSOCIATION</td>
<td>215,835,50</td>
</tr>
<tr>
<td>ASSOCIATION FOR INTERNATIONAL CANCER RESEARCH</td>
<td>215,598,85</td>
</tr>
<tr>
<td>AGENCIA GESTIO D'AJUTS UNIVERSITARIS</td>
<td>171,172,92</td>
</tr>
<tr>
<td>GRANTS MISCELANIOUS</td>
<td>142,028,95</td>
</tr>
<tr>
<td>HUMAN FRONTIERS SCIENCE</td>
<td>110,353,00</td>
</tr>
<tr>
<td>FUNDACIÓN RAMON ARECES</td>
<td>110,000,00</td>
</tr>
<tr>
<td>FUNDACIÓN ALICIA KOPLOWITZ</td>
<td>100,000,00</td>
</tr>
<tr>
<td>FONDATION RECHERCHE MÉDICALE</td>
<td>96,000,00</td>
</tr>
<tr>
<td>NOVARTIS FARMACEUTICA S.A.</td>
<td>96,000,00</td>
</tr>
<tr>
<td>EUROPEAN MOLECULAR BIOLOGY ORGANIZATION</td>
<td>86,916,37</td>
</tr>
<tr>
<td>FEDERATION EUROPEAN BIOCHEMICAL SOCIETIES</td>
<td>75,000,00</td>
</tr>
<tr>
<td>FONDS ZUR FÖRDERUNG DER WISSENSCHAFTLICHEN FOR</td>
<td>54,000,00</td>
</tr>
<tr>
<td>ASSOCIATION FRANÇAISE CONTRE LES MYOPATHIES AFM</td>
<td>51,500,00</td>
</tr>
<tr>
<td>NATIONAL UNIVERSITY OF IRELAND, GALWAY</td>
<td>50,000,00</td>
</tr>
<tr>
<td>DEUTSCHE FORSCHUNGSGEMEINSCHAFT (DFG)</td>
<td>43,674,00</td>
</tr>
<tr>
<td>FUNDACIÓN LA CAIXA</td>
<td>36,360,60</td>
</tr>
<tr>
<td>FONDATION JEROME LEJEUNE</td>
<td>20,000,00</td>
</tr>
<tr>
<td>DEUTSCHER AKADEMISCHER AUSTAUSCH DIENST</td>
<td>14,112,00</td>
</tr>
<tr>
<td>COMISION FULBRIGHT</td>
<td>13,518,00</td>
</tr>
<tr>
<td>AGILENT TECNOLOGIES SALES &amp; SERVICES GMBH&amp;CO.KG</td>
<td>10,000,00</td>
</tr>
<tr>
<td>ROCHE DIAGNOSTICS GMBH</td>
<td>10,000,00</td>
</tr>
<tr>
<td>APPPLERA HISPANIA S.A.</td>
<td>7,000,00</td>
</tr>
<tr>
<td>LABORATORIOS ALMIRALL S.A.</td>
<td>6,000,00</td>
</tr>
<tr>
<td>SANOFI-AVENTIS, S.A.</td>
<td>6,000,00</td>
</tr>
<tr>
<td>UNIVERSITAT POMPEU FABRA</td>
<td>5,500,00</td>
</tr>
<tr>
<td>FUNDACIÓN ESPAÑOLA PARA LA CIENCIA Y LA TECNOLOGIA.</td>
<td>5,000,00</td>
</tr>
<tr>
<td>SEQUENOM GMB</td>
<td>4,000,00</td>
</tr>
<tr>
<td>SOCIETAT CATALANA DE BIOLOGIA.</td>
<td>3,000,00</td>
</tr>
<tr>
<td>BOEHRINGER INGELHEIM FONDS (FOUNDATION FOR BASIC R</td>
<td>2,950,00</td>
</tr>
<tr>
<td>BECKMAN COULTER, INC</td>
<td>2,400,00</td>
</tr>
<tr>
<td>ABCAM LIMITED.</td>
<td>2,331,14</td>
</tr>
<tr>
<td>KUTXA ESPACIO DE LA CIENCIA</td>
<td>1,056,96</td>
</tr>
<tr>
<td>FEDERATION OF EUROPEAN NEUROSCIENCE SOCIETI</td>
<td>1,000,00</td>
</tr>
</tbody>
</table>

**Total Amount** ........................................... 18,266,648,94
APPENDIX 4

FINANCE & PERSONNEL EVOLUTION AT THE CRG

FINANCE

Millions €

12.00 –

10.00 –

8.00 –

6.00 –

4.00 –

2.00 –

0.00 –

2001 2002 2003 2004 2005 2006 2007

- Innovation, Universities & Enterprise Dept.
- Spanish Ministry of Science & Education
- International
- Health Dept.
- National
**Postdocs**

<table>
<thead>
<tr>
<th>Year</th>
<th>Nr. People</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>20</td>
</tr>
<tr>
<td>2004</td>
<td>21</td>
</tr>
<tr>
<td>2005</td>
<td>30</td>
</tr>
<tr>
<td>2006</td>
<td>59</td>
</tr>
<tr>
<td>2007</td>
<td>86</td>
</tr>
</tbody>
</table>

**PhD Students**

<table>
<thead>
<tr>
<th>Year</th>
<th>Nr. People</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>45</td>
</tr>
<tr>
<td>2004</td>
<td>54</td>
</tr>
<tr>
<td>2005</td>
<td>57</td>
</tr>
<tr>
<td>2006</td>
<td>73</td>
</tr>
<tr>
<td>2007</td>
<td>75</td>
</tr>
<tr>
<td>Year</td>
<td>Technicians</td>
</tr>
<tr>
<td>------</td>
<td>-------------</td>
</tr>
<tr>
<td>2003</td>
<td>31</td>
</tr>
<tr>
<td>2004</td>
<td>37</td>
</tr>
<tr>
<td>2005</td>
<td>34</td>
</tr>
<tr>
<td>2006</td>
<td>48</td>
</tr>
<tr>
<td>2007</td>
<td>61</td>
</tr>
</tbody>
</table>