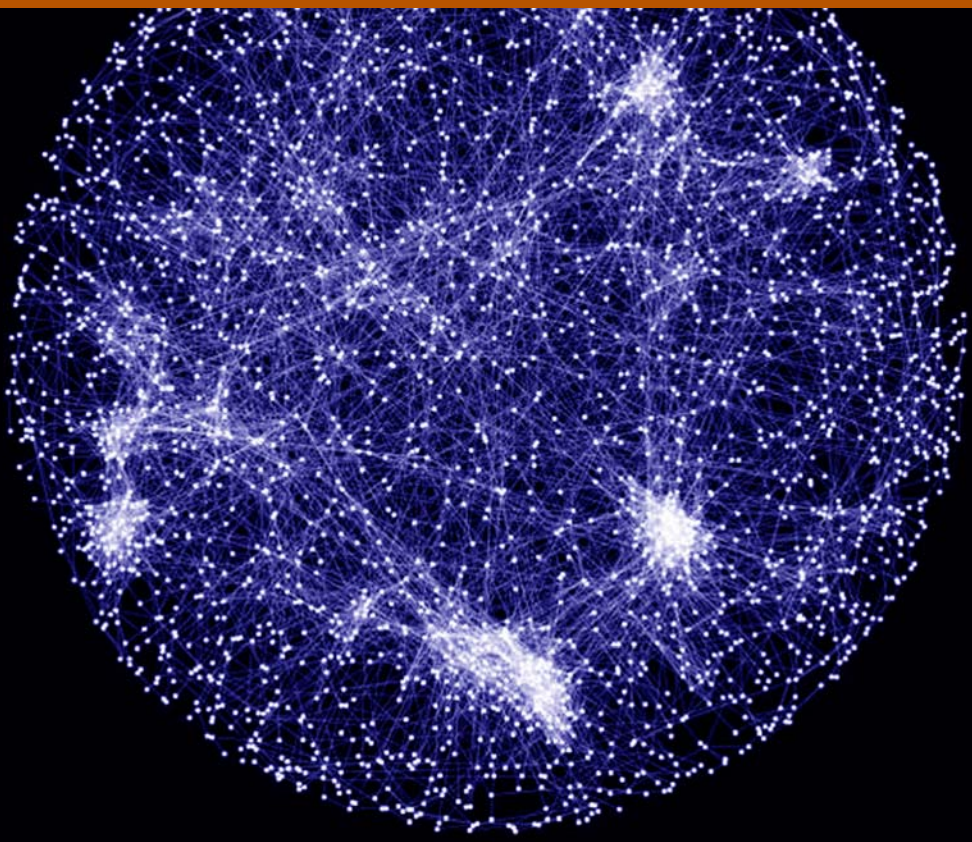
Joint Cold Spring Harbor Laboratory/Wellcome Trust Conference

