IMPPC
Institut de Medicina Predictiva
i Personalitzada del Càncer

4DCellFate

Integrative and Computational Biology
Joint symposium

V IMPPC Annual Conference and 4DCellFate Workshop

Barcelona 27 / 28 March 2014

welcome

Dear colleague,

Welcome to the conference Integrative and Computational Biology Joint symposium: V IMPPC Annual Conference (www.imppc.org) and 4DCellFate Workshop (www.4dcellfate.eu).

As a joint event between the Integrative Cancer Biology program of the Institute of Predictive and Personalized Medicine (IMPPC) and the FP7 Project 4DCellFate, coordinated by the CRG, we are sure that this conference will be of interest to computational biologists as well as experimental researchers working in the fields of genomics, epigenetics and cancer. Leading scientists of international standing in the fields of computational and integrative biology will discuss cutting-edge developments in current topics ranging from chromatin biology, gene regulation and cancer genome evolution.

One of the biggest opportunities in biology today is to make large-scale precise, systematic measurements and to computationally integrate large datasets, which can be used to address fundamental questions in biology. Our biggest challenge is how to exploit all this data in meaningful ways, analyze it in order to transform it to basic knowledge and also to translate it to a better understanding of diseases. We believe that this research will eventually lead to better diagnostics for diseases such as cancer, better predictions for prognosis and personalized treatments.

We are proud to present our program and we wish to thank the IMPPC and 4DCellFate for making the effort needed to put on a conference of this nature.

Yours sincerely,

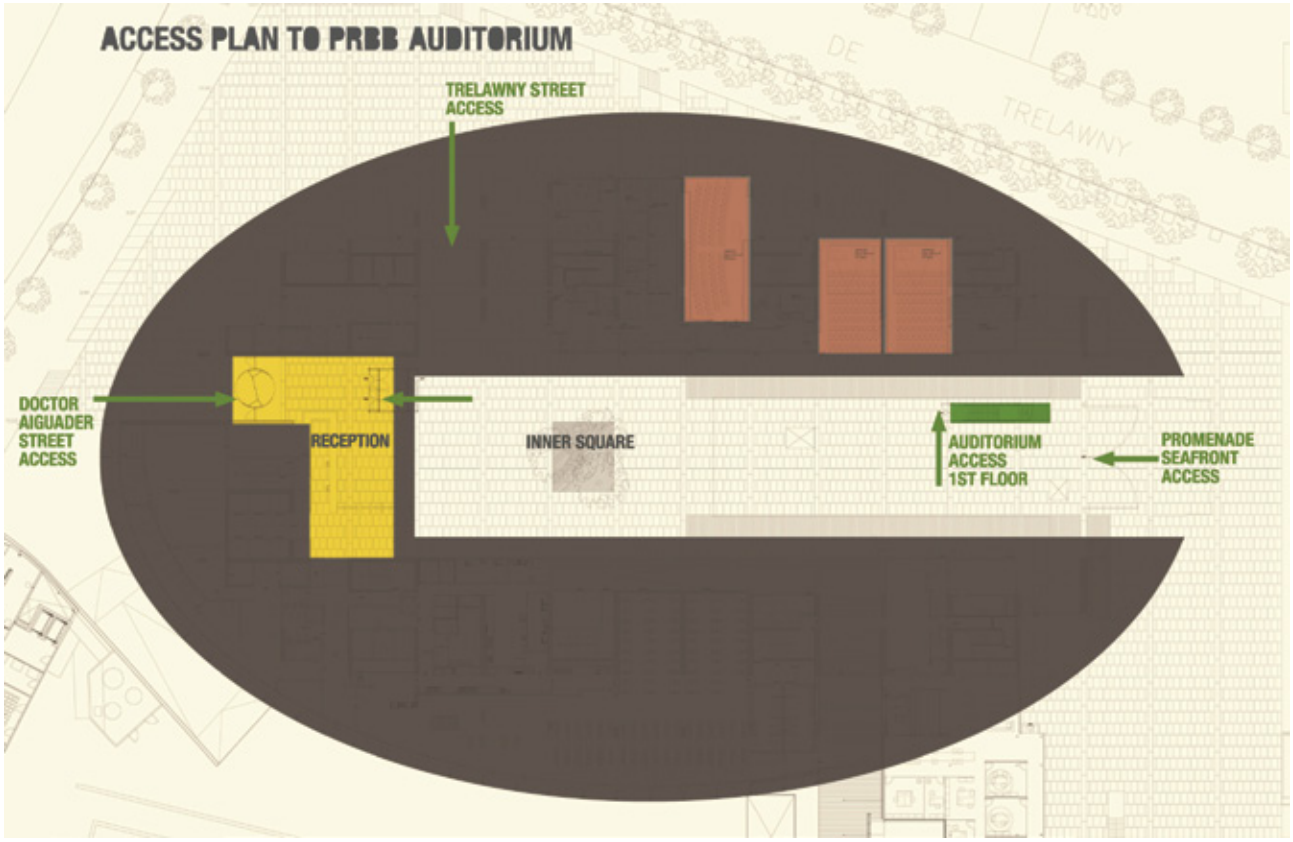
Ben Lehner - 4DCellFate, Centre for Genomic Regulation (CRG), Barcelona
Ana Rojas - Currently at the Institute of Biomedicine Sevilla (IBIS)
Tanya Vavouri - IMPPC

contents

PRBB

Our Venue
PRBB Auditorium
Dr Aiguader 88
08003 Barcelona
How to get here?

Entry
At the PRBB, you can reach the Auditorium via the staircase from the inner square (inside the horseshoe of the building).



Accreditations
Please wear your accreditation label at all times during the conference and have it visible when using the catering services. Remember to bring it with you for the second day.

Wifi Access
From the PRBB wifi access page choose the network “PRBB” and add the password darwin1809.

Smoking
The PRBB is a no-smoking building. The designated smoking areas are the inner square (downstairs) and anywhere outside the building.

program

Thursday 27 March

08.50 - 09.00 **Welcome from the Organizers**

09.00 - 10.50 Session I:
Big Biological Datasets

Chairman: Ben Lehner

09.00 - 09.30 **Luis Serrano,** Centre for Genomic Regulation (CRG), Spain
Protein competition as a means for cell specific signaling

09.30 - 10.00 **Olga Troyanskaya,** Lewis-Sigler Institute for Integrative Genomics
A tissue-specific view of human disease from integrated data analysis

10.00 - 10.30 **Roderic Guigo,** Centre for Genomic Regulation, Spain
RNA Heterogeneity in the eurkaryotic cell

10:30-10:50 **Eduardo Eyras,** Pompeu Fabra University
The landscape of alternative splicing alterations in cancer

10.50 - 11.20 **Coffee and Poster Session**

11.20- 13.30 Session II:
Modelling Chromatin

Chairman: Marc Marti-Renom

11.20 - 11.50 **Modesto Orozco,** Institute for Research in Biomedicine (IRB Barcelona), University of Barcelona (UB) and Barcelona Supercomputing Center (BSC), Spain
Chromatin analyzed with the eyes of physics

11.50 - 12.10 **Luca Giorgetti,** Institute Curie, France
Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription

12.10 - 12.30 **Ana M. Rojas,** Institute of Biomedicine of Seville (IBIS), Spain
Emergence and conservation of DNA Damage Response proteins affecting chromatin remodelling

12.30 - 12.50 **Vladimir Teif,** German Cancer Research Center (DKFZ)
Nucleosome re-arrangement as a feedback mechanism between DNA methylation and transcription factor binding

12:50-13:30 **Fast Track Session (I):** 8 Speed Talks
Speed presentations by selected poster authors

13.30 - 14.45 **Lunch and Poster Session**

14.45 - 16.05 Session III:
Cancer Genomics (Part I)

Chairman: Sergio Alonso

14.45 - 15.35 **Fast Track Session (II):** 9 Speed Talks
Speed presentations by selected poster authors

15.35 - 16.05 **Nuria Lopez-Bigas,** ICREA and Pompeu Fabra University, Spain
Identification of cancer drivers across tumor types

16.05 - 16.35 **Alfonso Valencia,** Spanish National Cancer Research Centre (CNIO), Spain
Analysis of Genomes and Cancer Genomes from an Evolutionary Perspective

16.35 - 17.05 **Coffee and Poster Session**

17.05 - 19.00 Session IV:
Cancer Genomics (part II)

Chairman: Alfonso Valencia

17.05 - 17.35 **Niko Beerenwinkel,** Swiss Federal Institute of Technology Zurich, Switzerland
Modeling cancer evolution from genomic data

17.35 - 18.05 **Elizabeth Murchison,** Wellcome Trust Sanger Institute, UK
Genetics and evolution of clonally transmissible cancers in dogs and Tasmanian devils

18.05 - 19.00 **Søren Brunak EMBO KEYNOTE LECTURE,** Center for Biological Sequence Analysis, Technical University of Denmark
Comorbidities and disease trajectories in cancer and beyond

19.00 - 20.00 **Poster Session**

Friday 28 March

09.30 - 11.30 Session V:
Regulatory Genomics

Chairman: Eduardo Eyras

09.30 - 10.00 **Anne-Ferguson-Smith,** University of Cambridge
Environmental modulation of epigenetic states in the sperm methylome and their consequences

10.00 - 10.20 **Tanya Vavouri,** Institute of Predictive and Personalized Medicine of Cancer (IMPPC)
The small RNA content of human sperm

10.20 - 10.50 **Mihaela Zavolan,** Biozentrum, University of Basel, Switzerland
Decoding miRNA-based regulation with low and high-throughput data

10.50 - 11.10 **Lorenzo Pasquali,** IDIBAPS
Pancreatic islet enhancer clusters enriched in type 2 diabetes risk-associated variants

11.10 - 11.30 **Tuncay Baubec,** Friedrich Miescher Institute for Biomedical Research
DNA methylation: (epi-)genomic views on pattern generation and interpretation

11.30 - 12.00 **Coffee and Poster Session**

12.00- 13.30 Session VI:
Integrative Biology

Chairman: Olga Troyanskaya

12.00 - 12.30 **Ewan Birney,** EMBL-EBI, Wellcome Trust Genome Campus, UK
EMBL-EBI and Elixir: A life science infrastructure

12.30 - 13.00 **Eileen Furlong,** EMBL Heidelberg, Germany
cis-regulatory control during development

13.00 - 13.30 **Nick Luscombe,** University College London and Cancer Research UK London Research Institute (joint appointments)
Guardians of the transcriptome: how hnRNP C protects you from unwanted axons

13.30 **Closing Remarks**

Invited talk

Luis Serrano
Centre for Genomic Regulation (CRG), Spain

Protein competition as a means for cell specific signaling

Abstract
The mechanisms of context-specific differences in signal transduction, such as those that occur among different cell types, are not fully understood. One possibility is that differences in the abundance of proteins change signaling outputs because these proteins compete for binding to hub proteins at critical network branch points. Focusing on the ErbB signaling, we created a protein interaction network that included information about protein domains and analyzed the role of competing protein interactions. By leveraging three dimensional protein structures to infer steric interactions among binding partners for a common binding domain or linear motif (node) and including information about protein abundance and interaction affinities, we identified a large number of competitive, mutually exclusive (XOR) protein interactions. Modeling changes in protein abundance with different patterns of partner proteins and XOR nodes (XOR motifs) revealed that each motif conferred a different response. We experimentally investigated the XOR motif containing the hub protein Ras and its binding partners RIN1 (Ras and Rab interactor 1) and CRAF (v-raf–leukemia viral oncogene 1). Consistent with the computational prediction, overexpression of RIN1 in cultured cells decreased the phosphorylation of CRAF and its downstream targets. Thus, our analyses provide evidence that variation in the abundance of proteins that compete for binding to XOR nodes could contribute to context-specific signaling plasticity.

Invited talk

Olga Troyanskaya
Lewis-Sigler Institute for Integrative Genomics

A tissue-specific view of human disease from integrated data analysis

Abstract
An immense molecular complexity forms the foundation of human disease. Our goal is to interpret and distill this complexity through accurate modeling of molecular networks and pathways, particularly those in which malfunction promotes the emergence of complex human disease. Although cell-lineage-specific gene expression and function underlie the development, function, and maintenance of diverse cell types within an organism, high-throughput data are rarely resolved with respect to specific cell lineages. In this talk, I will focus on our recent work developing integrative approaches that leverage functional genomics data collections to study how cellular pathways function in diverse cell types. Our work accounts for cell-lineage specificity in integrated models of gene expression and protein networks, including predictions of tissue-specific and cell-lineage-specific gene expression in human with accuracies higher then those of high-throughput experimental studies. I will also discuss how integrated analysis of functional genomics data can be leveraged to study tissue-lineage-specific protein function and interactions and to identify genes involved in disease in a way complementary to quantitative genetics approaches.

Invited talk

Roderic Guigo

Centre for Genomic Regulation (CRG), Pompeu Fabra University (UPF), Spain

RNA heterogeneity in the eukaryotic cell

Abstract

The unfolding of the instructions encoded in the genome is triggered by the transcription of DNA into RNA, and the subsequent processing of the resulting primary RNA transcripts into functional mature RNAs. RNA is thus the first phenotype of the genome, mediating all other phenotypic changes at the organism level caused by changes in the DNA sequence. While current technology is too primitive to provide accurate measurements of the RNA content of the cell, the recent development of Massively Parallel Sequencing Instruments has dramatically increased the resolution with which we can monitor cellular RNA. Using these instruments, the ENCODE project has surveyed the RNA content of multiple cell lines and subcellular compartments. The results of these surveys underscore pervasive transription, as well as great RNA heterogeneity between and within cells. Comparison of RNA surveys with other genome wide epigenetic surveys - such as those of binding sites for Transcription Factors, or of Histone modifications - reveals a very tightly coupling between the different pathways involved in RNA processing, transcription and splicing in particular.

Short talk

Eduardo Eyras

E. Eyras (presenting author), E. Sebestyen, G.P. Alamancos, A. Pages, R. Drews

Computational Genomics, Pompeu Fabra University (UPF), Spain

The landscape of alternative splicing alterations in cancer

Abstract

Current cancer genomics projects apply high-throughput technologies to discover recurrent genetic variations in patient samples. These efforts are crucial to describe the genetic diversity of cancer (Stephens et al. 2012) and to classify into novel subtypes for improved prognosis and therapeutics (Ciriello et al. 2013). However, alterations in Alternative Splicing, which hold many potential signatures for prognosis and therapy (Karni et al. 2007, Grosso et al. 2014), have not been yet thoroughly characterized. We have used RNA sequencing data from The Cancer Genome Atlas (TCGA) project for more than 1000 tumor samples and paired normals, to characterize the alternative splicing of genes in 13 different cancer types. Using different association measures we derive regulatory modules involving splicing factors and events that are recurrently altered across tumors, as well modules that are altered only in specific cancer types. Additionally, we find evidence for RNA binding proteins and chromatin remodeling factors to have a potential novel role as regulators of alternative splicing. We further investigate how the alterations of alternative splicing may potentially remodel the network of protein-protein interactions, and find a number of interactions that are predicted to be disrupted in specific tumors. Our results provide new cancer signatures and suggest new possible therapeutic strategies.
Stephens et al. (2012). The landscape of cancer genes and mutational processes in breast cancer. Nature, 486(7403), 400-4. doi:10.1038/nature11017 Ciriello et al. (2013). Emerging landscape of oncogenic signatures across human cancers. Nat Genet. 45(10):1127-1133 Grosso et al. (2014). Kumar (ed.), Nuclear Signaling Pathways and Targeting Transcription in Cancer, 313 Cancer Drug Discovery and Development, Springer, 2014 Karni et al. (2007). The gene encoding the splicing factor SF2/ASF is a protooncogene. Nat Struct Mol Biol. 2007 Mar;14(3):185-93.

Invited talk

Modesto Orozco

Institute for Research in Biomedicine (IRB Barcelona), University of Barcelona (UB) and Barcelona Supercomputing Center (BSC), Spain

Chromatin analyzed with the eyes of physics

Abstract

DNA is a complex molecule whose expression needs to be carefully regulated. Evolved organisms have developed complex regulatory networks that determine when gene is located and when it should be active or inactive. A significant part of gene regulation is controlled by chromatin structure. Compact forms of chromatin will difficult interaction of DNA with regulatory proteins, reducing gene expression, while loss of chromatin compaction typically leads to gene activation. Chromatin structure should be then carefully controlled by the cell. During my talk I will introduce a different view to chromatin structure, arising from a simple physical analysis of DNA properties. I will show that even in the absence of regulatory proteins DNA has hidden signals that allow the cell in the organization of chromatin structure and in the definition of gene regulation regions.

Invited talk

Luca Giorgetti

Luca Giorgetti[1], Elphège P. Nora[1#], Tristan Piolot[1], France Lam[1], Job Dekker[2], Guido Tiana[3], Edith Heard[1]

[1] Institut Curie, CNRS UMR3215, INSERM U934 26 rue d'Ulm, Paris F-75248, France. [2] Program in Systems Biology, Department of Biochemistry and Molecular. Pharmacology, University of Massachusetts Medical School, Worcester, MA, USA. [3] Dipartimento di Fisica, Università degli Studi di Milano and INFN, 20133 Milano, Italy.
Present address: Gladstone Institute of Cardiovascular Diseases, CA 94158 USA

Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription

Abstract

Transcriptional regulation occurs in the context of dynamic chromosome architecture. Recently a new level of chromosome organization, Topologically Associating Domains (TADs), was uncovered by chromosome-confirmation-capture (3C) techniques. To explore TAD structure and function, we developed a polymer model that can extract the full repertoire of chromatin conformations within TADs from population-based 3C data. The model predicts the degree to which chromosomal contacts vary between cells. It also identifies interactions within single TADs that stabilize boundaries between adjacent TADs. Combining the model's predictions with high-resolution DNA FISH and quantitative RNA FISH for TADs within the X-inactivation center (Xic), we dissect the relationship between transcription activity of a promoter and its spatial proximity to cis-regulatory elements in single cells. We demonstrate that contacts between potential regulatory elements occur in the context of fluctuating structures rather than stable loops and propose that such fluctuations can contribute to asymmetric expression in the Xic at the onset of X inactivation.

Invited talk

Ana Rojas
Institute of Biomedicine of Seville (IBIS), Spain

Emergence and conservation of DNA Damage Response proteins affecting chromatin remodelling

Abstract
The DNA Damage response is a very important network to mantain genome integrity. This network is largely regulated by post-translational modification that influences the chromatin stage close to the damaged sites. In a recent work, we have compiled the most complete network of human DDR pathways including regulatory aspects (Post-translational modifications), and studied its emergence within a global evolutionary framework. Many pathways have been lost in model organisms, having a potential profound impact in comparative genomics approaches and further downstream analyses. We will present these data in the framework of chromatin-related proteins involving DDR processes.