SYSTEMS BIOLOGY: NETWORKS

August 11–August 15, 2010

View metadata, citation and similar papers at core.ac.uk



CSH

wellcome trust

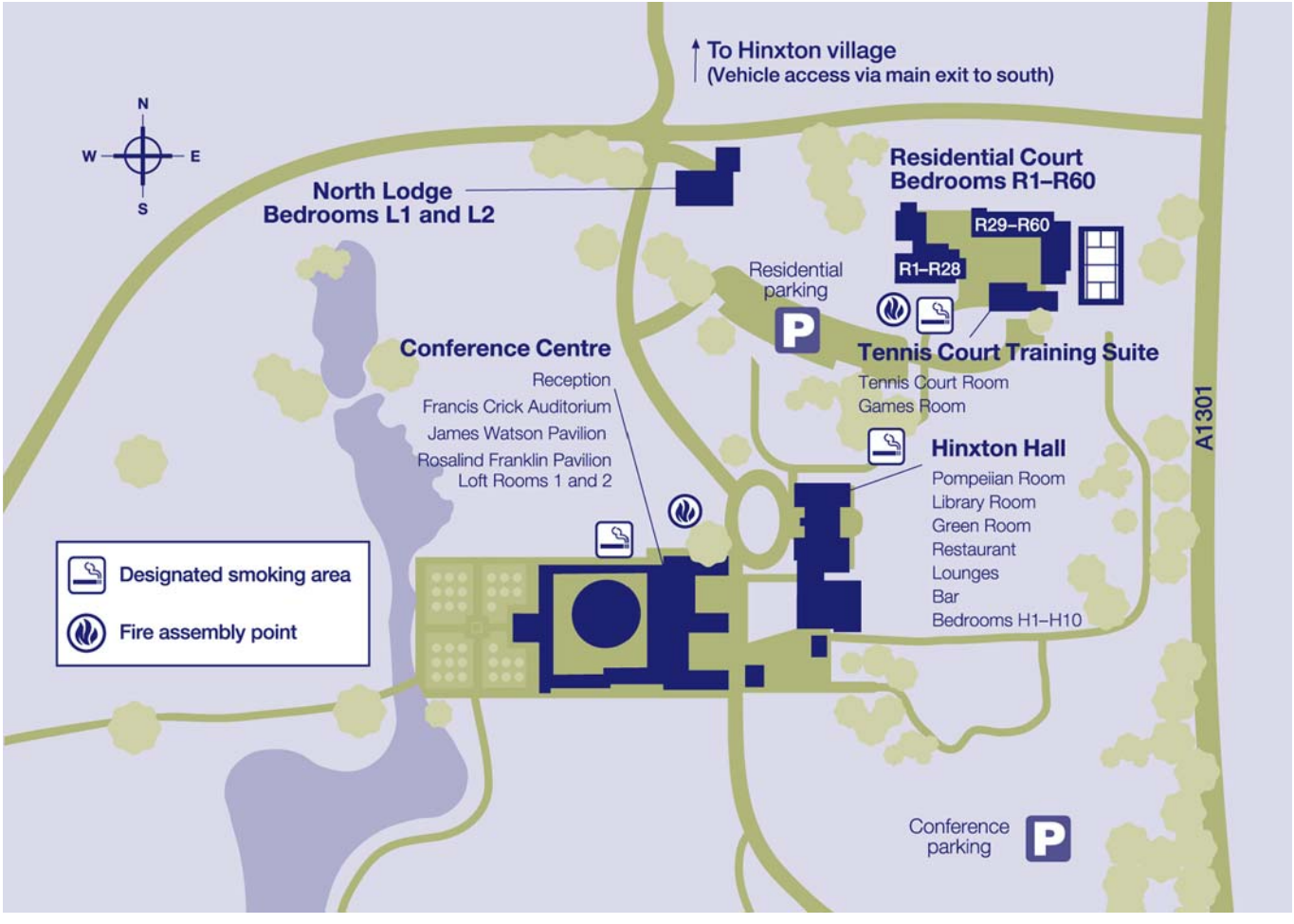
Joint Cold Spring Harbor Laboratory/Wellcome Trust Conference

SYSTEMS BIOLOGY: NETWORKS

August 11–August 15, 2010

Arranged by

Patrick Aloy, *Institute for Research Biomedicine, Spain*
Anne-Claude Gavin, *EMBL, Germany*
Trey Ideker, *University of California, San Diego, USA*
Marian Walhout, *University of Massachusetts Medical School, USA*


To Hinxton village
(Vehicle access via main exit to south)
N
W
E
S
North Lodge
Bedrooms L1 and L2
Residential Court
Bedrooms R1–R60
R29–R60
R1–R28
Residential
parking
P
Conference Centre
Reception
Francis Crick Auditorium
James Watson Pavilion
Rosalind Franklin Pavilion
Loft Rooms 1 and 2
Tennis Court Training Suite
Tennis Court Room
Games Room
Hinxton Hall
Pompeiian Room
Library Room
Green Room
Restaurant
Lounges
Bar
Bedrooms H1–H10
A1301
Designated smoking area
Fire assembly point
Conference
parking
P

SCHEDULE AT A GLANCE

Wednesday 11th August 2010

14.30-17.30	Session 1: Gene Regulatory Networks I
15.40-16.00	Break
16.00-17.30	Session 1, continued
17.30-18.30	Keynote Speaker: Charlie Boone
18.30-19.30	Wine and Cheese Party
19.30-21.00	Dinner

Thursday 12th August 2010

07.30-09.00	Breakfast
09.00-10.40	Session 2: Protein-Protein Interaction Networks
10.40-11:00	Morning Coffee
11:00-12:30	Session 2, continued
12.30-14.00	Lunch
14.00-15.10	Session 3: Genetic Networks
15.10-15.30	Break
15.30-17.00	Session 3, continued
17.00-19.00	Poster Session I
19.00-21.00	Dinner

Friday 13th August 2010

07.30-09.00	Breakfast
09.00-10.40	Session 4: Network Medicine
10.40-11.00	Morning Coffee
11.00-12.30	Session 4, continued
12.30-14.00	Lunch
14.00-15.20	Session 5: Gene Regulatory Networks II
15.20-15.40	Break
15.40-17.10	Session 5, continued
17.10-19.00	Poster Session II
19.00-21.00	Dinner

Saturday 14th August 2010

07.30-09.00	Breakfast
09.00-10.30	Session 6: Signalling Networks
10.30-10.50	Morning Coffee
10.50-12.20	Session 6, continued
12.30-14.00	Lunch
14.00-15.20	Session 7: Metabolic Networks
15.20-15.40	Break
15.40-16.40	Keynote Speaker: Edward Marcotte
17.00-19.00	Pre-dinner Drinks
19.00-21.00	Conference Dinner

General Information

Welcome to the Wellcome Trust Conference Centre and the **Systems Biology: Networks** Conference.

Conference Badges
Please wear your name badge at all times to promote networking and to assist staff in identifying you.

If you have advised us of any dietary requirements, you will find a small coloured dot on your badge. Please make yourself known to the catering team and they will assist you with your meal request.

Scientific Session Protocol
Photography, audio or video recording of the scientific sessions is not permitted.

Internet Access
Wireless internet access is available throughout the campus. Please inquire at reception for a Wireless Connection token.

Presentations
If you are an invited speaker or your abstract has been selected for an oral presentation, please provide an electronic copy of you talk to a member of the AV team who will be based in the auditorium.

Poster Sessions
Posters will be displayed throughout the conference. Please post your materials in the Cloisters on arrival. The abstract page number indicates to the assigned poster board number

Conference Meals
All meals will be served in the Hall Restaurant. Please refer to the conference programme in this book as times will vary based on the daily scientific presentations.

Social Events
The Hall Bar (cash bar) will be open from 19.00 – 23:00 each day.

Wednesday, 11 August – A Wine and Cheese reception will take place in the Conference Centre Cloisters from 18.30-19.30

Thursday, 12 August – A drinks reception will take place in the Conference Centre Cloisters from 17.00 during poster session I.

Friday, 13 August – A drinks reception will take place in the Conference Centre Cloisters from 17.10 during poster session II.

Saturday, 14 August – Pre dinner drinks will take place in the Conference Centre Cloisters from 17.00.

All conference meals and social events are for registered delegates only.

For Wellcome Trust Conference Centre Guests

Check in

If you are staying on site at the Wellcome Trust you may check into your room from 2.00pm. If you plan to arrive late at night you can check into your room as the conference centre reception is open 24 hours. Please note there will be no lunch or dinner facilities available outside of the conference timetable; however, there is a local public house (The Red Lion), serving both lunch and evening meals, located just 2 minute walk from the campus in the village of Hinxton.

Breakfast

Your breakfast will be served in the Hall restaurant from 07.30 – 09.00 every morning.

Telephone

If you are staying on-site and would like to use the telephone in your room, you will need to contact the Reception desk (Ext. 5000) to have your phone line activated - we will require your credit card number and expiry date to do so.

Departures

You must vacate your room by 10.00 on the day of your departure. Please ask at reception for assistance with luggage storage in the Conference Centre.

Red Lion Public House Guests

Check in

If you are staying at the Red Lion Public House, you are able to check into your room from 13.00-15.00 or 18.00-23.00

Breakfast

Your breakfast will be served at the Red Lion from 07.30-08.30 Monday - Friday and 08.30-09.30 Saturday - Sunday.

You must vacate your room by 10.30 on the day of your departure. A luggage store is available in the Conference Centre please ask at the reception.

Return Ground Transportation

Complimentary return transportation to Heathrow, Stansted Airport and the Cambridge Train Station and City Centre have been arranged for 09.30 on Sunday, 15 August. Please note: a sign-up sheet will be available at the registration desk. Places are limited so you are advised to book early.

Taxis

Please find a list of local taxi numbers should you require:

Panther – 01223 715715
Mid Anglia - Tel: 01223 836000
Phil's Taxi Services - Tel: 01223 521918
A&M Carriages (Airport Specialist) - Tel: 01223 513703

Messages and Miscellaneous

All messages will be posted on the registration desk in the Conference Centre Foyer.

A number of toiletry and stationery items are available for purchase at the conference centre reception. Cards for our self-service laundry are also available.

If you have any queries or comments, please do not hesitate to contact a member of staff who will be pleased to help you.