Short talk

Vladimir Teif
Deutsches Krebsforschungszentrum (DKFZ) & BioQuant, Im Neuenheimer Feld 280, 69120 Heidelberg, Germany

Nucleosome re-arrangement as a feedback mechanism between DNA methylation and transcription factor binding

Abstract
DNA cytosine methylation (5mC) and hydroxymethylation (5hmC) are among the most important epigenetic marks. The interplay of 5mC/5hmC marks, the arrangement of nucleosomes and transcription factors (TFs) links DNA methylation with cellular gene expression programs but the underlying mechanisms are poorly understood. We have addressed this issue by taking advantage of our recently determined genome-wide nucleosome occupancies maps in mouse embryonic stem cells and their neural progenitor and embryonic fibroblast counterparts (Teif et al., Nature Struct Mol Biol, 19, 1185-92 (2012)). The latter study has provided insight about cell-type- and protein-specific binding preferences of transcription factors to sites with either low (Myc, Klf4 and Zfx) or high (Nanog, Oct4 and Sox2) nucleosome occupancy as well as more complex patterns for CTCF. Here, we analyzed nucleosome positioning, DNA methylation and TF binding in conjunction with additional dinucleosome occupancy maps. Our study provides quantitative description for the relations between DNA methylation/demethylation, TF binding and nucleosome occupancy changes.

Invited talk

Nuria López-Bigas

ICREA and Pompeu Fabra University (UPF), Spain

Identification of cancer drivers across tumor types

Abstract

Identifying cancer drivers is a key first step towards precision cancer medicine. The availability of thousands of sequenced tumor genomes/exomes opens the possibility to carry out a genome-wide identification of cancer driver mutations, genes and pathways in an unbiased manner. We have developed IntOGen-mutations, a computational platform, to produce a comprehensive landscape of the role of mutations, genes and pathways in the emergence of the cancer phenotype across tumor types. We have recently analyzed somatic mutations identified in close to 7,000 samples from 30 tumor types to detect signals of positive selection in the pattern of somatic mutations in genes. Combining these signals, we identified 464 high-confidence mutational driver genes (CDs), some of which had not been linked before to tumorigenesis. Notably, most CDs are mutated at low frequencies --88% bear mutations in less than 1% of the samples in the dataset. We are investigating which CDs are targets of tailored cancer drugs, as a first step to reveal the landscape of possible personalized treatments applicable to the close to 7,000 patients represented in the dataset.

Invited talk

Alfonso Valencia

Spanish National Cancer Research Centre (www.cnio.es)

Development of personalized treatment strategies for hereditary cancer syndromes

Abstract

The fast progression of genomics is making of the use of personal genomic information a pressing daily reality. In this scenario Bioinformatics and Computational Biology play a central rôle. The organization and analysis of individual genomes is a complex task involving data organization, integration and interpretation challenges, and requires a blend of engineering and scientific developments at each step of the analysis, since it touches many areas in which the development of computational methods is far from complete.

In the context of a various cancer genome related projects my group is developing both the technical framework for handling the data and the methods required for the interpretation of the information. Based on our experience in these projects I will review some of the key problems in the analysis of high-throughput genomic information, including the prediction of the incidence of mutations with special emphasis in the application of co-evolution related methods, the implications of the alterations of the splicing machinery, and the comparative analysis of disease affected pathways.

References

- Valencia A & Hidalgo M (2012) Getting personalized cancer genome analysis into the clinic: the challenges in Bioinformatics. Genome Medicine, 461.
- Vazquez M, de la Torre V, Valencia A (2012) Chapter 14: Cancer Genome Analysis, in Translational Bioinformatics PLOS Computational Biology open access book.
- de Juan D, Pazos F, Valencia A. (2013) Emerging methods in protein co-evolution. Nat Rev Genet. 14:249-61.
- Rodriguez JM, Maietta P, Ezkurdia I, Pietrelli A, Wesselink JJ, Lopez G, Valencia A, Tress ML. (2013) APPRIS: annotation of principal and alternative splice isoforms. Nucleic Acids Res. 41:D110-7.
- Ibañez C, Boullosa C, Tabarés-Seisdedos R, Baudot A and Valencia A (2014) Molecular Evidence for the Inverse Comorbidity between Central Nervous System Disorders and Cancers detected by Transcriptomic Meta-analyses. Plos Genet in press.

Invited talk

Niko Beerenwinkel
ETH Zurich, Switzerland

Modeling cancer evolution from genomic data

Abstract
Cancer evolution is a stochastic evolutionary process characterized by the accumulation of mutations and responsible for tumor growth, clinical progression, immune escape, and drug resistance development. Evolutionary theory can be used to describe the dynamics of tumor cell populations and to make inference about the evolutionary history of a tumor from molecular profiling data. We discuss recent approaches to modeling the evolution of cancer, including population genetics models of tumorigenesis, phylogenetic methods of intra-tumor subclonal diversity, and probabilistic graphical models of tumor progression. Evolutionary modeling will play an increasingly important prognostic role in predicting disease progression and the outcome of medical interventions, such as targeted therapy.

Invited talk

Elizabeth P. Murchison
University of Cambridge, Department of Veterinary Medicine, Cambridge, UK
Wellcome Trust Sanger Institute, Cancer Genetics and Genomics, Hinxton, UK

Genome analysis of clonally transmissible cancers in dogs and Tasmanian devils

Abstract
Tasmanian devil facial tumour disease (DFTD) and canine transmissible venereal tumour (CTVT) are the only two known naturally occurring clonally transmissible cancers. These are cancers that can be transmitted between individuals by the physical transfer of living cancer cells. Thus DFTD and CTVT are long-lived somatic cell lineages that each first originated once as cancers in single individuals but that have now spread through their respective host populations as parasitic clonal cell lineages. DFTD is spread by biting and is threatening its host species, the Tasmanian devil, with extinction. CTVT is a sexually transmitted cancer that affects dogs and has spread around the world together with its host. The genome sequences of DFTD and CTVT have revealed features of the individuals that first spawned these two lineages as well as patterns of mutation and selection that have driven the evolution and characterised the emergence and spread of these two unusual long-lived cancers. Clonally transmissible cancers are a poorly understood type of infectious pathogen; although there are only two known diseases of this type, such diseases can emerge rapidly and have disastrous implications for species conservation.

EMBO Keynote Lecture

Søren Brunak
Technical University of Denmark & University of Copenhagen

Comorbidities and disease trajectories in cancer and beyond

Abstract
Electronic patient records remain a rather unexplored, but potentially rich data source for discovering correlations between diseases, drugs and genetic information in individual patients. Such data makes it possible to compute fine-grained disease co-occurrence statistics, and to link the comorbidities to the treatment history of the patients. A fundamental issue is to resolve whether specific adverse drug reaction stem from variation in the individual genome of a patient, from drug/environment cocktail effects, or both. Here it is essential to perform temporal analysis of the records for identification of ADRs directly from the free text narratives describing patient disease trajectories over time. We can then characterize the similarity of ADR profiles of approved drugs using drug-ADR networks and report on the relationship between the chemical similarity of drugs and their ADRs. Given the availability of longitudinal data covering long periods of time we can extend the temporal analysis to become more life-course oriented. We describe how the use of an unbiased, national registry covering 6.2 million people from Denmark can be used to construct disease trajectories which describe the relative risk of diseases following one another over time. We show how one can "condense" millions of trajectories into a smaller set which reflect the most frequent and most populated ones. This set of trajectories then represent a temporal diseaseome as opposed to a static one computed from non-directional comorbidities.

References
- Using electronic patient records to discover disease correlations and stratify patient cohorts. Roque FS et al., PLoS Comput Biol. 2011 Aug;7(8):e1002141.
- Mining electronic health records: towards better research applications and clinical care. Jensen PB, Jensen LJ, and Brunak S, Nature Reviews Genetics, 13, 395-405, 2012.
- A nondegenerate code of deleterious variants in mendelian Loci contributes to complex disease risk. Blair DR, Lyttle CS, Mortensen JM, Bearden CF, Jensen AB, Khiabanian H, Melamed R, Rabadan R, Bernstam EV, Brunak S, Jensen LJ, Nicolae D, Shah NH, Grossman RL, Cox NJ, White KP, Rzhetsky A. Cell. 155, 70-80, 2013.
- Dose-specific adverse drug reaction identification in electronic patient records, Robert Eriksson R, Werge T, Jensen LJ, Brunak S. Drug Safety, to appear 2014.
- Temporal disease trajectories condensed from population-wide registry data covering 6.2 million patients.
- Jensen AB, Moseley PL, Oprea TI, Ellesøe SG, Eriksson R, Schmock H, Jensen PB, Jensen LJ, Brunak S. Under revision 2014.

Invited talk

Anne Ferguson-Smith
University of Cambridge, UK

Abstract
Environmental factors during early life are critical for the later metabolic health of the individual and of future progeny. In a mouse model of maternal caloric restriction during pregnancy the metabolic physiology of offspring over two generations is affected, including via paternal transmission to the second generation. Here we explore whether the paternal experience of in utero undernutrition, with an impact on the health of his offspring, is an epigenetically inherited memory transmitted via his sperm methylome?

Invited talk

Tanya Vavouri
Institute of Predictive and Personalized Medicine of Cancer (IMPPC), Barcelona, Spain

The small RNA content of human sperm

Abstract
At the end of mammalian sperm development, sperm cells expel most of their cytoplasm and dispose of the majority of their RNA. Yet, hundreds of RNA molecules remain in mature sperm. The biological significance of the vast majority of these molecules is unclear. To better understand the processes that generate sperm small RNAs and what roles they may have, we sequenced and characterized the small RNA content of sperm samples from two human fertile individuals. We detected nearly two hundred microRNAs, some of which are highly abundant. The most abundant class of small non-coding RNAs in sperm are PIWI-interacting RNAs (piRNAs). The most abundant sperm piRNAs target LINE1 retrotransposons. This is in agreement with previous experiments showing that PIWI is required for LINE1 repression in adult testes. Our results show that LINE1 targeting piRNAs remain in mature sperm. Surprisingly, we found that human sperm cells contain piRNAs processed from pseudogenes. Clusters of piRNAs from human testes contain pseudogenes transcribed in the antisense strand and processed into small RNAs. Several human protein-coding genes contain antisense targets of pseudogene-derived piRNAs in the male germline and these piRNAs are still found in mature sperm. Our study provides the most extensive dataset and annotation of human sperm small RNAs to date and is a resource for further functional studies on the roles of sperm small RNAs. In addition, we propose that pseudogene-derived human piRNAs may regulate expression of their parent gene in the male germline.

Invited talk

Mihaela Zavolan
Biozentrum, Switzerland

Decoding miRNA-based regulation with low and high-throughput data

Abstract
miRNAs are small RNA regulators that guide Argonaute proteins to target mRNAs, decreasing their stability and protein output. Although it has been estimated that the majority of human genes are regulated by miRNAs and that a miRNA has on average hundreds of mRNA targets, experimental identification as well as computational prediction of miRNA targets remain challenging. Experimental methods based on the crosslinking and immunoprecipitation (CLIP) of Argonaute proteins have recently been proposed, but computational methods are needed to infer which miRNA guided the Argonaute interaction with these target sites. We developed a model to comparatively evaluate the likelihood of interaction of individual miRNAs with individual Argonaute binding sites and inferring the model's parameters from Ago2 CLIP data we found that they largely reflect previously known principles of miRNA-target interaction. Application of this model to various Ago2 CLIP data sets enabled us to identify a substantial number of miRNA binding sites that are non-canonical, meaning that they are not perfectly complementary to the miRNA 5' end. Through the analysis of the kinetics with which mRNAs respond to miRNA over-expression we determined that the loading of miRNAs into Argonaute proteins limits the speed of miRNA-dependent gene regulation.

Short talk

Lorenzo Pasquali

Lorenzo Pasquali (presenting author), Kyle J. Gaulton, Santiago A. Rodríguez-Seguí, Loris Mularoni[1], Irene Miguel-Escalada, Idem Akerman, Juan J. Tena, Ignasi Mor·n, Carlos Gómez-Marín, Martijn van de Bunt, Joan Ponsa-Cobas, Natalia Castro, Takao Nammo, Inés Cebola, Javier García-Hurtado, Miguel Angel Maestro, FranÁois Pattou, Lorenzo Piemonti, Thierry Berney, Anna L. Gloyn, Philippe Ravassard, José Luis GÛmez Skarmeta, Ferenc Müller, Mark I. McCarthy, Jorge Ferrer

Genomic Programming of Beta-cells Laboratory, Institut d'Investigacions Biomédiques August Pi i Sunyer (IDIBAPS), Spain. CIBER de Diabetes y Enfermedades Metabólicas Asociadas (CIBERDEM), Spain, The Wellcome Trust Centre for Human Genetics, University of Oxford, United Kingdom, Oxford Centre for Diabetes, Endocrinology, and Metabolism, Churchill Hospital, United Kingdom. Oxford NIHR Biomedical Research Centre, Churchill Hospital, United Kingdom. School of Clinical and Experimental Medicine

Pancreatic islet enhancer clusters enriched in type 2 diabetes risk-associated variants

Abstract

Type 2 diabetes affects over 300 million people, causing severe complications and premature death, yet the underlying molecular mechanisms are largely unknown. Pancreatic islet dysfunction is central for type 2 diabetes pathogenesis, and therefore understanding islet genome regulation could provide valuable mechanistic insights. We have now mapped and examined the function of human islet cis-regulatory networks. We identify genomic sequences that are targeted by islet transcription factors to drive islet-specific gene activity, and show that most such sequences reside in clusters of enhancers that form physical 3D chromatin domains. We find that sequence variants associated with type 2 diabetes and fasting glycemia are enriched in these clustered islet enhancers, and identify trait-associated variants that disrupt DNA-binding and islet enhancer activity. Our studies illustrate how islet transcription factors interact functionally with the epigenome, and provide systematic evidence that dysregulation of islet enhancers is relevant to the mechanisms underlying type 2 diabetes.

Short talk

Tuncay Baubec

Tuncay Baubec (presenting author) and Dirk Schübeler

Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058 Basel Switzerland

DNA methylation: (epi-)genomic views on pattern generation and interpretation

Abstract

Eukaryotic gene regulation is a highly orchestrated process that relies on interactions between regulatory proteins and the chromatinized genome. While recent genome-wide advances helped to broaden our knowledge on genomic distribution of regulatory proteins and epigenetic modifications, the molecular mechanisms related to establishment and interpretation of the identified patterns remain widely unexplored. This, for example, is the case for DNA methylation, a prevalent epigenetic modification involved in transcriptional repression. In mammals, disturbed DNA methylation results in impaired development and is associated with cancer. However, the contribution of genomic DNA methylation patterns to gene regulation remains widely unclear. We aim to address this by investigating how genomic DNA methylation patterns are generated and interpreted by the interplay of readers, writers and erasers. We have already profiled genomic targeting of the entire methyl-CpG-binding domain (MBD) family of proteins involved in readout of DNA methylation. Profiles of various MBD protein variants, including disease relevant mutations, and in various cell types revealed a linear dependency on local DNA methylation density that directs MBD proteins to methylated regulatory regions in the genome [1]. In unpublished work we have investigated the genomic binding preferences of DNA methyltransferases (DNMTs), the writers of this epigenetic mark. These profiles, in combination with functional analysis and computational modelling, allowed to uncover the binding logic of these relevant proteins to genomic sites. This identified histone modifications to play a central role in specifying DNA methylation patterns through direct recruitment of DNMTs to distinct genomic sites. I will present the quantitative framework that exposes genomic binding of methyl-CpG readers and writers, and identifies cis and trans components that determine target preference. 1. T. Baubec, R. Ivanek, F. Lienert, D. Sch¸beler, Methylation-Dependent and -Independent Genomic Targeting Principles of the MBD Protein Family, Cell 153, 480-492 (2013).

Invited talk

Ewan Birney
EMBL-EBI and Elixir

EMBL-EBI and Elixir

Abstract
Molecular biology is now a leading example of a data intensive science, with both pragmatic and theoretical challenges being raised by data volumes and dimensionality of the data. These changes are present in both “large scale” consortia science and small scale science, and across now a broad range of applications – from human health, through to agriculture and ecosystems. All of molecular life science is feeling this effect.

This shift in modality is creating a wealth of new opportunities and has some accompanying challenges. In particular there is a continued need for a robust information infrastructure for molecular biology. This ranges from the physical aspects of dealing with data volume through to the more statistically challenging aspects of interpreting it. Drawing on recent experience I will explore both the “blue collar” challenges of data volume and the “white collar” challenges of interpretation.

I will end with the serendipitous invention of using DNA for an entirely different reason – as a long-time horizon digital archiving material. I will describe this method and some of its benefits (as well as a few downsides) and explain how a future culture in 10,000 years time may still be able to read all of Shakespeare’s sonnets – and perhaps much more.

Invited talk

Eileen E. Furlong
EMBL, Heidelberg, Germany

Temporal regulation of transcription during development: It’s all in the timing

Abstract
One of the central challenges in biology is to understand how the genome is utilized to give rise to diverse cell types. Embryonic development occurs through the progressive restriction of cell fates, from pluripotent fields of cells to complex organs and tissues. This process requires a directed progression through interlinked regulatory states, each defined by the total set of active transcription factors. At each stage of development, the combined inputs of signalling and transcriptional networks regulate the expression of specific sets of genes that drive the transition to the next, often more specialized, state. Understanding how the underlying cis-regulatory networks produce spatial and temporal gene expression is therefore an essential step towards deciphering metazoan development and ultimately evolutionary change.
While genetic studies have identified a growing number of essential transcription factors (TFs) required for cell fate specification, little is known about the mechanisms by which these regulators function. Conversely, recent global approaches assaying TF binding enable the location and even combinatorial occupancy of enhancers to be experimentally measured at specific stages of development, at a genome-wide scale. A current major challenge is to interpret these TF binding data in terms of their resulting spatio-temporal cis-regulatory activity. Our work attempts to bridge this gap, by integrating genetic, genomic and computational approaches to understand how transcriptional networks drive cell fate selection, using mesoderm specification in Drosophila as a well-defined model system. I will present our current status on an on-going effort to integrate information on TF occupancy, changes in chromatin state and 4D enhancer interactions to understand how precise transcriptional regulation is controlled during embryonic development.

References
- Spitz F and Furlong EE (2012). Transcription factors: From enhancer binding to developmental control. Nature Reviews Genetics, Sep;13(9):613-26.
- Junion G*, Spivakov M*, Girardot C, Braun M, Gustafson EH, Birney E, Furlong EE (2012). A transcription factor collective defines cardiac cell fate and reflects the developmental history of this cell lineage. Cell, Feb 3;148(3):473-86.
- Bonn S*, Zinzen RP*, Girardot C*, Gustafson EH, Gonzalez AP, Delhomme N,
Ghavi-Helm Y, Wilczynski B, Riddell A and Furlong EE. (2012). Tissue specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. Nature Genetics, Jan 8;44(2):148-56.
- Wilczynski B, Furlong EE. (2010) Dynamic CRM occupancy reflects a temporal map of developmental progression. Mol Syst Biol. 6, 383 (22 June 2010).

Invited talk

Nick Luscombe
Joint appointments: University College London and Cancer Research UK London Research Institute, UK

Guardians of the transcriptome: how hnRNP C protects you from unwanted axons

Abstract
There are ~650,000 Alu elements in transcribed regions of the human genome. These elements contain cryptic splice sites, so they are in constant danger of aberrant incorporation into mature transcripts. Despite posing a major threat to transcriptome integrity, little is known about the molecular mechanisms preventing their inclusion. Here, we present a mechanism for protecting the human transcriptome from the aberrant exonization of transposable elements. Quantitative iCLIP data show that the RNA-binding protein hnRNP C competes with the splicing factor U2AF65 at many genuine and cryptic splice sites. Loss of hnRNP C leads to formation of previously suppressed Alu exons, which severely disrupt transcript function. Minigene experiments explain disease-associated mutations in Alu elements that hamper hnRNP C binding. Thus, by preventing U2AF65 binding to Alu elements, hnRNP C plays a critical role as a genome-wide sentinel protecting the transcriptome. The findings have important implications for human evolution and disease.

list of all abstracts

1. Yassen Assenov
Predicting primary origin of metastatic samples based on DNA methylation profiles

2. Anaïs Baudot
Inverse Comorbidity between Central Nervous System Disorders and Cancers interpreted by Transcriptomic Meta-analyses

3. Olga Bogatyrova
Mutations in regulators of the epigenome and their effects on the DNA methylome

4. Eduard Casas Masnou
miRNA regulation by DNA methylation

5. Elisabeth Castellanos
Hereditary cancer sub-exome sequencing approach for genetic diagnostics

6. Anna Díez
regioneR: an R package for the management and comparison of genomic regions

7. Julien Douet
Epigenomic analysis reveals two disctinct type of macroH2A-containing chromatin

8. Gabrijela Dumbovic
Transcriptional and epigenetic characterization of a tandemly repeated pericentromeric element in colorectal cancer

9. Miquel Duran-Frigola
Relating Chemical Fragments to Disease

10. Johannes Engelken
The disruption of trace element homeostasis due to aneuploidy as a unifying theme in the etiology of cancer

11. Guerau Fernandez
Towards understanding the effect of DNA demethylation on transcription regulation

12. Mirko Francesconi
The impact of genetic variation on gene expression dynamics in C. elegans development

13. Francisco Fuster Tormo
Atransferrinemia: A ultra-rare genetic anaemia.

14. Bernat Gel Moreno
Clusters of differentially expressed genes caused by regional genomic and epigenomic alterations help identify Malignant Peripheral Nerve Sheath Tumor drivers

15. Juan R Gonzalez
Detection of cancer signatures by integrating different features from omic data using regularized generalized canonical correlation

16. Laura Isus Díaz
A Network approach to Spinal Cord Injury

17. Samira Jaeger
A network biology approach to identify breast cancer drug targets and drug combinations

18. Rory Johnson
Genomic recycling: A new hypothesis for evolution of long noncoding RNA functions

19. David Juan
Late replicating CNVs as a source of new genes

20. Petra Kaferle
SGA: A versatile high throughput platform for yeast functional genomics in France

21. Pascal Kahlem
Scientific Network Management, SL An SME to support project drafting and management of scientific networks

list of all abstracts

22. Ivan Kulakovskiy
mTOR mRNA targets and transcription-translation rendezvous: a peep through sequence analysis keyhole

23. Pierre Luisi
Recent positive selection targets the center of the human protein-protein interaction network

24. Ettore Luzi
Role of microRNAs-Transcription Factors Gene Regulatory Networks dynamics in the onset and progression of tumors: the MEN1 syndrome as model

25. Vsevolod Makeev
Large non-redundant collection of binding motifs for human transcription factors

26. Izaskun Mallona
Towards understanding the role of repeat elements in genome biology: a semantic approach

27. Yulia Medvedeva
Effects of cytosine methylation on transcription factor binding sites

28. Arcadi Navarro i Cuartiellas
The EGA-CRG Elixir project

29. Chiara Pallara
Understanding CFC (cardio-facio-cutaneous) syndrome by molecular dynamics simulation

30. Vera Pancaldi
Evolutionary importance of expression variability and its application to stratification of CLL cases

31. Sílvia Pérez-Lluch
Transcriptional activation without chromatin marking in developmentally regulated genes

32. Daniel Rico Rodríguez
Epigenomic plasticity of human neutrophils and monocytes in cord and peripheral blood*

33. Martin Schaefer
Protein conservation and variation suggest mechanisms of cell type-specific modulation of signaling pathways

34. Christoph Schlaffner
Similarity scoring for modification and coding variant discovery

35. Maria Shutova
Induction of pluripotency leaves no reprogramming-specific trace in the human somatic cells-derived iPS cells

36. Mónica Suelves
Dynamic changes in DNA methylation during muscle-lineage commitment and terminal differentiation

37. David Torrents
Accurate characterization of complex structural variation in cancer by using a reference-free approach

38. Leonid Uroshlev
Ion prediction in apo-form of proteins

39. Ignacio Vazquez García
Probabilistic reconstruction of subclonal diversity in adaptive evolution

40. Ilya Vorontsov
Comparison and classification of transcription factor binding sites models with application to functional annotation of regulatory sequence variants

41. Dieter Weichenhan
Tagmentation-based whole genome bisulfite sequencing for very low input DNA amounts

list of fast track talks

Fast Track Session I

Thursday 27 March at 12.50

Yassen Assenov
Predicting primary origin of metastatic samples based on DNA methylation profiles

Anaïs Baudot
Inverse Comorbidity between Central Nervous System Disorders and Cancers interpreted by Transcriptomic Meta-analyses

Olga Bogatyrova
Mutations in regulators of the epigenome and their effects on the DNA methylome

Anna Díez
regioneR: an R package for the management and comparison of genomic regions

Mirko Francesconi
The impact of genetic variation on gene expression dynamics in C. elegans development

Bernat Gel Moreno
Clusters of differentially expressed genes caused by regional genomic and epigenomic alterations help identify Malignant Peripheral Nerve Sheath Tumor drivers

Juan R Gonzalez
Detection of cancer signatures by integrating different features from omic data using regularized generalized canonical correlation

Samira Jaeger
A network biology approach to identify breast cancer drug targets and drug combinations

Fast Track Session II

Thursday 27 March at 14.45

Rory Johnson
Genomic recycling: A new hypothesis for evolution of long noncoding RNA functions

David Juan
Late replicating CNVs as a source of new genes

Pierre Luisi
Recent positive selection targets the center of the human protein?protein interaction network

Arcadi Navarro i Cuartiellas
The EGA-CRG Elixir project

Sílvia Pérez-Lluch
Transcriptional activation without chromatin marking in developmentally regulated genes

Martin Schaefer
Protein conservation and variation suggest mechanisms of cell type-specific modulation of signaling pathways

Christoph Schlaffner
Similarity scoring for modification and coding variant discovery

David Torrents
Accurate characterization of complex structural variation in cancer by using a reference-free approach

Ignacio Vazquez García
Probabilistic reconstruction of subclonal diversity in adaptive evolution

Poster - 1

Yassen Assenov

(Yassen Assenov), Christoph Plass

Division of Epigenomics and Cancer Risk Factors, German Cancer Research Center, Heidelberg, Germany

Predicting primary origin of metastatic samples based on DNA methylation profiles

Abstract

Aberrant methylation is a common epigenetic event accompanying tumor progression. Frequently, short hypermethylated regions are observed against a background of global loss of methylation in malignant cell populations. Previously, we showed that DNA methylation profiles could prove valuable for translational purposes by identifying the tumor type origin of cancers of unknown primary origin. To this end, we extended this observation by comparing a selection of linear models for predicting origin of metastatic samples based on genome-wide methylation measurements. For training and validation we used a large collection of primary and metastatic tumor samples from The Cancer Genome Atlas, as well as several independent studies. The accuracy of many models is above 95 percent - an encouraging result for their clinical application. In addition to estimating probabilities for tumor of origin for a given metastatic sample, we provide interpretable results by listing the major differential methylation events that lead to the decisions made by the classifier.

Poster - 2

Anaïs Baudot

Kristina Ibánez*[1], César Boullosa*[1], Rafael Tabarés-Seisídos[2], Alfonso Valencia[1] and Anaïs Baudot[3] (presenting author)

[1] Structural Biology and Biocomputing Programme, Spanish National Cancer Research Centre (CNIO), Madrid, Spain. [2] Department of Medicine, University of Valencia, CIBERSAM, INCLIVA, Valencia, Spain. [3] Marseilles Institute of Mathematics (I2M), CNRS-AMU, Marseilles, France

Inverse Comorbidity between Central Nervous System Disorders and Cancers interpreted by Transcriptomic Meta-analyses

Abstract

Epidemiological evidences point to a lower-than-expected risk of cancer in patients with certain Central Nervous System (CNS) disorders. Such inverse comorbidity could arise, for instance, from environmental factors or drug treatments or be related to disease diagnosis. We and others hypothesized that the inverse comorbidities could also be driven by genetic factors common to both sets of complex diseases (Roe et al. Neurology (2013), Tabarés-seisdedos et al. Lancet Oncology (2011)). In this context, we recently published the first evidence of gene expression deregulations in opposite directions in inversely comorbid diseases (Ibánez et al. Plos Genetics (2014 Feb 20;10(2))). I will describe in this presentation the results obtained by meta-analyzing gene deregulations in 3 CNS disorders (Alzheimer's, Parkinson's and Schizophrenia) and 3 cancer types (Lung, Prostate, Colorectal). Strikingly, the comparison of the deregulations in both sets of diseases showed that a significant number of genes and pathways up-regulated in CNS disorders are down-regulated in Cancers, and vice-versa. For instance, the Pin1 gene and the p53 pathway, which were previously suggested to be implied in inverse comorbidity because of their key role in cell fate decision (Roe et al. Neurology (2013)), are down-regulated in the CNS disorders and up-regulated in Cancers. Novelties include the proteasome pathway - up-regulated in the 3 cancer types while down-regulated in the 3 CNS disorders, and the metallothioneins, MT1X, MT2A and MT1M - down-regulated in the 3 cancer types while up-regulated in the 3 CNS disorders. Overall, our analysis points to specific molecular processes the up-regulation of which could increase the risk of CNS disorders while reducing the risk for Cancer, while the down-regulation of another set of molecular processes would be implied in a reduced risk of CNS disorders while increasing the risk of Cancers. I will finally present my future projects grounded on this initial analysis, considering mutation and polymorphism informations, and the integration of all inverse comorbidities-related data in large-scale protein-protein and drug-target interaction networks for the initiation of drug repurposing strategies.

Poster - 3

Olga Bogatyrova

Olga Bogatyrova, Yassen Assenov, Christoph Plass

German Cancer Research Center (DKFZ), Germany

Mutations in regulators of the epigenome and their effects on the DNA methylome

Abstract

Genome wide profiling for genetic alterations in cancer has identified mutations in genes that are associated with epigenetic programming of genomes for DNA methylation patterns, histone modifications patterns and the positioning of nucleosomes. Here we describe a systematic evaluation of the available cancer genome profiling data established by large international consortia, in order to identify recurrently mutated genes or pathways. We establish a list of about 700 genes that are involved in either establishing an epigenetic mark (writers), modify these epigenetic marks (editors), or translate the epigenetic mark into a molecular signal (reader). Using currently available genome-wide datasets on genetic and epigenetic alterations in cancers (www.ICGC.org), we describe the distribution of mutations including gene mutations, copy number alterations, promoter methylation, and potential targeting by miRNA and lncRNA expression in a pan-cancer study. We have classified genes as potential oncogenic or those with tumor-suppressor function based on the location of mutations relative to functional domains and their frequency. Mechanisms of deregulation of epigenetic genes/pathways occur in a tumor-type specific manner. We have created a panel of 50 epigenetic genes (including: DNMTs, histones (H3F3A, HIST1H3B), histone editors (KDM5C, KDM6A) and writers (MLLs, SETD2, EZH2, ATM) that can promote epigenetic changes in cancer. Correlative analysis of publicly available methylation data (450K Illumina arrays from ICGC and TCGA web-site) with information on deregulated epigenetic driver genes provides a source for the identification and link between methylation groups and deregulated epigenetic genes/pathway. This bioinformatics approach to defects in regulators of the epigenome will help to understand mechanisms leading to distinct epigenetic patterns and will allow the molecular validation of defined correlation in experimental settings.

Poster - 4

Eduard Casas Masnou

Eduard Casas (presenting author), Guerau Fernéndez, Joaquim Custodio, Miguel A. Peinado, Tanya Vavouri

Institute of Predictive and Personalized Medicine of Cancer (IMPPC), Spain

miRNA regulation by DNA methylation

Abstract

Epigenetic alterations such as aberrant DNA methylation of the human genome are common in human diseases, especially cancer. Yet we do not fully understand the significance of the altered DNA methylation patterns and how this affects different types of genes. Although miRNAs have also been reported to be dysregulated in various cancer types, much remains unknown about their regulation. To gain insight into the effect of DNA methylation in the regulation of miRNA expression, we analyze a dataset of high-throughput small RNAs from human cell treated with 5-aza-2'-deoxycytidine, a demethylating agent. By measuring DNA methylation and microRNA expression levels in both treated and untreated cells, we aim to identify changes that are direct consequences of loss of DNA methylation.

Poster - 5

Elisabeth Castellanos

E. Castellanos[1] (presenting author), B. Gel[1], S. Santín[2], I. Rosas[1], P. Barrero(2), R. Pluvinet[2], J. Velasco[2], E. Terribas[1], E. Tornero[3], A Lopez-Doriga[3], L. Feliubadaló[3], M. Perucho[1], L. Sumoy[3], I. Blanco[3], G. Capellá[3], C. Làzaro[3], E. Serra[1]

[1] IMPPC-ICO Joint Program on Hereditary Cancer, Institute of Predictive and Personalized Medicine of Cancer (IMPPC), Spain. [2] Genomics Laboratory, Institut de Medicina Predictiva i Personalitzada del Càncer (IMMPC) Badalona, Barcelona, Spain. [3] Hereditary Cancer Program, Catalan Institute of Oncology (ICO), L'Hospitalet de Llobregat, Barcelona, Spain

Hereditary cancer sub-exome sequencing approach for genetic diagnostics

Abstract

We have developed a Next Generation Sequencing-based strategy, the Hereditary Cancer Sub-Exome Panel, for the comprehensive genetic diagnostics of known genes involved in hereditary cancer and RASopathies. Our strategy responds to the need to simplify the methodological procedures required for dealing with many different genes and genetic conditions, while maintaining or improving testing performance: sensitivity, specificity, turnaround time and cost-effectiveness. Furthermore, it aims to address major challenges such as genetic heterogeneity, uncertain clinical diagnostics or the presence of somatic mosaicism often encountered in our cases. Two versions of our SureSelect custom capture designs have been developed. The first version targeted the coding regions and intron-exon boundaries of 106 genes. DNA samples from 24 individuals with hereditary colorectal cancer, hereditary breast and ovarian cancer and neurofibromatoses were captured and sequenced using Ion Torrent PGM. The first version of the enrichment design allowed us to detect all known pathogenic and non-pathogenic variants in our samples, including deletions affecting more than one exon, except for point mutations located in homopolymer regions, frequently miss-called by Ion Torrent. After data analysis, an improved version was designed covering 125 genes and 46 cancer risk associated SNPs. The same 24 samples have been captured with version 2 and sequenced on MiSeq. With this data, we have compared the sequencing platform performance and checked the improvements in the capture design. We are performing a final validation with a larger set of samples to confirm the necessary requirements to bring this strategy to the clinics. Data available so far show that our approach leaves less than 10 out of 2525 exons uncovered and outperforms a commercial whole exome kit. In order to ensure the quality standards needed for genetic diagnostics we are developing a custom data analysis pipeline.

Poster - 6

Anna Díez

Anna Díez, Bernat Gel, Marcus Buschbeck, Miguel A. Peinado, Eduard Serra, Roberto Malinverni

Institute of Predictive and Personalized Medicine of Cancer (IMPPC), Spain

regioneR: an R package for the management and comparison of genomic regions

Abstract

Management and analysis of regional genomic information is increasingly important in biological studies, either as the main outcome of an analysis or as an additional layer in a dataset. Statistically assessing the spatial relations between region sets is a fundamental part of their analysis, but so far, the available options are lacking or limited in scope.

Here we present regioneR, an R package build on top of the Bioconductor's genomic regions functions with two main aims: (1) to offer a basic set of region manipulating functions with a simple interface and (2) to create a statistical framework based on customizable permutation tests to assess the relations between genomic region sets.

The core part of the package is a permutation test specifically designed to evaluate the relations between sets of genomic regions. All functions are prepared to work with a genome and a mask, either custom or automatically loaded from BSGenome, and custom masks can be used to deal with complex analysis. The randomization and evaluation functions are fully customizable and users can define their own functions. For example, in addition to the included evaluation functions dealing with overlaps, distances and base-level values, it is possible to evaluate other relevant information as GC content, methylation levels or position within the chromosomes. It is even possible to change the randomization process to take into account the structural complexity of the genome using alternative randomization strategies. In addition, the included plotting functionality creates publication-ready graphics representing the results of permutation tests. Besides its easy-to-use design, regioneR is a customizable and powerful tool to manage and analyze sets of regions, and a useful addition to the NGS and genome wide analysis toolbox.

Poster - 7

Julien Douet

Julien Douet (presenting author)*, Roberto Malinverni*, David Corujo and Marcus Buschbeck

Institute of Predictive and Personalized Medicine of Cancer (IMPPC), Spain

Epigenomic analysis reveals two disctinct type of macroH2A-containing chromatin

Abstract

Although two decades have elapsed since their initial discovery, the function of structural chromatin components containing macro domains remains a mystery. Facultative incorporation of any of the so-called macroH2A histone variants places a metabolite-binding macro domain on the structural unit of chromatin. The expression and function of macroH2A variants is tightly linked to cellular differentiation. They contribute to developmental processes; act as tumor suppressors in cancer and form part of the epigenetic barrier inhibiting somatic cell reprogramming. In the context of gene transcription both repressive and activating functions have been attributed to macroH2A. But how macroH2A acts on the molecular level remains unknown. Here, we present the analysis of four ChIPseq experiments realized with the two macroH2A forms in two different human cell lines: HCT116 and HepG2. The results refined the idea we had about the macroH2A positioning. Indeed, macroH2A variants do not only forms large domains associated to H3K27me3 as previously described, but also some very intriguing peaks sitten on small sequences, showing very high amounts of macroH2A and a strong association to H3K9me3.