Joint Cold Spring Harbor Laboratory/Wellcome Trust conferences at Hinxton are supported in part with funding courtesy of The Wellcome Trust.

These abstracts should not be cited in bibliographies. Material contained herein should be treated as personal communication and should be cited as such only with consent of the author.

Printed on 100% recycled paper.

*Front Cover: T*he genetic landscape of a yeast cell as described by Costanzo et al., Science. 2010 Jan 22;327(5964):425-31. The image was generated by Anastasia Baryshnikova.

PROGRAM

WEDNESDAY, August 11—2:30 PM

SESSION 1 GENE REGULATORY NETWORKS I

Chairperson: **L. Serrano,** Centre for Genomic Regulation, Barcelona, Spain

Pascal Braun
Presenter affiliation: Dana-Farber Cancer Institute.

Protein-protein interaction network of post-transcriptional gene expression regulation in *Trypanosoma cruzi*
Christian M. Probst, Henrique Preti, Daniela P. Pavoni, Fabricio K. Marchini, Newton M. Vidal, Marco A. Krieger.
Presenter affiliation: Carlos Chagas Institute - ICC/FIOCRUZ, Curitiba, Brazil. 1

High-resolution transcriptome-wide identification of HuR regulatory elements
Neelanjan Mukherjee, Jeff Nusbaum, David Corcoran, Uwe Ohler, Thomas Tuschl, Jack Keene.
Presenter affiliation: Duke University Medical Center, Durham, North Carolina. 2

Individual variation in gene expression in genetic and network studies
Renuka R. Nayak, Vivian G. Cheung.
Presenter affiliation: Howard Hughes Medical Institute, University of Pennsylvania, Philadelphia, Pennsylvania. 3

Investigating the connection between the nuclear organization and the primary structure in senescent chromatin
Tamir Chandra, Jean-Yves Thuret, Lixiang Xue, Rekin's Janky, Shamith A. Samarajiwa, Scott Newman, Masako Narita, Agustin Chicas, Paul A. Edwards, Madan M. Babu, Scott W. Lowe, Simon Tavare, Masashi Narita.
Presenter affiliation: Cancer Research UK, Cambridge, United Kingdom. 4

WEDNESDAY, August 11—5:30 PM

KEYNOTE SPEAKER

WEDNESDAY, August 11—6:30 PM

Wine and Cheese Party

THURSDAY, August 12—9:00 AM

SESSION 2 PROTEIN-PROTEIN INTERACTION NETWORKS

Chairperson: **V. Cheung,** University of Pennsylvania, Philadelphia, USA

SESSION 3 GENETIC NETWORKS

Chairperson: **H. Zhu,** Johns Hopkins University, Baltimore, Maryland, USA

POSTER SESSION I

FRIDAY, August 13—9:00 AM

SESSION 4 NETWORK MEDICINE

Chairperson: **P. Braun,** Dana-Farber Cancer Institute, Boston, Massachusetts, USA

FRIDAY, August 13—2:00 PM

SESSION 5 GENE REGULATORY NETWORKS II

Chairperson: **G. Superti-Furga,** Austrian Academy of Sciences, Vienna, Austria

FRIDAY, August 13—5:10 PM

POSTER SESSION II

See Poster Session I for list of posters.

SESSION 6 SIGNALLING NETWORKS

Chairperson: **A. Godzik,** Burnham Institute for Medical Research, La Jolla, California, USA

SATURDAY, August 14—2:00 PM

SESSION 7 METABOLIC NETWORKS

Chairperson: **D. Pe'er,** Columbia University, New York, New York, USA

SATURDAY, August 14—3:45 PM

KEYNOTE SPEAKER

Deaf plants, bleeding yeast, and other surprising disease models from deeply conserved gene networks
Edward M. Marcotte, Kriston L. McGary, Tae Joo Park, Hye Ji Cha, John O. Woods, Martin Blom, John B. Wallingford.
Presenter affiliation: University of Texas, Austin, Texas. 73

SATURDAY, August 14

CONFERENCE BANQUET

5:00 PM Drinks 7:00 PM Dinner

PROTEIN-PROTEIN INTERACTION NETWORK OF POST-TRANSCRIPTIONAL GENE EXPRESSION REGULATION IN TRYPANOSOMA CRUZI

Christian M Probst[1,2], Henrique Preti[1], Daniela P Pavoni[1], Fabricio K Marchini[1,2], Newton M Vidal[2], Marco A Krieger[1]