Poster - 8

Gabrijela Dumbovic

Gabrijela Dumbovic, Johanna K. Samuelsson[1], Sergio Alonso, Sonia Forcales and Manuel Perucho

Institute of Predictive and Personalized Medicine of Cancer (IMPPC), Spain

[1] Sanford-Burnham Medical Research Institute, La Jolla, California

Transcriptional and epigenetic characterization of a tandemly repeated pericentromeric element in colorectal cancer

Abstract

Somatic DNA hypomethylation especially at repetitive elements has been described in several types of cancer. We proposed a “wear & tear” model linking aging with cancer through the gradual demethylation of pericentromeric repetitive sequences and the consequent increased risk for mitotic errors and chromosomal aberrations. We identified a frequently hypomethylated region in colon tumors that belongs to the family of a moderately repetitive pericentromeric element called SST1. This element shows demethylation in approximately 20% of colorectal cancer patients. While the majority of tumors exhibit gradual demethylation, conforming to the “wear & tear” model, a few patients displayed a more severe demethylation that turned to be age-independent. In colon cancer cell lines with different SST1 methylation profiles, we observed that hypomethylated SST1 co-occurs with high levels of H3K27me3 and lower levels of H3K9me3, while methylated SST1 correlates with low H3K27me3 and high H3K9me3 levels. Furthermore, induced demethylation of SST1 by 5-aza-2’-deoxycytidine treatment leads to an increase in H3K27me3 and a reduction of H3K9me3 levels at the SST1 element. The data suggest that when SST1 elements are demethylated, they are reprogrammed to a more relaxed and plastic heterochromatin. In agreement with this observation, we have detected increased transcription of this element in 5-aza-2’-deoxycytidine treated cells. We found that the non-coding RNA originating from SST1 elements is predominantly non-polyA, and transcribed by RNA polymerase II. We investigated the subcellular localization of SST1 RNA and we found that it is mainly associated to chromatin. This data further reinforces a non-coding function of SST1 RNA. Our data shows that chromatin reprogramming of the SST1 pericentromeric region associates with some colorectal cancers. We are now working to clarify the mechanism causing SST1 demethylation and the specific cross-talk between H3K27me3 and DNA methylation regulating SST1 expression. Further studies are also needed to find out whether SST1 non-coding RNA could have a role in the disruption of genome integrity.

Poster - 9

Miquel Duran-Frigola

Miquel Duran-Frigola (presenting author), David Rossell and Patrick Aloy

[1,3] Joint IRB-BSC Program in Computational Biology, Institute for Research in Biomedicine (IRB) Barcelona, c/Baldiri Reixac 10-12, 08028 Barcelona, Spain. [2] Biostatistics and Bioinformatics Unit, IRB Barcelona, Spain. [2] Department of Statistics, University of Warwick, Coventry, UK. [3] Institució Catalana de Recerca i Estudis Avançats (ICREA), Pg. Lluís Companys 23, 08010 Barcelona, Spain

Relating Chemical Fragments to Disease

Abstract

Efforts to compile the phenotypic effects of drugs and environmental chemicals offer the opportunity to adopt a chemo-centric view of human health. Small molecules play a crucial role in biological processes and, in many cases, exploring their structural features helps to predict their behavior. This is of particular value when the exact molecular and physiological events are unknown. We explored thousands of chemicals and analyzed their relationship with respect to a comprehensive collection of adverse and beneficial outcomes. Our study included molecules that are related to the etiology of 934 conditions and also those that are useful to treat as many as 835 diseases. In our chemo-centric disease models, we balanced interpretation and prediction efficiency through reporting molecular fragments associated with each of the phenotypic effects. Using these fragments, we built plausible predictors for approximately 23% of the diseases. Among fragment-disease relationships, we found privileged and liable structures that could inform the drug discovery process. Also, our observations reveal that not all disease classes are equally suitable for a chemo-centric view, although we anticipate that consistent, systematic amassment of chemical records will increase our ability to build disease models in the future. Overall, while we believe that systems biology will ultimately provide a full understanding of mechanisms of action, we propose that the rapid accumulation of data on small molecules could readily facilitate the anticipation of the phenotypic effects and the design of more effective compounds. In this regard, our analysis is the most exhaustive study to date of the spectrum of effects that small molecules have on our health. Moreover, it allows for relating phenotypes based on their underlying chemistry, shedding light on the chemical basis of human diseases.

Poster - 10

Johannes Engelken

Johannes Engelken[1,2,3], Matthias Altmeyer[4], Renty B. Franklin

[1] Institute of Evolutionary Biology (CSIC – Universitat Pompeu Fabra), 08003 Barcelona, Spain. [2] Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany. [3] Bioinformatics and Genomics, Center for Genomic Regulation (CRG), 08003 Barcelona, Spain. [4] Chromosome Stability and Dynamics Group, Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Blegdamsvej 3B, 2200 Copenhagen, Denmark. [5] Department of Oncology and Diagnostic Sciences, University of Maryland Dental School and The University of Maryland Greenebaum Cancer Center, Baltimore, 21201 Maryland, USA.

The disruption of trace element homeostasis due to aneuploidy as a unifying theme in the etiology of cancer

Abstract

While decades of cancer research have firmly established multiple “hallmarks of cancer”, cancer’s genomic landscape remains enigmatic – in particular the phenomenon of aneuploidy – gains and losses of large genomic regions. An additional characteristic of many different cancers is the deregulation of the homeostasis of trace elements, such as copper, zinc and iron. Concentrations of copper are markedly increased in cancer tissue and the blood plasma of cancer patients, while zinc levels are typically decreased. Here we discuss the hypothesis that the disruption of trace element homeostasis and the phenomenon of aneuploidy might be linked. Our tentative analysis of genomic data from diverse tumor types mainly from The Cancer Genome Atlas (TCGA) project suggests that gains and losses of metal transporter genes occur frequently and correlate well with transporter gene expression levels. Candidate driving events may include, among others, the gains of zinc transporter genes SLC39A1 and SLC39A4 on chromosome arms 1q and 8q, respectively, and the losses of zinc transporter genes SLC30A5, SLC39A14 and SLC39A6 on 5q, 8p and 18q. The recurrent gain of 3q might be associated with the iron transporter gene TFRC and the loss of 13q with the copper transporter gene ATP7B. By altering cellular trace element homeostasis such events might contribute to the initiation of the malignant transformation. Consistently, it has been shown that zinc affects a number of the observed hallmark characteristics including DNA repair, inflammation and apoptosis. We term this model the “aneuploidy metal transporter cancer” (AMTC) hypothesis and find it compatible with the cancer-promoting role of point and focal mutations in established tumor suppressor genes and oncogenes (e.g. MYC, MYCN, TP53, PIK3CA, BRCA1, ERBB2). We suggest a number of approaches for how this hypothesis could be tested experimentally and briefly touch on possible implications for cancer etiology, metastasis, drug resistance and therapy.

Poster - 11

Guerau Fernandez

Guerau Fernàndez (presenting author), Joaquin Custodio, Miguel A. Peinado, Tanya Vavouri

Institute of Predictive and Personalized Medicine of Cancer (IMPPC), Spain

Towards understanding the effect of DNA demethylation on transcription regulation

Abstract

Regulation of gene transcription is a tightly controlled process. DNA methylation seems to play a major role in transcriptional silencing as well as in other important processes such as X chromosome inactivation, genomic imprinting and control of endogenous retroelements. Importantly, common human diseases such as cancer and more specifically leukemia are significantly associated with mutations in the DNA methylation pathway. The usage of demethylating drugs like azacytidine has been a promising approach to treat patients with malignancies such as myelodysplastic syndrome (MDS). Although effective, the mechanism by which azacytidine leads to a health benefit for MDS patients remains unknown. Because DNA methylation is found throughout the entire genome in different contexts and due to its involvement in a big number of processes, we still do not fully understand its role in gene regulation. To address this, we are carrying out a systematic study of the transcriptional consequences of genome-wide changes in DNA methylation levels in human cells.

Poster - 12

Mirko Francesconi

Mirko Francesconi[1,2] (presenting author), Ben Lehner[1,2,3]

[1] EMBL-CRG Systems Biology Unit, Centre for Genomic Regulation (CRG), Dr. Aiguader 88, 08003 Barcelona, Spain. [2] Universitat Pompeu Fabra (UPF), Dr. Aiguader 88, 08003 Barcelona, Spain. [3] Institució Catalana de Recerca i Estudis Avançats (ICREA), Pg. Lluís Companys 23, 08010 Barcelona, Spain

The impact of genetic variation on gene expression dynamics in C. elegans development

Abstract

Development of a multicellular organism and physiological responses require massive coordinated changes in gene expression across multiple cell- and tissue-types. Expression quantitative trait locus (eQTL) analysis exploits natural genetic variation to identify genomic loci that underpin gene expression differences. eQTLs have been found in multiple species, however there has been no comprehensive characterization of how sequence variants influence the complex dynamic patterns of gene expression that occur during development and in physiology. Here we describe a novel and efficient experimental design to infer gene expression dynamics from single expression profiles in multiple genotypes and apply it to characterize the impact of local (cis) and distant (trans) genetic variation on gene expression at high temporal resolution throughout a 12-hour period of the development of C. elegans.

Poster - 13

Francisco Fuster Tormo

Francisco Fuster[1], Cristina Díaz de Heredia[2], Eunice S. Edison[3], Katja Moser[4], Érica Morán[1], Rekha Athiyarath[3], Jéssica Aranda[1], Ana M. Rojas[5], Mayka Sánchez[1]

[1] IMPPC-UDGAEMH, Unit for Advanced Genetic Diagnostics for Rare Iron Metabolism Disorders. Badalona, Barelona. [2] Hospital Universitari Vall d'Hebron. Pediatric Haemathology-Oncology Service, Barcelona. [3] Christian Medical College Hospital. Haemathology department, Vellore, India. [4] Klinik f,r Kinder- und Jugendmedizin Hospital, Pediatrics and Neonatology department. Ascha?enburg, Germany. [5] Barcelona Supercomputing Center (BSC), Life Sciences department. Barcelona

Atransferrinemia: A ultra-rare genetic anaemia

Abstract

Transferrin (TF) is a glycoprotein that transports iron in the blood. Transferrin synthesis is increased in iron deficiency anaemia. Atransferrinemia (OMIM #209300, ORPHA1195) is an ultra-rare autosomal recessive disease characterized by a severe deficiency of serum transferrin, causing ineficient erythropoiesis leading to a severe hypochromic microcytic anaemia with iron overload in non-haematopoietic tissues. To date, there are only 14 cases described in the literature and only 4 have been characterized at the molecular level. Our unit for advanced genetic diagnostics for rare iron metabolism diseases is involved in the clinical, genetic, molecular and computational characterization of 6 new wordwide cases of atransferrinemia: 2 Spanish, 2 Turkish and 2 from India. The existence of these cases underlines the fact that atransferrinemia should be considered, even if it is an ultra-rare disease, as the cause of a previously unexplainable hypochromic microcytic anaemia with iron overload. An appropriate diagnosis would lead to the correct implementation of an adequate treatment to minimize the severe health complications related to iron overload. Therefore, it is critical to distinguish atransferrinemia from other hereditary anaemias thalassemia, non-sideroblastic anaemia, sideroblastic anaemia, acquired anaemias celiac disease, myelodisplastic syndrome or nutritional iron deficiency anaemia to improve patients' quality of life.

Poster - 14

Bernat Gel Moreno

Bernat Gel[1] (presenting autor), Ernest Terribas[1], Tapan Mehta[2], Josep Biayna[1], Juana Fernandez-Rodriguez[3], Conxi Lazaro[3], Eduard Serra[1]

[2] University of Alabama at Birmingham, Birmingham, Al, USA. [3] Institut Català d'Oncologia (ICO) - IDIBELL, L'Hospitalet de Llobregat

Clusters of differentially expressed genes caused by regional genomic and epigenomic alterations help identify Malignant Peripheral Nerve Sheath Tumor drivers

Abstract

One of the major clinical complications of Neurofibromatosis type 1 (NF1) patients is the development of malignant peripheral nerve sheath tumors (MPNSTs). Approximately 8-13% of NF1 patients develop MPNSTs, which are the leading cause of NF1-related mortality. MPNSTs have a poor prognosis, with a 5-year survival rate of NF1 patients with MPNST of 21%. Our aim is to identify genes and mechanisms driving the development and progression of MPNSTs by integrating different sources of genomic and epigenomic data, centered in genes and in genomic regions. The SNP-array analysis of 14 primary MPNSTs and 5 cell lines revealed a high degree of hyperploidy and LOH, and a global landscape of genomic alterations consistent with a chromoplexy phenomenon. In such an aberrant genome, we expected that regional genomic alterations would heavily influence gene expression. Thus, in addition to study the differential gene expression between benign and malignant tumors, we considered the genomic position of differentially expressed genes. We identified groups of frequently up-regulated or down-regulated genes clustering together in specific genomic regions, what we call transcriptional imbalances (TIs). We found that many TIs can be explained by the underlying copy-number alterations detected. However, down-regulated TIs only partially correlate with copy-number losses, what prompted us to study their association with regional epigenetic regulation. We can conclude that an important part of the differential expression seen in MPNSTs is determined by regional alterations. Analogously to the fact that not every gene within a somatic copy-number altered region would be expected to be a driver, we show that TIs are capturing an important part of differentially expressed passenger genes. Following this view, we have explored different ways of using TIs in an integrative analysis of genomic MPNST data to uncover interesting genes and mechanisms for tumor progression, which are currently under experimental validation.

Poster - 15

Juan R Gonzalez

Juan R González (presenting author), Mikel Esnaola, Alejandro Càceres, Carles Hernéndez-Ferrer, Marcos López, Luis Pérez-Jurado

[1-5] Center for Research in Environmental Epidemiology (CREAL), Spain. [6] University of Pompeu Fabra (UPF), Spain

Detection of cancer signatures by integrating different features from omic data using regularized generalized canonical correlation

Abstract

The main goal of personalized medicine is to use medical models to customize health care by clustering individuals based on particular profiles. The use of 'omic' data has played a major role in these profiling efforts. Data integration procedures are required before performing such clustering methods. However, data integration has to face the challenge of dealing with the huge dimensionality of multi-modal data while maximizing their individual information. Several techniques have been applied to reduce the dimension of the space of predictors in 'omic' data, but none of them guarantees the interpretability of the results, or even biological sense of the latent variables. To address this limitation, we propose to reduce the dimensionality by extracting the relevant features from each 'omic' data. For instance, genomic data can provide additional information about gene-sets, point mutations, copy number variants, inversions and mosaic events (http://www.creal.cat/jrgonzalez/software.htm). Alternatively, RNA-seq data can inform about global transcriptomic expression and alternative splicing events. Once data are summarized, we use a novel approach to perform data integration called regularized generalized canonical correlation. This method incorporates two different ways of combining datasets. Fist, it provides a consensus between tables (e.g. similar to PCA for more than two tables). Second, it includes information about how datasets are linked (i.e. data at different time-points, causality relationships,...). Outputs are then used to: 1) describe profiles of individuals having similar phenotypes to define "disease signatures", and 2) determine those individuals with similar patterns which may help to stratify the population or describe new phenotypes. Initial results include analysis on data from the TCGA (The Cancer Genome Atlas) project including genomic, epigenomic, transcriptomic and clinical information, to study the determinants of cancer prognosis. Our approach finds biomarkers that correlate with complex combinations of different features obtained from 'omic' data that improves individuals' prognosis prediction.

Poster - 16

Laura Isús Díaz

Laura Isús (presenting author)[1], Jofre Tenorio-Laranga[1], Mireia Herrando-Grabulosa[2], Caty Casas[2] and Patrick Aloy[1,3]

[1] Institute for Research in Biomedicine (IRB Barcelona), Joint IRB-BSC Program in Computational Biology, Spain. [2] Group of Neuroplasticity and Regeneration Department of Cell Biology, Physiology and Immunology Institute of Neurosciences, Universitat Autónoma de Barcelona, and Centro de Investigacion Biomédica en Red sobre Enfermedades Neurodegenerativas (CIB)

A Network approach to Spinal Cord Injury

Abstract

Spinal Cord Injury (SCI) represents a severe health problem associated with lifetime disabilities. Immediate cell death occurring after SCI is followed by a progressive death of neurons and degeneration. Depending on the severity and the proximity to the soma of the axonal injury, spinal motoneurons (MNs) may evolve to a retrograde degeneration reaction or to a regenerative process. Combining proteomic data with physical and functional interaction information specifically containing disease-associated genes and their direct interactors, can provide further insights into the dynamic behavior and mechanisms involved in the degeneration and regeneration of spinal motoneurons. We propose a directed integrative approach to decipher the distinct molecular and cellular changes that contribute to each type of process and particularly in the death mechanisms and the characteristic neuropatic pain associated with the degenerative process. Comparing GSEA analysis of our lists of candidates and networks with classical functional enrichment analysis (DAVID) we have been able to identify distinct enriched pathways (motives) between the degenerative and the regenerative process instead of general GO terms and KEGG pathways common to both models and many other disorders. In conclusion, some motives become significant only when direct interactors were included in the GSEA (e.g., Anoikis and Autophagosome fusion events) showing that by mapping our candidates to an interaction network we are increasing the statictical power of our analysis. Our directed approach helps us to rationalize our findings and may provide new candidate and interesting proteins for further analyses.

Poster - 17

Samira Jaeger

Samira Jaeger[1], Rodrigo Arroyo[1], Patrick Aloy[1,2]

[1] Institute for Research in Biomedicine, Spain. [2] Institució Catalana de Recerca i Estudis Avançats (ICREA), Spain

A network biology approach to identify breast cancer drug targets and drug combinations

Abstract

Breast cancer is, after lung cancer, the second most common cancer type in the world, and the leading cause of cancer death in women (Hutchinson 2010). Clinically, it is a very heterogeneous disease regarding the underlying molecular alterations, the cellular composition of tumors, and the different clinical outcomes (Polyak 2011). Ongoing efforts in (breast) cancer drug discovery yielded a shift towards mechanism-based and target-oriented strategies, particularly aiming at modulating specific molecular pathways and the tumor microenvironment (Gibbs 2000, Mills 2012) which improved the management of the disease significantly. Yet, despite the therapeutic progress and the expanding repertoire of new anti-cancer agents, therapy failure due to primary or acquired drug resistance remains a major challenge in the treatment of breast cancer (Tsang and Finn 2012). In this work, we present a network biology approach to identify novel therapeutic targets, i.e., proteins contributing to mechanisms related to breast cancer development and progression, and drug combinations for breast cancer. In line with novel network pharmacological approaches, we aim for finding multi-target intervention strategies, i.e., new candidates and combinations that target particular groups of proteins involved in perturbed complexes or pathways (Hopkins 2008). We identify novel breast cancer drug targets from a breast cancer specific network by exploiting primary targets of successful drugs as their annotated functions and pathways reflect mechanisms and processes which need to be modulated pharmacologically. We show that the resulting candidates provide promising individual and combinatorial drug targets for simultaneous inhibition with breast cancer agents. Furthermore, we developed a measure for inferring drug combinations with respect to prevent drug resistance by taking functional redundancy and pathway crosstalk into account. Our strategy is based on a crosstalk inhibition measure which is able to quantify the amount of pathway crosstalk that can be prevented by inhibiting two (sets of) targets in combination. We apply this measure to the set of candidate as well as primary breast cancer targets to infer novel drug combinations. A selected number of inferred drug combinations is curren.

Poster - 18

Rory Johnson

Centre for Genomic Regulation, Spain

Genomic recycling: A new hypothesis for evolution of long noncoding RNA functions

Abstract

The human genome contains thousands of long noncoding RNAs (lncRNAs) of unknown function. Our present challenge is how to infer lncRNA function given sequence alone, which in turn requires an understanding the nature of lncRNA functional domains. In fact, the literature contains several examples of lncRNA domains that originate from transposable elements (TEs). We propose that this reflects a general phenomenon of lncRNA functional evolution through insertion of functional fragments of TEs. We will explore how this model can explain the observed rapid evolution and functional diversity of lncRNA. We present a genome wide catalogue of putative TE-derived lncRNA domains. These sequences bear several hallmarks of functionality. In summary, this hypothesis of TE exaptation in driving lncRNA functional evolution holds promise in functional prediction for lncRNA.

Poster - 19

David Juan

David Juan[1], Daniel Rico[2], Tomas Marques-Bonet[1], Oscar Fernandez-Capetillo[3] and Alfonso Valencia[1]

[1] Structural Biology and Biocomputing Programme, Spanish National Cancer Research Centre (CNIO), Madrid, Spain. [2] Institut Catala de Recerca i Estudis Avancats (ICREA) and Institut de Biologia Evolutiva (UPF/CSIC), Barcelona, Spain. [3] Genomic Instability Group; Spanish National Cancer Research Centre (CNIO); Madrid, Spain

Late replicating CNVs as a source of new genes

Abstract

Asynchronous replication of the genome has been associated with different mutation rates and copy number variation (CNV) in human populations. The late replication is generally associated to heterochromatic regions that tend to suffer high replicative stress. CNVs typically involve intermediate to large regions, providing a potential substrate for the generation of new genes through gene duplication. Here, we elucidate the possible relevance of the association of CNV regions with later DNA replication times on gene birth and evolution. Our analyses show that most human genes duplicated in the Primate lineage are located in late replicating CNV regions. We have traced the relationship between replication timing and the evolutionary age of duplicated genes. Strikingly, we have found that the more recently a gene has been duplicated the later it is replicated in human and mouse cells. These results suggests that the currently active accumulation of CNVs in late replicating regions has being also persistently and extensively influencing genome evolution throughout animal history. We propose a new evolutionary model based on these observations in which chromatin structure and replication dynamics conditions gene birth and the rising of new protein functions. This scenario would have shaped the available evolutionary outcomes while facilitating fast tissue specialization in metazoans species. Our model establishes a realistic framework that allows understanding apparently controversial observations done by evolutionary and molecular biologists. In particular, it has the potential to help us to interpret the functional and structural consequences of the massive somatic DNA damage observed in many cancer cells.

Poster - 20

Petra Kaferle

Petra Kaferle[1], Manolis Papamichos-Chronakis[1]

[1] Nuclear Dynamics and Genome Plasticity, Institut Curie/Section de Recherche, Paris, France

SGA: A versatile high throughput platform for yeast functional genomics in France

Abstract

With complete sequence of yeast genome in 1996 [1] and construction of single-deletion collection in year 2000 [2], era of whole-genome experimental approaches started to evolve with high speed. One of the earliest techniques that emerged was Synthetic Genetic Array (SGA) [3]. Crossing of query strain, which carries a mutation in gene of interest, with desired collection, sporulation and selections for specific markers, enable us to introduce fast new perturbation in vast number of strains. Newly introduced mutation represents a hub in genetic network, whose interactors we would like to annotate. In addition, external perturbations, e.g. drugs or semi-permisive temperature, can also give additional information about system response. With usage of many different libraries we can study effects of depletion, over-expression and localization of gene products in cells under specific conditions. Today SGA represents a very powerful tool in the field of yeast functional studies. We have created a fully equipped laboratory, dedicated to SGA technique, which is available to all interested collaborators. The core platform is based on the 48-plate robotic manipulator for replicating colonies from/to solid and liquid media in 96, 384 and 1536 format. In addition, the robot is equipped with a re-arraying tool that enables custom-designed organization of collections or isolation of individual strains from plates. Image acquisition and analysis are also fully automated [4]. Furthermore, we provide different collections of S.cerevisiae. Besides technical equipment we also offer full customer support before starting an experiment, all necessary troubleshooting and help with data analysis, which is done in collaboration with the Bioinformatics and Computational Systems Biology of Cancer department in Institut Curie. References: 1. Science. 1996 Oct 25;274(5287):546, 563-7. 2. Comp Funct Genomics. 2001;2(4):236-42. 3. Science. 2001 Dec 14;294(5550):2364-8. 4. http://sgatools.ccbr.utoronto.ca/

Poster - 21

Pascal Kahlem

Patricia Carvajal (presenting author), Jonathan Mackinnon and Pascal Kahlem

Scientific Network Management SL, Calle Ateneu 17; 08960 Sant Just Desvern; Spain Tel: 0034 698707986 - info@s-n-m.eu

Scientific Network Management, SL An SME to support project drafting and management of scientific networks

Abstract

Scientific Network Management S.L. (SNM) is a privately owned company, registered in Spain since January 2013, which focuses on the drafting and management of scientific research projects and networks. The company offers two major activities: I. The first activity is the drafting of scientific projects and scientific manuscript editing. The company has previous experience in project drafting for three Framework Programme 7 calls of the European Commission. In the three cases, the company had prepared both stage1 and stage2 proposals. Currently, the company is supporting a consortium preparing a grant proposal for a Horizon2020 2-stages call, which, if funded, would start by the end 2014. Currently SNM has the capacity to undertake up to four grant proposals simultaneously. II. The second activity is the management of scientific networks in the field of life science. The success of scientific international networks relies on coordination skills, which are often exceeding the workload of the scientific coordinators involved, hence the need for the specialized function of project manager within such consortia. This includes: 1) Preparing, executing and post-processing of major project meetings. 2) Implementing and maintaining the project infrastructure. 3) Optimizing internal and external communication. 4) Organizing of conferences and training. 5) Managing the reporting. 6) Writing and maintaining of the consortium agreement. Capitalizing on over 8 years of experience of its managing director and coworkers as managers in more than 7 scientific networks funded either by the Framework Programmes 6 and 7 of the European Commission or by other funding bodies, SNM covers most requirements for the management of international consortia. SNM is validated as an SME by the European Commission under the PIC number 951025831.

Poster - 22

Ivan Kulakovskiy

Irina A. Eliseeva[1], Ilya E. Vorontsov[2], Kirill E. Babeyev[3], Sofya M. Buyanova[4], Maria A. Sysoeva[5], Fyodor A. Kondrashov[6,7,8], Ivan V. Kulakovskiy[2,9] (presenting author)

[1] Institute of Protein Research RAS, Russia. [2] Vavilov Institute of General Genetics RAS, Russia. [3] Gymnasium MBGO, Russia. [4] Lomonosov Moscow State University, Russia. [5] National Mining Resources University, Russia. [6] Centre for Genomic Regulation, Spain. [7] Universitat Pompeu Fabra, Spain. [8] Institució Catalana de Recerca i Estudis Avançats, Spain. [9] Engelhardt Institute of Molecular Biology, Russia

mTOR mRNA targets and transcription-translation rendezvous: a peep through sequence analysis keyhole

Abstract

Many mRNAs responding to growth-dependent translation carry a well-known sequence motif, 5'-terminal oligopyrimidine tract, TOP, in their 5' UTRs. In particular, translation of such mRNAs is regulated by mTOR signaling pathway. Among many functions, the mTOR pathway plays important role in cancer development, cell proliferation and aging. Modern high throughput data allows studying detailed sequence features of mTOR mRNA targets and respective transcription start sites (TSS) at DNA level. Massive identification of mRNAs regulated by the mTOR pathway was done in 2012 by Hsieh et al. in PC3 cells using ribosome profiling. The study reported a novel pyrimidine-rich translational element (PRTE), as a key regulatory motif, that, unlike 5' TOP, did not exhibit any preferences to 5' ends of 5'-UTRs. To clarify the picture, we coupled ribosome profiling data of Hsieh et al. with the TSS data obtained in 2011 by Kanamori-Katayama et al. with HeliScopeCAGE, cap-analysis of gene expression on a single-molecule sequencer. By bioinformatics sequence analysis we confirm the canonical TOP motif and its strong positional preferences to 5' ends of 5'-UTRs. At the DNA level, respective oligopyrimidine tracts (OP) straddle the experimentally validated TSS regions. Therefore, transcription from OP segments creates the 5'-terminal TOP in the corresponding mRNAs. Importantly, downstream regions of 5'-UTRs of mTOR targets are not enriched with oligopyrimidine motifs, calling existence of PRTE into question. Finally, we demonstrate several mTOR target genes with broad and multimodal TSS spanning dozens of nucleotides and carrying a distinct subtype of OP/TOP motif. Some of those TSS are only partially covered with OP occurrences. Thus, transcription from such TSS produces TOP motif only for a fraction of transcribed mRNA and only this fraction may respond to mTOR via TOP mechanism. We hypothesize that the interplay between transcription and translation may play a crucial role in the regulation of the mTOR response.

Poster - 23

Pierre Luisi

Pierre Luisi[1], David Alvarez-Ponce[2,3], Marc Pybus[1], Mario A Fares[2,3], Jaume Bertranpetit[1] and Hafid Laayouni[1]

[1] Institute of Evolutionary Biology (Universitat Pompeu Fabra-CSIC), CEXS-UPF-PRBB, Barcelona, Catalonia, Spain. [2] Integrative Systems Biology Group, Instituto de Biología Molecular y Celular de Plantas, Consejo Superior de Investigaciones Científicas (CSIC)-Universidad Politécnica de Valencia (UPV), Valencia, Spain. [3] Smurfit Institute of Genetics, University of Dublin, Trinity College, Dublin, Ireland

Recent positive selection targets the center of the human protein-protein interaction network

Abstract

Genes vary in their likelihood to undergo adaptive evolution. The genomic factors that determine adaptability, however, remain poorly understood. Genes function in the context of molecular networks, with some occupying more important positions than others and thus being likely to be under stronger selective pressures. However, how positive selection distributes across the different parts of molecular networks is still not fully understood. Here, we inferred long- and short-term positive selection events through the comparison of 10 mammalian and 270 human genomes, respectively, and observed that positive selection affected different parts of the human protein-protein interaction network (PIN) at different evolutionary time-scales. In agreement with previous results, genes that underwent long-term adaptation tend to act at the periphery of the network (i.e., they exhibit fewer interactions). Signatures of recent positive selection, however, concentrate at the center of the network. Therefore, network adaptation occurs through intra-specific adaptive leaps affecting key network genes, followed by fine-tuning adaptations in less important network regions.

Poster - 24

Ettore Luzi

Laboratory of Neuroendocrine Complex Diseases Center on Endocrine Hereditary Tumors, AOUC, Department of Surgery and Translational Medicine , University of Firenze

Role of microRNAs-Transcription Factors Gene Regulatory Networks dynamics in the onset and progression of tumors: the MEN1 syndrome as model

Abstract

Biological functions, and thus dysfunction in cancer, are based on networks, nonlinearities and dynamic processes. A basic notion of modern system biology is that biological functions are performed by groups of genes that act in an synergic and interdependent way (network). These complex networks (gene regulatory networks) can be divided into simpler regulatory patterns called network motifs, composed by three or four interacting components that are able to perform elementary signal processing functions. Network motifs can be divided into two categories: feedback and feed-forward loops. We have now evidence that microRNA-transcription factors are recurrent network motifs that enhance the robustness of gene regulation in mammalian genome by buffering the impact of noise on gene expression. Multiple endocrine neoplasia type 1 (MEN1) syndrome is a rare hereditary cancer disorder characterized by tumors of the parathyroids, of the neuroendocrine cells, of the gastro-entero-pancreatic tract, of the anterior pituitary, and by non-endocrine neoplasms and lesions. MEN1 gene, a tumor suppressor gene, encodes menin protein. Loss of heterozygosity at 11q13 is typical of MEN1 tumors, in agreement with the Knudson's two-hit hypothesis. We investigated this possibility by analysis of miR-24-1 expression profiles in parathyroid adenomatous tissues from MEN1 gene mutation carriers, in their sporadic non-MEN1 counterparts, and in normal parathyroid tissue. Interestingly, the MEN1 tumorigenesis seems to be under the control of a "negative feedback loop" between miR-24-1 and menin protein, that mimics the second hit of Knudson's hypothesis and that could buffer the effect of the stochastic factors that contribute to the onset and progression of this disease. Our data show an alternative way to MEN1 tumorigenesis and, probably, to the "two-hit dogma". The functional significance of this regulatory mechanism in MEN1 tumorigenesis is also the basis for opening future developments of RNA antagomir(s)-based strategies in the in vivo control of tumorigenesis in MEN1 carriers.

Poster - 25

Vsevolod Makeev

Ilya E. Vorontsov[1] Ivan V. Kulakovskiy[1,2] Vsevolod J. Makeev[1,2,3,4] (presenting author)

[1] Vavilov Institute of General Genetics, Russian Academy of Sciences, Moscow, Russia. [2] Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, Russia. [3] State Research Institute of Genetics and Selection of Industrial Microorganisms GosNIIgenetika, Moscow, Russia. [4] Moscow Institute of Physics and Technology, Dolgoprudny, Moscow Region, Russia

Large non-redundant collection of binding motifs for human transcription factors

Abstract

The results of experimental studies of DNA-protein binding are currently accumulated in several databases, with are usually biased in a sense that each collection contains binding motifs obtained by some preferred experimental method. For a popular TF, the aggregated collection of binding motifs usually contains several binding motifs. Additionally, some TFs may share the same DNA binding domain, thus recognizing very similar motifs at DNA. Our objective was to obtain a non-redundant collection of human TFBS recognition motifs, containing as many TFs as possible but with a controlled minimal similarity between each pair of motifs. Such collection of motifs is very appealing from practical point of view, and theoretically allows revealing overall landscape of TF binding motifs. We started from positional weight matrices (PWMs) from 5 most representative motif collections (HOCOMOCO, HOMER, JASPAR, SWISSREGULON, and the result of recent high-throughput SELEX experiments) initially containing a total of 1802 motifs for 681 human TFs. We used Jaccard based distances to obtain the distance matrix between PWM. The PWMs were then clustered by UPGMA approach. The resulting collection contains 856 distinct representative motifs of 856 clusters for about 530 factors, thus containing motifs for more than 30% of all known TFs.

Poster - 26

Izaskun Mallona

Izaskun Mallona (presenting author), Anna Díez, Mireia Jordà, Miguel A. Peinado

Institute of Predictive and Personalized Medicine of Cancer (IMPPC), Spain

Towards understanding the role of repeat elements in genome biology: a semantic approach

Abstract

About half of the human genome is derived from repetitive elements, whereas coding sequences represent less than the 2%. Alu elements are a primates-specific lineage of abundant retrotransposons belonging to the short interspersed elements (SINE) family. It has been reported their implication in multiple phenomena such as nucleosome positioning, alternative splicing or the generation of novel transcription factors binding sites. Nonetheless, the role of Alus and other repeats in shaping the genome remains unclear. The usage of ontologies is a relatively new field in bioinformatics, although some controlled vocabularies are widely used. For instance, Sequence Ontology (SO) offers a hierarchy of concepts and relationships to be used to annotate genomic data; and Gene Ontology (GO) provide a set of terms to describe molecular functions, biological processes and cellular locations of genes and gene products. However, formal reasoning over ontologies are not so widespread. In this work we describe a knowledgebase for the human Alus. The ontology is linked to SO and GO and is devoted to address the genomic context of the Aluome. For each Alu element, the closest gene and transcript are stored, as well their functional annotation according to GO, the state of the surrounding chromatin and the transcription factors binding sites inside the Alu. Semantic rules have been used to aid efficient data storage, querying and integration. The Alu ontology has been subjected to reasoners like Pellet and it is being transferred to a Semantic MediaWiki. Its formalization is freely available at http://gattaca.imppc.org/groups/maplab/imallona/ontologies/merged.owl.

Poster - 27

Yulia Medvedeva

Yulia A. Medvedeva[1] (presenting author), Abdullah Khamis, Ivan V. Kulakovskiy, Wail Ba-Alawi, Md. Shariful I. Bhuyan, Hideya Kawaji, Timo Lassmann, Alister R.R. Forrest, Matthias Herbers, FANTOM consortium, Vladimir B. Bajic[2]

[1] Institute of Predictive and Personalized Medicine of Cancer (IMPPC), Spain
[2] King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia

Effects of cytosine methylation on transcription factor binding sites

Abstract

DNA methylation of promoters is linked to gene repression [1]. Yet, the question remains whether DNA methylation is a cause or consequence of gene repression. In the former case, DNA methylation may affect the affinity of transcription factors (TFs) towards their binding sites (TFBSs). In the latter case, gene repression caused by chromatin modification, is stabilized by DNA methylation. Both above-mentioned scenarios have been supported by non-systematic evidences and have not been tested for a wide spectrum of TFs. In this study we found that for 16.6% of cytosines methylation profile and the expression profile of neighboring transcriptional start site show significant negative correlation. We name CpG that correspond to such cytosines *traffic lights*. CpG *traffic lights* are mostly located within CpG islands in gene promoters. We hypothesize that if CpG *traffic lights* are not induced by average methylation of a silent promoter, they may affect binding of TFs to their binding sites and therefore regulate transcription. We observed a strong selection against CpG *traffic lights* within TFBSs, more pronounced for *core* position of the TFBS, supporting the damaging role of CpG *traffic lights* for a TFBS. Surprisingly, we found selection to be stronger for repressors then for activators or multifunctional TFs. We conclude that single cytosine methylation may play a role in transcriptional regulation. At the same time, blocking of TFBS by selective methylation is likely to be restricted to special cases and cannot be considered as a general regulatory mechanism of methylation-dependant transcription. This puts the current common perception of the link of methylation and gene expression into a different perspective. This work is part of the FANTOM5 project, providing genome-wide expression data (Forrest et al., submitted) across various cell types using cap analysis of gene expression (CAGE) [2]. 1. Bird, A.P., DNA methylation versus gene expression. J Embryol Exp Morphol, 1984. 83 Suppl: p. 31-40. 2. Salimullah, M., et al., NanoCAGE: a high-resolution technique to discover and interrogate cell transcriptomes. Cold Spring Harb Protoc, 2011. 2011(1): p. pdb prot5559.

Poster - 28

Arcadi Navarro i Cuartiellas

(Jordi Rambla[1,2]); Angel Carreño[1,2]; Oscar Martínez Llobet[1,2]; Marc Sitges[1,2]; Arcadi Navarro[1,2]

[1] Centre for Genomic Regulation (CRG), Dr. Aiguader 88, 08003 Barcelona, Spain.
[2] Universitat Pompeu Fabra (UPF), 08003 Barcelona, Spain.