[1]Carlos Chagas Institute - ICC/FIOCRUZ, Functional Genomics Laboratory, Curitiba, 81350-010, Brazil, [2]Carlos Chagas Institute - ICC/FIOCRUZ, Bioinformatics Laboratory, Curitiba, 81350-010, Brazil

Trypanosoma cruzi is a protozoan parasite, affecting more than 15 million people in South and Central Americas and having inefficient treatment. Besides being a public health problem, it has peculiar characteristics that make it an interesting study object for basic biology; also, it is an ancient eukaryote, and its study can provide clues to cellular processes distinct from those of other model organisms. One of its major features is the regulation of gene expression, which occurs post-transcriptionally, and studying the processing, storage and degradation of mRNA is extremely important for understanding the gene expression control mechanisms of this organism. RNA-binding proteins are extremely important in regulating these processes, as through their interaction with specific mRNAs they can define mRNA fate. We are conducting several large scale analyses of *T. cruzi* molecular biology, including RNA-seq and mass spectrometry, which are providing important clues about *T. cruzi* biology. Their integration is expected to improve our knowledge, but this is not trivial; a possible way to improve the initial integration is to build a map of protein-protein interaction (PPI). Aiming to create an initial map of RNA-binding PPI, we have constructed a *T. cruzi* ORFeome, comprising ~8,000 proteins. We have selected ~400 proteins, consisting of all putative RNA-binding or those that are part of potential interacting functional complexes, as ribosome, splicing and decay machinery, among others and have screened them by a yeast two-hybrid (Y2H) array approach. Using two stringency criteria, we have identified 82 and 34 interactions, of 73 and 29 proteins, respectively. Some of these interactions were also identified in other interactome datasets, but only for a small fraction. Although this raises some concerns about false positives, we hypothesize that, due to its large evolutionary distance from other model organisms whose interactome is well studied, and to the specificity of its post-transcriptional gene expression regulation, comparison of interactome datasets based on orthologs is of low power to detect true interactions. We are currently performing *in vivo* analysis of selected interactions to assess the specificity of our results. Besides that, we are integrating the several large scale data we have generated in our group over the interaction network, as well as increasing the coverage of the whole *T. cruzi* interactome.

HIGH-RESOLUTION TRANSCRIPTOME-WIDE IDENTIFICATION OF HUR REGULATORY ELEMENTS

Neelanjan Mukherjee[1], Jeff Nusbaum[2], David Corcoran[3], Uwe Ohler[3], Thomas Tuschl[2], Jack Keene[1]

[1]Duke University Medical Center, Molecular Genetics and Microbiology, Durham, NC, 27110, [2]Rockefeller University, Laboratory of RNA Molecular Biology, New York, NY, 10065, [3]Duke University Institue for Genome Science and Policy, Biostatistics & Bioinformatics, Durham, NC, 27710

RNA-binding proteins (RBPs) and noncoding RNAs are posttranscriptional regulatory factors that control the fate of each mRNA species. RBPs coordinate all aspects of RNA processing including splicing, export, localization, stability, and translation. HuR, a member of the ELAV/Hu family of RBPs, has been shown to positively regulate the stability and translation of target mRNAs through AU-rich elements (AREs) typically found in the 3' untranslated region (UTR) of mRNAs encoding many immediate early genes, inflammatory cytokines, and growth factors. While many studies have identified mRNAs associated with HuR, the precise binding sites have not been identified on a genome-wide scale. Furthermore, the extent to which HuR regulates mRNA processing events, such as splicing and poly-adenylation, has not been examined globally. We identified precise HuR-RNA interaction sites transcriptome-wide at a high resolution utilizing a recently developed technology, PAR-CLIP (Photoactivatable-Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation). As expected there were many binding sites in the 3' UTR of mRNAs. Surprisingly, there were a substantial number of intronic binding sites, suggesting HuR is involved in pre-mRNA processing. Utilizing Affymetrix exon arrays, we identified mRNAs that require HuR for maintaining their expression levels and RNA processing. Integration of binding data with knockdown expression data reveals that HuR positively regulates the expression of target mRNAs. Furthermore, the degree of binding was proportional to extent of HuR-dependent regulation of gene expression. Unexpectedly, the positive regulation remained true for mRNAs that only contained intronic HuR binding sites.
Integration of the data revealed known functions of HuR, for example HuR positively regulates the expression of functionally related proteins involved in the regulation of cell cycle. We also discovered novel potential cellular functions of HuR. Our data suggest that HuR directly regulates the expression of mRNAs encoding amino-acyl tRNA-synthetases (ARSs). The majority of the twenty ARSs are positively regulated by HuR and contain HuR binding sites. Nine ARSs form a macromolecular complex which include 3 other non-ARS proteins. The mRNAs encoding these three proteins are also directly bound and regulated by HuR. The tRNA multi-synthetase macromolecular complex, a highly multifunctional complex important in many diseases, may represent a post-transcriptional RNA operon coordinated by HuR.

INDIVIDUAL VARIATION IN GENE EXPRESSION IN GENETIC AND NETWORK STUDIES

Renuka R Nayak[1], Vivian G Cheung[2]

[1]Medical Scientist Training Program, University of Pennsylvania, Philadelphia, PA, 19072, [2]HHMI/ U Pennsylvania, Pediatrics & Genetics, Philadelphia, PA, 19072

We use the extensive variation in expression levels of human genes to study gene functions and regulations. We measured gene expression of human B-cells from hundreds of unrelated and related individuals. Using these data, we carried out genetics of gene expression (GOGE) studies and constructed gene coexpression networks.

In GOGE studies, we treated expression levels of genes ("expression phenotypes") as quantitative traits and identified regulatory variants that influence individual differences in gene expression. We found that about 10% of these regulatory variants act in cis to influence gene expression, the remaining ones act in trans. We validated these findings molecularly by gene knockdown and metabolic assays. The results include over 1,000 polymorphic regulator-target gene pairs.

In parallel, we use gene expression data of >8 million gene pairs to construct coexpression network. The resulting networks consist mainly of large connected components and have scale-free topologies. We found that neighboring genes often share similar functions; thus, the networks allowed us to identify biological processes and characterize functions of genes.

In this presentation, I will describe the results of using variation in human gene expression in genetic and network studies to identify gene functions and regulatory relationships. I will also discuss implications of these findings for human diseases.

INVESTIGATING THE CONNECTION BETWEEN THE NUCLEAR ORGANIZATION AND THE PRIMARY STRUCTURE IN SENESCENT CHROMATIN

Tamir Chandra[1], Jean-Yves Thuret[2], Lixiang Xue[1], Rekin's Janky[3], Shamith A Samarajiwa[1], Scott Newman[4], Masako Narita[1], Agustin Chicas[5], Paul A Edwards[4], Madan M Babu[3], Scott W Lowe[5], Simon Tavare[1], Masashi Narita[1]

[1]CRUK , Cambridge Research Institute, Cambridge, CB20RE, United Kingdom, [2]CEA , iBiTec-S, Paris, 91191, France, [3]Medical Research Council - Laboratory of Molecular Biology, Structural Studies Division, Cambridge, CB20QH, United Kingdom, [4] Cambridge University, Department of Pathology and Hutchison/MRC Research Centre, Cambridge, CB2 0XZ, United Kingdom, [5]Cold Spring Harbor Laboratory, Howard Hughes Medical Institute, Cold Spring Harbor, NY, 11724

We have previously identified a global and progressive heterochromatic structure within senescent human diploid fibroblasts – senescence associated heterochromatic foci (SAHFs) – which have since been widely used as a marker of senescence. Our current model is that SAHFs contribute to the altered gene expression profile seen in senescent cells. However, the structural detail of SAHFs is still unclear. Here we take cell-biological and genomic approaches to analyze in great detail how histone marks are reorganized during SAHF formation. We find that SAHFs are distinct multi-layer structures, in which H3K9me3 is enriched in the 'core' of SAHFs, surrounded by an H3K27me3 layer, which separates the core from the transcriptionally active H3K4/36me3 region. Consistent with our model, chromatin IP coupled with deep sequencing (ChIP-seq) for histone marks in these three layers reveals that these marks are dynamically redistributed at subsets of genic regions, including cyclin A and p16INK4A. Strikingly, however, the data also show that the global 'landscapes' of the silencing marks (H3K9me3 and H3K27me3) are largely unchanged, despite such a drastic alteration in chromatin structure. This indicates that the non-genic 'stable' H3K9me3 mark nucleates SAHFs, and that the global re-distribution or accumulation of H3K9/27me3 is not necessary for SAHF formation. Thus our combined approach of cell biology and biochemistry provides insight into the relationship between the higher-order and primary structures of silencing histone marks.