The EGA-CRG Elixir project

Abstract

We present our efforts to make the European Genome-phenome Archive (EGA) , currently a service of the European Bioinformatics Institute (EBI), a jointly managed service between the EBI and the Centre de Regulació Genómica (CRG). The EGA is a permanent archive for all types of potentially identifiable genetic and phenotypic data that has been consented for use in biomedical research, but not for open public distribution. Accepted submissions include manufacturer raw data from genome sequence, transcriptome, epigenome or proteomics experiments. The EGA also stores called variants, genotypes, study summary statistics and associated sample phenotypes. The EGA follows strict protocols for information management, data storage, security and dissemination. Authorized access to the data is managed in partnership with the data providing organizations. The EGA includes major reference data collections for rare and common diseases as well as control sets that can be used in addition to the public reference panels such as the 1000 Genomes project.

Poster - 29

Chiara Pallara

Chiara Pallara (presenting author)[1], Fabian Glaser[2], Juan Fernandez-Recio[3]

[1] Joint BSC-IRB research program in Computational Biology, Life Sciences Department, Barcelona Supercomputing Center, Spain.
[2] Bioinformatics Knowledge Unit, The Lorry I. Lokey Interdisciplinary Center for Life Sciences and Engineering Technion, Israel

Understanding CFC (cardio-facio-cutaneous) syndrome by molecular dynamics simulation

Abstract

MEK (mitogen activated kinase) and its only know substrate ERK (extracellular signal-regulated kinase) have been reported to have an essential role in fundamental cellular activities such as cell survival, proliferation, motility and differentiation. Moreover, recent studies revealed as many as 17 germline mutations (on 7 different positions) in MEK1 associated with the CFC syndrome, which result in an elevation of MEK1 basal kinase activity [1]. Here we show a preliminary study focused on the molecular basis of CFC syndrome-causing mutations in MEK1 and their effect on the structure and interactions of the protein. In order to investigate the intrinsic propensity for MEK1 active-inactive transition in CFC, we performed a total of four 100ns long NPT-MD simulations, using AMBER12 package [2] on the WT and three mutants linked to CFC (Y130C, E203K and Q56P). The analysis of the simulations suggests that these mutations tend to destabilize the overall MEK1 structure. More interestingly, the mutants simulations show that some specific regions, such as the P-loop (involved in the placing of the γ-phosphate of the ATP for catalysis) and the T-loop (responsible for the MEK activation), drift away from the initial closed (inactive) conformation and towards the kinase open (active) form. These remarkable results encourage us to direct our studies to the high-throughput regime now available thanks to the high-performance computing systems, in order to perform longer simulations and investigate all the CFC syndrome-related mutations. We hope that this work will allow a better understanding of the basic dysfunction caused by CFC-MEK1 mutations and lead to the development of effective treatments, still unavailable. [1] Rodriguez-Viciana, P., and Rauen, K.A. (2008). Biochemical Characterization of Novel Germline BRAF and MEK Mutations in Cardio-Facio-Cutaneous Syndrome. Methods Enzymol. 2008;438:277-89. [2] D.A. Case, T.A. Darden, T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A.W. Goetz, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, R.M. Wolf, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui.

Poster - 30

Vera Pancaldi

Vera Pancaldi (presenting author), Simone Ecker, Daniel Rico, Alfonso Valencia

Structural and Computational Biology Group, Spanish National Cancer Research Centre (CNIO), Madrid, Spain

Evolutionary importance of expression variability and its application to stratification of CLL cases

Abstract

The Spanish Chronic Lymphocytic Leukemia - ICGC project has generated extensive genomics and epigenomics datasets for this disease. One of the major findings, which is in line with the current literature, is that there is a non-trivial correspondence between the transcriptome and methylome of the different patients (Kulis et al. 2012). The differences in DNA methylation patterns were found to be directly related to the two clinical subtypes of the disease. These two subtypes are defined based on the level of mutation in the IgVH region, a higher mutation level (M-CLL subtype) correlates with better prognosis compared to the others (U-CLL subtype). On the contrary, the analysis of the transcriptome discovered an alternative subdivision of patients, which also correlates with disease progression prognosis but is not related to the mutated/unmutated subtypes (Ferreira et al. 2013). Analysing published gene expression data (Kulis et al., Ferreira et al., 2013;Fabris et al., 2008) we noticed that the U-CLL patients present a significantly wider distribution of expression levels of individual genes compared to M-CLL patients. This observation might reconcile the apparent discrepancy between epigenetics and gene expression results. From a biological standpoint, our findings suggest a relationship between the variability of gene expression across patients and variability that could be observed in multiple samples of the same patient or in multiple single cells from the same sample. To translate these ideas in practical terms we demonstrate that a classifier trained solely with information on the variability of the level of expression of each gene with respect to the average distribution of the expression in the population (distance of each gene expression value from the average expression values of each gene across all patients) is able to†distinguishes the two CLL subtypes. Fabris, S. et al., 2008. Molecular and transcriptional characterization of 17p loss in B-cell chronic lymphocytic leukemia. Genes, chromosomes & cancer, 47(9), pp.781-93. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18521849 [Accessed January 9, 2014]. Ferreira, P.G. et al., 2013. Transcriptome characterization by RNA sequencing identifies a major molecular and clinica.

Poster - 31

Sílvia Pérez-Lluch

Sílvia Pérez-Lluch[1,2], Enrique Blanco[3], Hagen Tilgner[1,2,4], Joao Curado[1,2,5], Montserrat Corominas[3] and Roderic Guigo[1,2]

[1] Centre for Genomic Regulation (CRG). Catalonia, Spain. [2] Universitat Pompeu Fabra (UPF). Catalonia, Spain. [3] Departament de GenËtica i Institut de Biomedicina (IBUB) de la Universitat de Barcelona. Catalonia, Spain. [4] Graduate Program in Areas of Basic and Applied Biology, Abel Salazar Biomedical Sciences Institute, University of Porto, Portugal. [5]

Transcriptional activation without chromatin marking in developmentally regulated genes

Abstract

The interplay of activating and repressing histone modifications at promoter regions is assumed to play a key role in the regulation of gene expression. Challenging this generally accepted model, here we show that activation of genes that are temporally and spatially regulated during metazoan development occurs in the absence of canonically activating histone modifications. Strong chromatin marking, in contrast, leads to transcriptional stability and tighter regulation of splicing. We also show that lack of chromatin marking is not a transient genomic state, but a constitutive property of developmentally regulated genes. Finally, we show that the promoter sequence of developmentally regulated genes is under greater selective pressure than that of constitutively expressed ones. Our results support a dual model of chromatin associated transcription regulation, in which chromatin marking is associated to stable, tightly controlled production of RNA, while a more flexible, unmarked chromatin state would permit rapid gene activation and de-activation during development. In these genes, Transcription Factors binding to chromatin would play the predominant regulatory role.

Poster - 32

Daniel Rico Rodríguez

The BLUEPRINT Consortium**. Presenting author: Daniel Rico[1]

[1] Structural Biology and BioComputing Programme, Spanish National Cancer Research Center, CNIO, Madrid, Spain

(*) Submitted for publication

(**) For more information, see www.blueprint-epigenome.eu

Epigenomic plasticity of human neutrophils and monocytes in cord and peripheral blood*

Abstract

Neutrophils and monocytes represent two of the most abundant nucleated myeloid cells in the blood and as part of the innate immune system they provide a first line of defense against infections. Here we report the detailed and integrated analysis of RNA-seq data, ChIP-seq of six histone modifications and of DNA methylation by bisulfite sequencing at base pair resolution for monocytes and neutrophils extracted from peripheral blood and cord blood. We identified specific epigenetic marks and gene-regulatory regions that are primarily determined by ontogenic cell states and others that primarily reflect the maturity of the cells. Comparing the two cell types revealed that monocytes show a substantially higher transcriptional activity than neutrophils, in line with functional analysis revealing that gene expression and translation pathways are significantly more active in monocytes. For the monocytes we characterized the sequence patterns of regulatory factors in the accessible sites using DNaseI-footprint analysis. We found that SNPs associated with Crohn's disease are significantly enriched in these monocyte-specific accessible sites. Genome segmentation identified genomic elements with different chromatin states in the two cell types. In comparison to neutrophils, monocytes have four times more cell-specific enhancer regions, probably reflecting a more plastic epigenome that can still undergo further differentiation into macrophages, dendritic cells and osteoclasts. In contrast, most of the neutrophil-specific features tend to be characterized by repressed chromatin. Finally, we observed that many pathways that showed differential expression, such as IFN signaling, also have important differences at the epigenomic level. Altogether, our study provides a comprehensive epigenetic chart of chromatin states in primary human neutrophils and monocytes, as well as new insights into the regulatory program of neutrophils and monocytes, thus providing a valuable resource for studying the regulation of the human innate immune system.

Poster - 33

Martin Schaefer

Martin H. Schaefer (presenting author)[1,2], Jae-Seong Yang[1,2], Luis Serrano[1,2,3], Christina Kiel[1,2]

[1] EMBL/CRG Systems Biology Research Unit, Centre for Genomic Regulation (CRG), Barcelona, Spain. [2] Universitat Pompeu Fabra (UPF), Barcelona, Spain. [3] Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

Protein conservation and variation suggest mechanisms of cell type-specific modulation of signaling pathways

Abstract

Many proteins and signaling pathways are present in most cell types and tissues and yet perform specialized functions. To elucidate mechanisms by which these ubiquitous pathways are modulated, we overlaid information about cross-cell line protein abundance and variability, and evolutionary conservation onto functional pathway components and topological layers in the pathway hierarchy. We found that the input (receptors) and the output (transcription factors) layers evolve more rapidly than proteins in the intermediary transmission layer. In contrast, protein expression variability decreases from the input to the output layer. We observed that the differences in protein variability between the input and transmission layer can be attributed to both the network position and the tendency of variable proteins to physically interact with constitutively expressed proteins. Differences in protein expression variability and conservation are also accompanied by the tendency of conserved and constitutively expressed proteins to acquire somatic mutations, while germline mutations tend to occur in cell type-specific proteins. Thus, conserved core proteins in the transmission layer could perform a fundamental role in most cell types and are therefore less tolerant to germline mutations. In summary, we propose that the core signal transmission machinery is largely modulated by a variable input layer through physical protein interactions. We hypothesize that the bow-tie organization of cellular signaling on the level of protein abundance variability contributes to the specificity of the signal response in different cell types.

Poster - 34

Christoph Schlaffner

Christoph Schlaffner[1,2] (presenting author), James Wright[1], Jyoti Choudhary[1]

[1] Wellcome Trust Sanger Institute, UK. [2] University of Applied Sciences Upper Austria, Austria

Similarity scoring for modification and coding variant discovery

Abstract

One of the key goals in biomedical research is to find genetic alterations that underline specific phenotypes. Genomics has proven successful in identifying somatic variants at a large scale. However, proteomics is also capable of detecting coding variation as well as revealing regulatory post translational modifications that are not accessible by genomics. Proteomics uses sequence database searching to match experimental spectra to theoretical spectra leading to assignment of peptides from which proteins and their modifications can be identified. This approach can only assign peptides represented in the database leaving many spectra in a standard experiment unidentified. Data mining has revealed that in part this is due to the occurrence of unanticipated modifications shifting the mass of precursor and fragment ions. An approach for finding these unexpected modifications is to pair spectra from the same dataset and then find the delta mass shift between the precursor ions in the two fragment ion spectra. Here we describe a method to find pairs of unmodified and modified spectra and introduce a new scoring system based on a general spectral dot-product. Furthermore, our method allows generation of networks of spectra with delta mass differences for discovery of coding variants and modifications. To demonstrate the effectiveness of this method, spectral pairing experiments were conducted on a published datasets of 2400 synthetic peptides (over 7000 spectra) and a well curated experiment (over 74000 spectra). We compared our results with that obtained from high standard database searches. Overall, the results showed that our method can effectively pair spectra that differ by a single modification or amino acid substitution.

Poster - 35

Maria Shutova

Maria Shutova[1], Anastasia Surdina[1], Vladimir Naumov[2], Dmitry Ischenko[2], Ekaterina Vassina[1], Maria Lagarkova[1,2], Sergey Kiselev[1]

[1] Vavilov Institute of General Genetics RAS, Moscow, Russia. [2] Scientific Research Institute of Physical-Chemical Medicine, Moscow, Russia

Induction of pluripotency leaves no reprogramming-specific trace in the human somatic cells-derived iPS cells

Abstract

Human induced pluripotent stem (hiPS) cells are a promising source of different types of patient-specific cells for future use in biomedicine. However, it remains controversial whether hiPS cells are molecularly indistinguishable from human embryonic stem (hES) cells, which serves as a standard for pluripotency. By comparing genetically identical human ES and iPS cells, we show that their expression (Illumina HT12v.4) and methylation (Illumina 450k) landscapes are quite similar. Different clones of hiPS cells showed a small number of iPS-specific genes and CpGs, but all of them has clone-specific or lab-specific batch effects and cannot be considered as a consequence of reprogramming process itself. Interestingly, hiPS cells derived from different parental lines (neurons, retinal pigmented epithelium, and fibroblast-like cells) share only 6-9% of the genes and CpGs which are at the same level of expression or methylation between these lines and isogenic ES cells line. That core set of pluripotency genes and CpG loci represent mostly the targets of Oct4, Sox2 and Nanog ES-specific transcription factors as well as of histone modifications H3K27me3 and Polycomb. Functionally they represent the genes involved mostly in negative regulation of epithelial cell lineage development. Genes related to mitosis and cell proliferation were found within more diverse set of pluripotency-related genes. Thus, the pluripotent state of newly developed hiPS cells was controlled mostly by transcriptional and DNA methylation lock of the developmental genes. On the other hand, the ability of hiPS cells to proliferate is controlled at the transcriptional level by recovery of the ES-like proliferative pattern. We also showed that so called “somatic memory” genes or CpGs in iPS cells don’t indicate a cell of origin and this type of differences decreases upon extended culture of iPS cell lines. Interestingly, only 13% of iPS specific genes preserve their state during cultivation. Together, these data suggest that despite the minor transcriptional and epigenetical fluctuations human iPS cells are indistinguishable from human ES cells and don’t retain any specific markers of reprogramming.

Poster - 36

Mónica Suelves Esteban

Elvira Carrió[1], Izaskun Mallona[1], Anna Díez[1], Ildefonso Cases[2], Miquel Ángel Peinado[1] and Mónica Suelves[1]

[1] Institute of Predictive and Personalized Medicine of Cancer, Barcelona, Spain.
[2] Genomics and Bioinformatics Platform of Andalusia, Seville, Spain

Dynamic changes in DNA methylation during muscle-lineage commitment and terminal differentiation

Abstract

Cellular differentiation involves widespread epigenetic reprogramming, including modulation of DNA methylation patterns. DNA methylation is a covalent biochemical modification restricted to CpG dinucleotides in mammals that controls chromatin structure and gene expression. We have investigated the role of DNA methylation during muscle-lineage commitment and muscle terminal differentiation. Mouse embryonic stem cells, muscle stem cells, terminal differentiated myotubes and adult muscle tissue at genome-wide scale have been analyzed by AIMS-Seq. We have identified near to 1.000 differentially methylated regions (DMRs) during muscle-lineage commitment and muscle terminal differentiation. Our results show a major gain of DNA methylation during muscle lineage commitment, and an interesting loss of DNA methylation in non-repetitive CpG-poor regions during differentiation. Hypermethylated amplicons in muscle cells are significantly enriched in regulatory regions (-2Kb to -50Kb) and intergenic regions, meanwhile the hypomethylated sequences are significantly enriched in 5-UTR, exons and promoter regions. Interestingly, hypomethylated regions show enrichment of the two main myogenic master regulators MyoD and Myogenin according to ENCODE data, supporting the participation of DNA methylation in controlling myogenesis. In addition, we have identified two important muscle-regulatory genes and two pluripotency related genes differentially methylated in myogenic cells versus non-muscle lineage cells. Ongoing studies will clarify the functionality of these differentially DNA methylation changes in the establishment and identity of the muscle lineage. This proyect has been supported by Ministerio de Ciencia e Innovación (SAF2009-08128 and SAF2012-37427) and from Generalitat de Catalunya (2009 SGR1356). EC is a FPI Fellow (MCINN). AD is supported in part by PTA2011-5655-I (MCINN).

Poster - 37

David Torrents

Valentí Moncunill[1], Santi Gonzalez[1], Lise Andrieux[1], Sílvia Beà[2], Itziar Salaverria[2], Cristina Royo[2], Laura Martinez[1], Montserrat Puiggrós[1,3], Maia Segura-Wang[4], Adrian M. St,tz[4], Alba Navarro[2], Romina Royo[1,3], Josep Ll. Gelpí[1,3,5], Ivo G. Gut[6], Carlos López-Otín[7], Modesto Orozco[1,5,8], Jan O. Korbel[4], Elias Campo[2], Xose S. Puente[7], David Torrents[1,9] (presenting author)

[1] Joint IRB-BSC Program in Computational Biology, Barcelona Supercomputing Center, 08034 Barcelona. [2] Department of Pathology, Hematopathology Unit, Hospital Clínic, August Pi I Sunyer Biomedical Research Institute (IDIBAPS), Hospital Clinic of Barcelona, University of Barcelona,08036 Barcelona, Spain. [3] Computational Bioinformatics, National Institute of Bioinformatics, Spain. [4] European Molecular Biology Laboratory, Genome Biology Research Unit, Meyerhofstr. 1, Heidelberg, Germany. [5] Departmen

Accurate characterization of complex structural variation in cancer by using a reference-free approach

Abstract

The development of highthroughput sequencing technologies has changed our understanding of cancer. However, despite the increasing demand to identify the genetic alterations in tumor cells, the characterization of somatic structural variants in cancer still remains a challenge. Current strategies depend on the alignment of reads to a reference genome, a step that restricts the complete definition of structural variation. In this work we developed a reference-independent approach called SMUFIN (Somatic MUtation FINder), which is able to accurately identify all types of somatic variation, from substitutions to large structural variants (SVs), at base pair resolution. Performance tests showed average sensitivity of 92% and 74% for SNVs and SVs, with specificities of 95% and 91%, respectively. Analysis of two aggressive forms of solid and hematological tumors revealed that this procedure identifies breakpoints associated with chromothripsis and chromoplexy with high specificity. Taken together, SMUFIN constitutes the first reference-free and integrated solution for an accurate and complete characterization of somatic variation in cancer.