TRANSCRIPTION NETWORKS IN THE DNA DAMAGE STRESS RESPONSE IN ES CELLS*

R Siddappa, L von Stechow, J Carreras Puigvert, B van de Water, E H.J. Danen

Leiden University/LACDR, Division of Toxicology, Leiden, 2300RA, Netherlands

Cells experience constant genomic insult by exogenous and endogenous DNA damaging agents. If the damage is unrepaired or escapes the surveillance machinery, this can lead to ageing, degenerative diseases and cancer. DNA damage is sensed by ATM/ATR protein kinases and activates downstream checkpoint kinases and a p53 signaling cascade to control repair, cell cycle and survival. In this project, we make use of a gene family-wide short interference (si)RNA transcription factor library to identify transcription factors that are involved in the apoptotic response to various DNA-damaging drugs in ES cells. Primary high-throughput screening for cell viability has revealed potential regulators of the sensitivity of ES cells to the DNA damaging drug cisplatin. The hits are validated in a secondary deconvolution screen for cell survival and further evaluated in a real time apotosis assay using automated imaging. From ~1600 transcription factors tested in the primary screen, 140 have been identified as potential hits, and have entered the validation process. In parallel, transcriptomics data have been obtained from similar treatments of ES cells exposed to Cisplatin. Bioinformatics analysis of these two datasets is used to link identified transcription factors to transcriptional regulation of their target genes. We aim to construct a global transcriptional network model for DNA damage signaling in ES cells. This will be further tested for multiple compounds and cell types, including cancer cells. This will allow us to define potential biomarkers and new druggable targets in order to improve chemotherapeutic treatments.
*Supported by the Netherlands Toxicogenomics Center.

IMEX – WHEN THE MAJOR MOLECULAR INTERACTION DATABASES COLLABORATE TOGETHER TO SERVE YOU BETTER.

Samuel Kerrien, The IMEx Consortium, Henning Hermjakob

European Bioinformatics Institute, Proteomics Services Team, Hinxton, CB101SD, United Kingdom

Molecular interaction data is now collected by an increasing number of databases, leaving the user with a bewildering choice of where to look for information on their protein of interest. Several resources exist which merge the content of a number of databases, but the management of redundant data is often challenging. The International Molecular Exchange consortium (IMEx [1]) is a group of major public interaction data providers sharing curation effort who have agreed to

· Develop and work to a single set of curation rules when capturing data from both directly deposited interaction data or from publications in peer-reviewed journals
· Capture full details of an interaction in a "deep" curation model
· Perform a complete curation of all protein-protein interactions experimentally demonstrated within a publication
· Make these interaction available in a single search interface on a common website
· Provide the data in standards compliant download formats
· Release a significant number of IMEx records every year
· Make all IMEx records freely accessible under the Creative Commons Attribution License

Current IMEx members are IntAct, MINT, DIP, MPact, MatrixDB, MPIDB, BioGRID and Molecular Connections. From production start in February 2010 to May 2010, the IMEx consortium has released 50,359 interactions as IMEx records. The dataset will automatically increase with every release by a participating, member database. Data is made available in standards-compliant tab-delimited and XML formats, enabling users to visualize the data in a wide range of tools such as Cytoscape and the R Bioconductor software. IMEx data and documentation is available from [1]. IMEx data access is based on PSICQUIC, the PSI Common QUery InterfaCe. PSICQUIC is an effort from the HUPO Proteomics Standard Initiative (HUPO-PSI) to standardise the access to molecular interaction databases programmatically. PSICQUIC is a client-server technology that allows users to effortlessly access the IMEx data and over 2,000,000 binary interactions that have been shared across 12 data providers so far. The PSICQUIC Registry [2] is available to expose the list of services currently available. The PSICQUIC view [3] is a client that enables users to query all PSICQUIC registered service from a single web site.

[1] http://imex.sourceforge.net/
[2] http://www.ebi.ac.uk/Tools/webservices/psicquic/registry/registry?action=STATUS
[3] http://www.ebi.ac.uk/Tools/webservices/psicquic/view/

THE GENETIC LANDSCAPE OF A CELL

Charles Boone

University of Toronto, Banting and Best Department of Medical Research, Toronto, M4T 1K9, Canada

A genome-scale genetic interaction map was constructed by examining 5.4 million gene-gene pairs for synthetic genetic interactions, generating quantitative genetic interaction profiles for ~75% of all genes in the budding yeast, Saccharomyces cerevisiae. A network based on genetic interaction profiles reveals a functional map of the cell in which genes of similar biological processes cluster together in coherent subsets and highly correlated profiles delineate specific pathways to define gene function. The global network identifies functional cross-connections between all bioprocesses, mapping a cellular wiring diagram of pleiotropy. Genetic interaction degree correlated with a number of different gene attributes, which may be informative about genetic network hubs in other organisms. Extensive and unbiased mapping of the genetic landscape also provides a key for interpretation of chemical-genetic interactions and drug target identification.