Poster - 38

Leonid Uroshlev

(presenting author) Leonid Uroshlev, Vsevolod Makeev

[1] Institute of Molecular Biology RAS, Moscow, Russian Federation. [2] Institute of General Genetics RAS, Moscow, Russian Federation

Ion prediction in apo-form of proteins

Abstract

Motivation: Proteins with metal ion cofactors make up from a third to a half of all proteins. Computational prediction of ion binding site in protein form, obtained without ion (so-called Apo-form) is important for practical applications. For example, it might be utilized to verify crystallized structure of a given protein or to rationally design protein to be able able to bind a given type of ion. Results: With help of our algorithm, PIONCA, we predicted Zinc, Calcium and Magnesium binding for Apo and Holo-pairs with different RMS. Ion position in Holo form was predicted from the structure of Apo form. Using a library of Apo-Holo structure pairs we evaluated how prediction accuracy depended on the RMS distance between the Apo and the Holo forms and on the resolution of the Apo form structure. The prediction success rate was reaching from 50% up to 94% depending on the RMS distance between Apo and Holo stuctures.

Poster - 39

Ignacio Vazquez García

Ignacio Vázquez García[1,2] (presenting author), Andrej Fischer[2], Ville Mustonen[2]

[1] DAMTP, Centre for Mathematical Sciences, University of Cambridge, Cambridge, CB3 0WA, United Kingdom. [2] The Wellcome Trust Sanger Institute, Hinxton, Cambridge, CB10 1SA, United Kingdom

Probabilistic reconstruction of subclonal diversity in adaptive evolution

Abstract

Adaptation to a new environment commonly requires the acquisition of genetic alterations, which can arise as point mutations, copy number changes and rearrangements. Due to beneficial “driver” alterations accumulating in the progeny of a single cell, unique clones can emerge, often carrying other “passenger” events without distinct fitness effects. Such population structure is common in asexual evolution, ranging from bacterial, parasitic or viral infections to cancer progression. Examining the clonal hierarchy of these genomic alterations in heterogeneous populations sheds light into their adaptive trajectories. However, reconstructing and demixing the aggregate signal of a subclonal population is a challenging problem, as short reads from next-generation sequencing do not yield direct information on long-range haplotypes when applied to mixed cell populations. Here we introduce a probabilistic inference method that exploits whole genome sequencing to infer the number of subclones, their fractions and their genotype posterior probabilities from a collection of mutation and copy number alterations. The algorithm groups small genic and large chromosomal alterations into major subclones, estimates their number and enables a high-definition reconstruction of their genotypes. We show that this framework can be used to systematically track how populations respond to strong selection using time-series data collected from somatic cell populations of human cancer.

Poster - 40

Ilya Vorontsov

Ivan Kulakovskiy Ilya Vorontsov (presenting author) Grigory Khimulya Vsevolod Makeev

[1,4] Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Russia
[1,2,3,4] Vavilov Institute of General Genetics, Russian Academy of Sciences, Russia

Comparison and classification of transcription factor binding sites models with application to functional annotation of regulatory sequence variants

Abstract

Positional weight matrix (PWM) remains most popular for quantification of transcription factor (TF) binding. PWM supplied with a score threshold defines a set of putative transcription factor binding sites (TFBS), thus providing a TFBS model. TFBS identified by different experimental methods produce similar but not identical PWMs. Nowadays, even for human TFs, there are many different collections of PWMs often containing similar-but-not-the-same models for a given TF. Several models also exist for similar TFs from the same structural family. Such redundancy complicates downstream computational analysis, such as functional annotation of single nucleotide variants lying in gene regulatory regions. Thus, it is an important task to measure properly the similarity between PWM-defined TFBS models. Common approaches to PWM comparison are based on comparison of matrix elements and do not take respective score thresholds into account. However, difference between two TFBS sets recognized by a given pair of PWMs can heavily depend on the score thresholds. We propose a variant of the Jaccard index as a practical approach to compare two TFBS models, each consisting of a PWM and the respective scoring threshold. We show how to effectively compute the proposed measure using dynamic programming approach and present the efficient software implementation: MACRO-APE (MAtrix CompaRisOn by Approximate P-value Estimation). Using several up-to-date collections of known binding motifs for human transcription factors we show how MACRO-APE can be utilized to classify existing TFBS models of different collections and reduce redundancy of the joint PWM set. Finally, basing on this methodology, we present a sister tool, PERFECTOS-APE (PrEdicting Regulatory Functional Effect by Approximate P-value Estimation), aimed at functional annotation of disease-associated single-nucleotide variants, located in cis-regulatory regions of genes and possibly affecting transcriptional regulation via TF binding. As a case study we applied PERFECTOS-APE to analyze several polymorphisms, associated with an increased risk of breast cancer.

Poster - 41

Dieter Weichenhan

Qi Wang[1], Lei Gu[1,2], Andrew Adey[3], Bernhard Radlwimmer[4], Wei Wang[4], Volker Hovestadt[4], Marion B‰hr[2], Stephan Wolf[5], Jay Shendure[3], Roland Eils[1], Christoph Plass[2], Dieter Weichenhan[2] (presenting author)

[1] Theoretical Bioinformatics, German Cancer Research Center (DKFZ), Heidelberg, Germany. [2] Division of Epigenomics and Cancer Risk Factors, German Cancer Research Center (DKFZ), Heidelberg, Germany. [3] Department of Genome Sciences, University of Washington, Seattle, USA. [4] Division of Molecular Genetics, German Cancer Research Center (DKFZ), Heidelberg, Germany. [5] Core Facility High Throughput Sequencing, German Cancer Research Center (DKFZ), Heidelberg, Germany

Tagmentation-based whole genome bisulfite sequencing for very low input DNA amounts

Abstract

Aberrations in the DNA methylome contribute to onset and progression of cancer and other diseases. Whole genome bisulfite sequencing (WGBS) is the gold standard in methylome analysis. The large DNA amount of about 5 µg required for WGBS is prohibitive, however, in studies where only quantities in the ng range are available. Alternative methods with less input usually interrogate CpG-dense regions and, hence, only minor portions of the methylome. However, only the complete methylome offers the view on all functionally important genomic features including enhancers and noncoding RNAs in regions of low CpG density. We have established tagmentation-based WGBS (TWGBS) which makes use of a hyperactive transposase for DNA fragmentation and adapter attachment in a single step. TWGBS circumvents several DNA input and time consuming steps inherent to the conventional WGBS library preparation. T-WGBS requires not more than 10-30 ng DNA and, hence, is ideally suited for studying precious biological specimens such as sorted cells or micro dissected tissue samples. We compared conventional WGBS with TWGBS and present the high reliability of TWGBS with respect to genome-wide CpG coverage, correlation of methylation levels, sequence duplication frequency and consistency between independent replicates.

(GIGA)n SCIENCE

www.gigasciencejournal.com

Editor-in-Chief:
Laurie Goodman (USA)

Executive Editor:
Scott Edmunds (Hong Kong)

华大基因 BGI

BioMed Central
The Open Access Publisher

delegate contacts

María del Rosario Abellán Sánchez
Unidad de Genotipado y Diagnóstico Genético. Fundación Investigación Clínico de Valencia - INCLIVA
m.rosario.abellan@uv.es

Alberto Acedo Bécares
Instituto de Biología y Genética Molecular (UVa-CSIC)
acedo@ibgm.uva.es

Vera Ademà
Josep Carreras Leukaemia Research Institute
vadema@carrerasresearch.org

Helena Aguilar
IDIBELL
haguilar@iconcologia.net

Jesús Albèniz Domínguez
Lifesequencing
jesus.albeniz@lifesequencing.com

Angel Alonso
Servicio de Genética, Complejo Hospitalario de Navarra
aalonsos@cfnavarra.es

Sergio Alonso
IMPPC
salonso@imppc.org

Monica Alvarado
Kaiser Permanente Southern California
Monica.X.Alvarado@kp.org

Miguel Alvarez Tejado
Roche Applied Science
miguel.alvarez-tejado@roche.com

Montserrat Andreu
Hospital del Mar
mandreu@parcdesalutmar.cat

Balasubramanyam Annangi
Department of Genetics and Microbiology, Faculty of Bioscience, Universitat Autonoma de Barcelona
basucall@gmail.com

Antonis C Antoniou
University of Cambridge, U.K.
a.antoniou@srl.cam.ac.uk

Jessica Aranda Cebrian
IMPPC
jaranda@imppc.org

Maria del Mar Arias Alonso
Hospital Virgen del Camino, Pamplona
mm.arias.alonso@navarra.es

Maria Armengol
IGTP
mparmengol@igtp.cat

Daniel Azuara
ICO - IDIBELL
dazuara@iconcologia.net

Neus Baena Diaz
UDIAT, Hospital de Sabadell
nbaena@tauli.cat

Francesc Balaguer
Hospital Clinic
fprunes@clinic.ub.es

Milagros Balbín Felechosa
Hospital Universitario Central de Asturias
mbalbin@hca.es

Luisa Balestrino
S.C. Genetica Medica, ASL8 Cagliari PO, Binagh
luibale@hotmail.com

Judith Balmaña
Vall d'Hebron Hospital, Barcelona
jbalmana@vhebron.net

Elizabeth Bancroft
Royal Marsden NHS Foundation Trust & Institute of Cancer Research
elizabeth.bancroft@rmh.nhs.uk

Maria Joao Baptista
Josep Carreras Leukaemia Research Institute
mjbaptista@carrerasresearch.org

Victor Barbera
Hospital General Universitario de Elche
barbera_vicjua@gva.es

Tina Bedenhem
Oxford Molecular Genetics Laboratory
tina.bedenham@ouh.nhs.uk

Fernando Bellido Molias
ICO - IDIBELL
fbellido@iconcologia.net

Beatriz Bellosillo
Hospital del Mar
94161@hospitaldelmar.cat

Javier Benítez Ortiz
Spanish National Cancer Research Centre (CNIO)
jbenitez@cnio.es

Elena Beristain Mendizabal
Hospital Universitario Araba-Txagorritxu
eberistain001@ikasle.ehu.es

Javier Bertol Chorro
Vall Hebron research institute
javier.bertol.chorro@gmail.com

Josep Biayna Rodríguez
IMPPC
jbiayna@imppc.org

Inga Bjornevoll
St. Olavs University Hospital in Trondheim
inga.bjornevoll@stolav.no

Ignacio Blanco
ICO-IDIBELL
iblanco@iconcologia.net

Ana Blanco Pérez
Fundación Póblica Galega de Medicina Xenómica
ana.blanco@usc.es

Pilar Blay
Hospital Universitario Central de Asturias
pilarblayalbors@gmail.com

Sandra Bonache
Vall d'Hebron
sbonache@vhebron.net

Ester Borràs Flores
ICO-IDIBELL
eborras@iconcologia.net

Nina Bosch Pages
Vall d'Hebron
nbosch.vhebron@gmail.com

Joan Brunet
ICO
jbrunet@iconcologia.net

Raquel Buj Gómez
IMPPC
RBuj@imppc.org

John Burn
Institute of Genetic Medicine, University of Newcastle United Kingdom
john.burn@newcastle.ac.uk

Marcus Buschbeck
IMPPC
mbuschbeck@imppc.org

Trinidad Caldes
Hospital Clínico San Carlos
trinidad.caldes@salud.madrid.org

Maria Adelaide Caligo
University Hospital
adelaide.caligo@gmail.com

Olga Campos Mata
ICO
ocampos@iconcologia.net

Jordi Camps Polo
Fundació Clínic per a la Recerca Biomèdica
jcamps@clinic.ub.es

Gabriel Capellá
ICO-IDIBELL
gcapella@iconcologia.net

Sabela Carballal Ramil
Hospital Clinic (Barcelona)
CARBALLAL@clinic.ub.es

Estela Carrasco López
Hospital Valle Hebron
ecarrasco@vhebron.net

Ineabel Carrillo Ortiz
Multiplicom NV
ineabel.carrillo-ortiz@multiplicom.com

Patricia Casano
Merck
arancha.despedida@hotmail.es

Alberto Cascón
Spanish National Cancer Research Centre (CNIO)
acascon@cnio.es

Ildefonso Cases
IMPPC
icases@imppc.org

Elisabeth Castellanos
IMPPC
ecastellanos@imppc.org

Marta Castells Zaragoza
Parc de Salut Mar
mcastells@parcdesalutmar.cat

Ester Castellsagué Torrents
McGill University
ester.castellsague@mail.mcgill.ca

Adela Castillejo Castillo
Elche University Hospital
castillejo.ade@gmail.com

Elena Castro Marcos
The Institute of Cancer Research, London, UK
elena.castro@icr.ac.uk

Emili Cid Roldós
IMPPC
ecid@imppc.org

Kathleen Claes
Center for Medical Genetics, Ghent University Hospital
kathleen.claes@ugent.be

Llorenç Coll Mulet
IMPPC
lcoll@imppc.org

Iñaki Comino Mendez
Spanish National Cancer Research Centre (CNIO)
icomino@cnio.es

Osmel Companioni Nàpoles
ICO-IDIBELL
ocompanioni@idibell.cat

Anna Cordeiro Santanach
UB, Universitat de Barcelona
acordesa11@alumnes.ub.edu

David Cordero Romera
ICO
dcordero@iconcologia.net

Martina C. Cornel
VU University Medical Center, The Netherlands
MC.Cornel@vumc.nl

Diana Coronado
coronado.dianam@gmail.com

Julien Swen Crettaz
Hospital Universitario Donostia
julienswen.crettaz@osakidetza.net

Miriam Cuatrecasas
Hospital Clinic
mcuatrec@clinic.ub.es

Raquel Cuesta Llorente
ICO
rcuesta@iconcologia.net

Alvaro Curiel García
Instituto de Biología y Genètica Molecular (CSIC-UVa)
acurielg@alumni.unav.es

delegate contacts

Esther Darder Bernabeu
ICO
edarder@iconcologia.net

Rafael de Cid
IMPPC
rdecid@imppc.org

Aguirre de Cubas
Spanish National Cancer Research Centre (CNIO)
aadcubas@cnio.es

Inmaculada de Juan Jiménez
Hospital Universitario La Fe
dejuan_inm@gva.es

Miguel de la Hoya
Instituto de Investigacion Sanitaria San Carlos (IdISSC)
mdhoya@hotmail.com

Cèsar del Rio
Merck Millipore
cesar.del.rio@merckgroup.com

Jesús del Valle
ICO
jdelvalle@iconcologia.net

Emmanouil Dermitzakis
University of Geneva Medical School
Emmanouil.Dermitzakis@unige.ch

Lohmann Dietmar
Universitäts klinikum Essen, Eye Cancer Genetics Research Group
dietmar.lohmann@uni-due.de

Vera Adema
Josep Carreras Leukaemia Research Institute (IJC)
vadema@carrerasresearch.org

Federico Agostini
Centre for Genomic Regulation (CRG)
federico.agostini@crg.es

Sergio Alonso
IMPPC
SAlonso@imppc.org

Maria Henar Alonso Aguado
IDIBELL
mhalonso@iconcologia.net

Teresa Alonso Carnicero
IMPPC
malonso@imppc.org

Patrick Aloy
Institute for Research in Biomedicine
patrick.aloy@irbbarcelona.org

Eduardo Andres Leon
IBIS
eleon-ibis@us.es

Xerxes Arsiwalla
Pompeu Fabra University (UPF)
x.d.arsiwalla@gmail.com

Irena Artamonova
Vavilov Institute of General Genetics RAS
irenart@gmail.com

Yassen Assenov
German Cancer Research Center (DKFZ)
y.assenov@dkfz.de

Jordi Banus Martí
IMPPC
jordi.banus18@gmail.com

Nuno Barbosa-Morais
Instituto de Medicina Molecular
nmorais@fm.ul.pt

Anaïs Baudot
CNRS AMU
anais.baudot@univ-amu.fr

Miguel Beato
Centre for Genomic Regulation (CRG)
migue.beato@crg.es

Niko Beerenwinkel
ETH Zurich
niko.beerenwinkel@bsse.ethz.ch

Carla Bello Cabrera
Centre for Genomic Regulation (CRG)
Carla.Bello@crg.eu

Sergi Beltran
Centro Nacional de Análisis Genómico (CNAG)
sbeltrana@pcb.ub.cat

Daniela Bezdan
Centre for Genomic Regulation (CRG)
bezdan.daniela@googlemail.com

Josep Biayna Rodríguez
IMPPC
jbiayna@imppc.org

Ewan Birney
EMBL-EBI
niko.beerenwinkel@bsse.ethz.ch

Enrique Blanco
Centre for Genomic Regulation (CRG)
enrique.blanco@crg.eu

Olga Bogatyrova
DKFZ (German Cancer Research Center)
o.bogatyrova@dkfz.de

Gireesh Bogu
Centre for Genomic Regulation (CRG)
girishbogu@gmail.com

Sarah Bonnin
Centre for Genomic Regulation (CRG)
sarah.bonnin@crg.eu

Alessandra Breschi
Centre for Genomic Regulation (CRG)
alessandra.breschi@crg.es

Søren Brunak
Technical University of Denmark
brunak@cbs.dtu.dk

Raquel Buj Gómez
IMPPC
rbuj@imppc.org

Marcus Buschbeck
IMPPC
mbuschbeck@imppc.org

Neus Cantariño Iglesias
IMPPC
ncantarino@imppc.org

Valentina Caputo
Imperial College London
v.caputo@imperial.ac.uk

Silvia Carbonell Sala
Centro Nacional de Análisis Genómico (CNAG)
scarbonell@pcb.ub.cat

Lucas Carey
Pompeu Fabra University (UPF)
lucas.carey@upf.edu

Angel Carreño Torres
Centre for Genomic Regulation (CRG)
angel.carreno@crg.eu

Anna Carreras
IMPPC
acarreras@imppc.org

Elvira Carrió Gaspar
IMPPC
ecarrio@imppc.org

Patricia Carvajal Vallejos
Scientific Network Management (S-N-M SL)
pcarvajal.sr2012@student.ie.edu

Eduard Casas Masnou
IMPPC
ecasas@imppc.org

Ildefonso Cases Cases
GBPA
icases@bioinfomgp.org

Elisabeth Castellanos
IMPPC
ecastellanos@imppc.org

Esra Nurten Cece
Pompeu Fabra University (UPF)
esra.cece@gmail.com

Maria Chatzou
Center for Genomic Regulation (CRG)
maria.chatzou@crg.es

Davide Cirillo
Centre for Genomic Regulation (CRG)
Davide.Cirillo@crg.es

Adrià Closa Mosquera
IDIBELL
aclosa@iconcologia.net

Laia Codo
Barcelona Supercomputing Center(BSC)
laia.codo@bsc.es

David Cordero
Catalan Institute of Oncology (ICO)
dcordero@iconcologia.net

Marc Corrales
Centre for Genomic Regulation (CRG)
marc.corrales@crg.eu

Alfred Cortes Closas
ICREA and CRESIB
alfred.cortes@cresib.cat

David Corujo García
IMPPC
dcorujog@gmail.com

Anton Crombach
Centre for Genomic Regulation (CRG)
anton.crombach@crg.eu

Joao Curado
Centre for Genomic Regulation (CRG)
joao.curado@crg.eu

Joaquin Custodio Rojo
IMPPC
jcustodio@imppc.org

Sarbashis Das
Institute of Molecular Biology
s.das@imb-mainz.de

Rafael de Cid
IMPPC
rdecid@imppc.org

Luciano di Croce
4DCellFate
Luciano.DiCroce@crg.eu

Anna Díez
IMPPC
adiez@imppc.org

Sarah Djebali
Centre for Genomic Regulation (CRG)
sarah.djebali@crg.es

Juliane Dohm
Centro Nacional de Análisis Genómico (CNAG)
jdohm@crg.es

Julien Douet
IMPPC
jdouet@imppc.org

Gabrijela Dumbovic
IMPPC
gdumbovic@imppc.org

Miquel Duran-Frigola
IRB Barcelona
miquel.duran@irbbarcelona.org

Johannes Engelken
IBE (Institute of Evolutionary Biology, CSIC-UPF)
johannes.engelken@upf.edu

Jose Espinosa-Carrasco
Centre for Genomic Regulation (CRG)
joseantonio.espinosa@crg.eu

delegate contacts

Harvey Evans
IMPPC
hevans@imppc.org

Eduardo Eyras
Pompeu Fabra University (UPF)
eduardo.eyras@upf.edu

Mario Ezquerra
IDIBAPS
marioezquerrat@hotmail.com

Fariha Faiyaz
MINARAT AR RIYADH SCHOOL
mfhaque@ksu.edu.sa

Anne Ferguson-Smith
University of Cambridge
afsmith@mole.bio.cam.ac.uk

Guerau Fernandez
IMPPC
gfernandez@imppc.org

Oscar Flores
IRB Barcelona - BSC
oscar.flores@irbbarcelona.org

Sonia Forcales
IMPPC
sforcales@imppc.org

Roald Forsberg
CLC bio
rforsberg@clcbio.com

Mirko Francesconi
Centre for Genomic Regulation (CRG)
mirko.francesconi@crg.es

Leonor Frias Moya
Centro Nacional de Análisis Genómico (CNAG)
leonor.frias@gmail.com

Marc Friedlander
Centre for Genomic Regulation (CRG)
marc.friedlaender@crg.eu

Raquel Fueyo Arévalo
IBMB-CSIC
rfabmc@ibmb.csic.es

Eileen Furlong
EMBL Heidelberg
eileen.furlong@embl.de

Laura I. Furlong
GRIB (IMIM-UPF)
lfurlong@imim.es

Francisco Fuster Tormo
IMPPC
ffuster@imppc.org

Carolina Gallo Lopez
Centre for Genomic Regulation (CRG)
carolina.gallo@crg.es

Alejandra Garcia
Molecular Biology Institute of Barcelona (IBMB-CSIC)
agabmc@ibmb.csic.es

Pablo M. Garcia-Roves
IDIBAPS
pgarciar@clinic.ub.es

Bernat Gel Moreno
Institut de Medicina Predictiva i Personalitzada del Cancer
bgel@imppc.org

Josep Ll. Gelpi
University of Barcelona (UB)
gelpi@ub.edu

Luca Giorgetti
Institute Curie, France
luca.giorgetti@curie.fr

Andre Gohr
MPI of Molecular Genetics
bioinf.ag@gmail.com

Antonio Gómez
IDIBELL
agomezm@idibell.cat

David Gomez-Cabrero
Karolinska Institutet
david.gomezcabrero@ki.se

Ramon Goñi
Barcelona Supercomputing Center(BSC)
ramon.goni@bsc.es

Santi Gonzalez
Barcelona Supercomputing Center(BSC)
santiago.gonzalez@bsc.es

Juan R González
Centre for Research in Environmental Epidemiology (CREAL)
jrgonzalez@creal.cat

Beatriz Gonzalez Alonso
IMPPC
BGonzalez@imppc.org

Juan Carlos González Sánchez
Centro Andaluz de Biología del Desarrollo (CABD)
juancarlosgonzs@gmail.com

Alba Gonzalez-Franquesa
IDIBAPS
agonzalf@clinic.ub.es

Ana Rita Grosso
Instituto de Medicina Molecular
agrosso@fm.ul.pt

Roderic Guigo
Centre for Genomic Regulation (CRG)
roderic.guigo@crg.eu

Alba Gutierrez
Research Programme on Biomedical Informatics (GRIB)
alba.gutierrez@upf.edu

Jessica Hernandez Rodriguez
Institut de Biologia Evolutiva - UPF
jessica.hernandez@upf.edu

Carles Hernandez-Ferrer
Centre for Research in Environmental Epidemiology (CREAL)
chernandez@creal.cat

Irene Hernando Herraez
Pompeu Fabra University (UPF)
irene.hernando@upf.edu

Heinz Himmelbauer
Centre for Genomic Regulation (CRG)
heinz.himmelbauer@crg.es

Susana Iraola Guzmán
IMPPC
siraola@imppc.org

Ibai Irastorza Azkarate
Centro Andaluz de Biología del Desarrollo (CABD)
iiraazc@upo.es

Manuel Irimia
University of Toronto
mirimia@gmail.com

Laura Isús Díaz
IRB Barcelona
laura.isus@irbbarcelona.org

Andrea Izquierdo Bouldstridge
IBMB-CSIC
aibbmc@ibmb.csic.es

Samira Jaeger
IRB Barcelona
samira.jaeger@irbbarcelona.org

Rory Johnson
Centre for Genomic Regulation (CRG)
rory.johnson@crg.eu

Mireia Jorda Ramos
IMPPC
mjorda@imppc.org

David Juan
Spanish National Cancer Research Centre - CNIO
dadejuan@cnio.es

Teresa Juan Blanco
IRB Barcelona
teresa.juanblanco@irbbarelona.org

Philippe Julien
Centre for Genomic Regulation (CRG)
philippe.julien@crg.es

Petra Kaferle
Institut Curie
petra.kaferle@curie.fr

Pascal Kahlem
Scientific Network Management, SL
pkahlem@s-n-m.eu

Ivan Kulakovskiy
Engelhardt Institute of Molecular Biology
ivan.kulakovskiy@gmail.com

Michael Lappe
CLC bio, a QIAGEN Company
mlappe@clcbio.com

François Le Dily
Centro Nacional de Análisis Genómico (CNAG)
françois.ledily@crg.es

Ben Lehner
4DCellFate
lehner.ben@gmail.com

Carmen Maria Livi
Centre for Genomic Regulation (CRG)
carmenmaria.livi@crg.eu

Verónica Lloréns-Rico
Centre for Genomic Regulation (CRG)
veronica.llorens@crg.eu

Sergio Lois Olmo
Vall d'Hebron Institut de Recerca (VHIR)
sergio.lois@gmail.com

Marcos López
Centre for Research in Environmental Epidemiology (CREAL)
marcos.lopezs@e-campus.uab.cat

Marcos López
Centre for Research in Environmental Epidemiology (CREAL)
mlopez3@creal.cat

Nuria Lopez-Bigas
Pompeu Fabra University (UPF)
nuria.lopez@upf.edu

Eva Loureiro
Catalan Institute of Oncology (ICO)
eloureiro@iconcologia.net

Pierre Luisi
Instituto de Biologia Evolutiva (UPF-CSIC)
pierre.luisi@upf.edu

Nick Luscombe
University College London and Cancer Research UK
nicholas.luscombe@cancer.org.uk

Ettore Luzi
University of Firenze
eluzi@unifi.it

Jonathan Mackinnon
Scientific Network Management
Jonathanagmackinnon@gmail.com

Vsevolod Makeev
Vavilov Institute of General Genetics, Russian Academy of Sciences
vsevolod.makeev@vigg.ru

Roberto Malinverni
IMPPC
rmalinverni@imppc.org

Izaskun Mallona
IMPPC
imallona@imppc.org

Oscar Marín Sala
VHIR
oscar.marin@vhir.org

delegate contacts

Francisco Martinez Jimenez
Centro Nacional de Análisis Genómico (CNAG)
fmartinez@pcb.ub.es

Kirsten McEwen
MRC Clinical Sciences Centre
kirsten.mcewen@csc.mrc.ac.uk

Yulia Medvedeva
IMPPC
ju.medvedeva@gmail.com

Maria Mendez Lago
Centro Nacional de Análisis Genómico (CNAG)
mmendezlago@pcb.ub.es

Alexandre Mendoza
Institut Biologia Evolutiva
alexmendozasoler@gmail.com

Angelika Merkel
Centro Nacional de Análisis Genómico (CNAG)
amerkel@pcb.ub.es

Belen Miñana
Centre for Genomic Regulation (CRG)
belen.minana@crg.es

Andre Minoche
Centre for Genomic Regulation (CRG)
andre.minoche@crg.es

Eva Morales
Centre for Research in Environmental Epidemiology (CREAL)
embarto@hotmail.com

Victor Moreno
Catalan Institute of Oncology (ICO)
v.moreno@iconcologia.net

Francisco D. Morón-Duran
ICO-IDIBELL
fdmoron@iconcologia.net

Adrià Mosquera
IDIBELL
aclosa@iconcologia.net

Josep Lluis Mosquera
University of Barcelona (UB)
jlmosquera@gmail.com

Sadaf Mughal
German Cancer Research Center (DKFZ)
s.mughal@dkfz.de

Loris Mularoni
Hospital Clinic de Barcelona
loris.mularoni@gmail.com

Elizabeth Murchison
Wellcome Trust Sanger Institute and Cambridge University
em4@sanger.ac.uk

Nikolina Nakic
Imperial College London
n.nakic@imperial.ac.uk

Arcadi Navarro i Cuartiellas
Centre for Genomic Regulation (CRG)
arcadi.navarro@upf.edu

Aaron New
Centre for Genomic Regulation (CRG)
aaron.new@crg.eu

Lara Nonell Mazelon
IMIM
lnonell@imim.es

Yaiza Núñez Álvarez
IMPPC
ynunez@imppc.org

Rocio Nuñez Torres
IBIS
rnuneztorres@gmail.com

Modesto Orozco
IRB, UB and BSC
modesto@mmb.pcb.ub.es

Stephan Ossowski
Centre for Genomic Regulation (CRG)
stephan.ossowski@crg.eu

Anna Palau
IMPPC
apalau@imppc.org

Chiara Pallara
Barcelona Supercomputing Center(BSC)
chiara.pallara@bsc.es

Laura Palomo Sanchis
Josep Carreras Leukaemia Research Institute (IJC)
lpalomo@carrerasresearch.org

Vera Pancaldi
Spanish National Cancer Research Centre (CNIO)
vpancaldi@cnio.es

Ida Paramonov
IMPPC
iparamonov@imppc.org

Solip Park
Centre for Genomic Regulation (CRG)
imagineyd@gmail.com

Lorenzo Pasquali
IDIBAPS
pasquali@clinic.ub.es

Solenn Patalano
Babraham Institute, Cambridge
solenn.patalano@babraham.ac.uk

Miguel A Peinado
IMPPC
mpeinado@imppc.org

Vera Pendino
European School of Molecular Medicine
vera.pendino@gmail.com

Alexandre Perera
Universitat Politècnica de Catalunya
Alexandre.perera@upc.edu

Gael Perez Alamancos
Pompeu Fabra University (UPF)
gael.perez@upf.edu

Sílvia Pérez-Lluch
Centre for Genomic Regulation (CRG)
silvia.perez@crg.eu

Inna Pertsovskaya
IDIBAPS - Hospital clinic
inna.perts@gmail.com

Manuel Perucho
IMPPC
conference2014@imppc.org

Janet Piñero Gonzalez
IMIM
jpinero@imim.es

David Pizarro Martinez
Centro Andaluz de Biología del Desarrollo (CABD)
dapizarro87@gmail.com

Merce Planas Félix
Barcelona Supercomputing Center(BSC)
merce.planas@bsc.es

Cristian Polo Rivera
IMPPC
cristian_noik@hotmail.com

Melanija Posavec
IMPPC
mposavec@imppc.org

Javier Prado-Martinez
IBE (UPF-CSIC)
javier.prado@upf.edu

Leszek Pryszcz
Centre for Genomic Regulation (CRG)
lpryszcz@crg.es

Emanuele Raineri
Centro Nacional de Análisis Genómico (CNAG)
emanuele.raineri@gmail.com

Veronica Raker
Centre for Genomic Regulation (CRG)
veronica.raker@crg.eu

Ravishankar Ramanathan
Centro Nacional de Biotechnologia (CNB), CSIC
rramanathan@cnb.csic.es

Jordi Rambla
Centre for Genomic Regulation (CRG)
jordi.rambla@crg.eu

Gabriel Rech
IMPPC
rechgab@gmail.com

Oscar Reina Garcia
IRB Barcelona
oscar.reina@irbbarcelona.org

Dmitry Repchevskiy
Barcelona Supercomputing Center(BSC)
dmitry.repchevski@bsc.es

Daniel Rico Rodríguez
Centro Nacional de Investigaciones Oncológicas - CNIO
drico@cnio.es

Casandra Riera Ribas
Vall d'Hebron Institut de Recerca (VHIR)
casandra.riera@vhir.org

Ana Rojas
IBIS, Spain
arojas-ibis@us.es

Rosa Romero
VHIR
rosaromero@gmail.com

Miguel Romero Durana
Barcelona Supercomputing Center(BSC)
mromero@bsc.es

Carlota Rubio Pérez
GRIB-UPF/IMIM
carlotarp@gmail.com

Mayka Sanchez
IMPPC
msanchez@imppc.org

Hurtado Sarah
IMPPC
hurtado-sarah@hotmail.com

Martin Schaefer
Centre for Genomic Regulation (CRG)
martin.schaefer@crg.eu

Gabriele Scheler
Correns Foundation
gscheler@gmail.com

Christoph Schlaffner
Wellcome Trust Sanger Institute
cs25@sanger.ac.uk

Michael Schroeder
GRIB, UPF-IMIM
michael.schroeder@upf.edu

Ronald Schuyler
Centro Nacional de Analisis Genomico
ron.schuyler@gmail.com

Arnau Sebé-Pedrós
Institut de Biologia Evolutiva (CSIC-Universitat Pompeu Fabra)
arnau.sebe@ibe.upf-csic.es

Jennifer Semple
Centre for Genomic Regulation (CRG)
jennifer.semple@crg.eu

Eduard Serra
IMPPC
eserra@imppc.org

delegate contacts

Enric Serra Sanz
Centro Nacional de Análisis Genómico (CNAG)
eserra@pcb.ub.cat

Luis Serrano
Centre for Genomic Regulation (CRG)
luis.serrano@crg.eu

Priyanka Sharma
Centre for Genomic Regulation (CRG)
priyanka.sharma@crg.eu

Maria Shutova
VIGG RAS
mashutova@gmail.com

Francesc Sole
Josep Carreras Leukaemia Research Institute (IJC)
fsole@carrerasresearch.org

Carolina Soriano Tarraga
IMIM
csoriano@imim.es

Petra Stockinger
Centre for Genomic Regulation (CRG)
petra.stockinger@crg.eu

Francesco Strino
CLC bio
francesco.strino@clcbio.com

Ian Sudbery
University of Oxford
ian.sudbery@dpag.ox.ac.uk

Mónica Suelves Esteban
IMPPC
msuelves@imppc.org

Lauro Sumoy
IMPPC
lsumoy@imppc.org

Fran Supek
Centre for Genomic Regulation (CRG)
fran.supek@crg.eu

Hana Susak
Centre for Genomic Regulation (CRG)
hana.susak@crg.eu

David Tamborero
Pompeu Fabra University (UPF)
david.tamborero@upf.edu

Gian Gaetano Tartaglia
Centre for Genomic Regulation (CRG)
gian@tartaglialab.com

Vladimir Teif
German Cancer Research Center (DKFZ)
v.teif@dkfz-heidelberg.de

Ernest Terribas
IMPPC
eterribas@imppc.org

Sudhir Thakurela
Institute of Molecular Biology
s.thakurela@imb-mainz.de

David Torrents
ICREA-BSC
torrents.david@gmail.com

Kiana Toufighi
Centre for Genomic Regulation (CRG)
kiana.toufighi@crg.es

Olga Troyanskaya
Princeton University
ogt@genomics.princeton.edu

Baubec Tuncay
Friedrich Miescher Institute for Biomedical Research
tuncay.baubec@fmi.ch

Marek Tyl
Imperial College London
mtyl@imperial.ac.uk

Evrim Besray Unal
Centre for Genomic Regulation (CRG)
besray.unal@crg.es

Leonid Uroshlev
Insitute Of Molecular Biology RAS
leoniduroshlev@gmail.com

Alfonso Valencia
CNIO
valencia@cnio.es

Vanesa Valero Lázaro
IMPPC
vvalero@imppc.org

Tanya Vavouri
IMPPC
tvavouri@imppc.org

Ignacio Vázquez García
University of Cambridge and Wellcome Trust Sanger Institute
i.vazquez-garcia@damtp.cam.ac.uk

Carme Vidal
Hospital Son Espases
mariac.vidal@ssib.es

Natalia Vilor Tejedor
Centre for Research in Environmental Epidemiology (CREAL)
nvilor@creal.cat

Anna Vlasova
Centre for Genomic Regulation (CRG)
anna.vlasova@crg.eu

Adrià Voltes
DCEXS - Pompeu Fabra University
adria.voltes01@estudiant.upf.edu

Ilya Vorontsov
Vavilov Institute of General Genetics, RAS
vorontsov.i.e@gmail.com

Dieter Weichenhan
DKFZ Heidelberg
d.weichenhan@dkfz.de

Kyoung Jae Won
University of Pennsylvania
wonk@mail.med.upenn.edu

Roni Wright
Centre for Genomic Regulation (CRG)
roni.wright@crg.es

Fumiichiro Yamamoto
IMPPC
fyamamoto@imppc.org

Miyako Yamamoto
IMPPC
myamamoto@imppc.org

Jae-Seong Yang
Centre for Genomic Regulation (CRG)
jae-seong.yang@crg.es

Luis Zapata
Centre for Genomic Regulation (CRG)
luis.zapata@crg.es

Mihaela Zavolan
Barcelona Supercomputing Center(BSC)
mihaela.zavolan@unibas.ch

For further information on the IMPPC please write to:

Harvey Evans
Communications Manager
hevans@imppc.org
(+34) 93 554 3054
www.imppc.org

Integrative and Computational Biology Joint symposium

V IMPPC Annual Conference and 4DCellFate Workshop

http://www.imppc.org/congress/v_annual_conference/program/

For more information about 4DCellFate please contact:

Veronica A. Raker, Ph.D.
Scientific Project Manager
Center for Genomic Regulation (CRG)
Veronica.Raker@crg.eu
Tel. (+34) 933160374
http://www.4dcellfate.eu/

Photographs: Albert Uriach
Designed by: malabars (www.malabars.com)

Conference Organizers

IMPPC Institut de Medicina Predictiva i Personalitzada del Càncer

4DCellFate

With the Support of

Generalitat de Catalunya Departament de Salut

CRG Centre for Genomic Regulation

7TH FRAMEWORK PROGRAMME

Silver Sponsors

THE COMPANY OF Biologists

Sponsors

EMBO 50 years

www.s-n-m.eu Scientific Network Management

Media Partners

(GIGA)n SCIENCE