BIOCHEMICAL APPROACHES TO BIOMOLECULAR NETWORKS

Anne-Claude Gavin

EMBL, Heidelberg, D-69117, Germany

Since the sequencing of the first eukaryotic genome, *Saccharomyces cerevisiae*, more than 10 years ago, explosion of new analytical tools in the fields of transcriptomics, proteomics and metabolomics contributes ever-growing molecular repertoires of the building blocks that make up a cell. Biology does not rely on biomolecules acting in isolation. Biological function depends on the concerted action of molecules acting in protein complexes, pathways or networks. Biomolecular interactions are central to all biological functions. In human, for example, impaired or deregulated protein–protein or protein–metabolite interaction often leads to disease. Recent strategies have been designed that allow the study of interactions more globally at the level of entire biological systems. We will discuss the use of these biochemical approaches to genome-wide screen in model organisms.

A DYNAMIC VIEW OF DOMAIN-MOTIF INTERACTION NETWORKS

Hanah Margalit, Eyal Akiva

The Hebrew University of Jerusalem, Microbiology and Molecular Genetics, Jerusalem, 91120, Israel

Cellular protein interaction networks involve many transient interactions that are mediated by domains that bind short linear motifs. These domain-motif pairs define functional units that are often involved in key cellular processes and need to be tightly regulated. A common strategy of the cell to control protein function and interaction is by post-translational modifications of specific residues, especially phosphorylation. Indeed, there are examples where domain-motif interactions are enhanced or prevented by phosphorylation of the motif's residues. Here we carry out a large-scale integrative analysis to examine the extent of this mechanism in the cellular networks. By integrating large-scale experimental data of domain-motif interactions and of phosphorylation events we demonstrate an intriguing coupling between the two, where residue phosphorylation in or near the motifs is implied to be associated with prevention or enabling of the domain-motif interaction. We study the evolutionary traces of the coupling between domain-motif interaction and phopshorylation and its functional implications. Our results suggest that phosphorylation plays a role as a specificity switch between proteins that bind to the same protein regions either through the same domain or through different domains. Our findings suggest that in many cases the interaction between domains and linear motifs depends on the phopsphorylation state of the motifs, adding a dynamic aspect to domain-motif interaction and defining interaction-regulation modules.

NETALIGNER – FAST AND ACCURATE ALIGNMENT OF PROTEIN INTERACTION NETWORKS.

Roland A Pache[1,2], Patrick Aloy[1,2,3]

[1]Institute for Research in Biomedicine (IRB) Barcelona, Structural and Computational Biology, Barcelona, 08028, Spain, [2]Barcelona Supercomputing Center (BSC), Life Sciences, Barcelona, 08034, Spain, [3]Institució Catalana de Recerca i Estudis Avançats (ICREA), Life & Medical Sciences, Barcelona, 08010, Spain

Recent advances in genome sequencing and large-scale interaction detection provided comprehensive lists of the macromolecules present in an organism and how they interact to create complex biological systems. To understand and interpret this huge amount of data, we need novel bioinformatics approaches. Based on the success of sequence alignment in unveiling genome function, organization and evolution, alignment of protein interaction networks is expected to vastly increase our understanding of biological processes, their evolution and adaptation to changing environmental conditions. However, existing network alignment programs still suffer from certain limitations that prevent them from being widely applied.

Here, we present a novel network alignment algorithm, NetAligner, that addresses those limitations, featuring both fast alignment of query pathways or protein complexes to whole species interactomes and fast interactome to interactome alignment for finding conserved complexes or subnetworks between two organisms. NetAligner is able to align networks of arbitrary topology and to accurately model evolutionary duplication events through many-to-many homology relationships. In addition to considering interaction reliabilities to address potential false positives, NetAligner is the first network alignment algorithm that offers the option to predict conserved interactions to counter the high amount of false negatives in current interactomes, considerably increasing alignment performance in some cases.

We conducted several independent benchmarks, in which NetAligner consistently outperformed the current standard in the field, while requiring only a fraction of the runtime. This, together with its fast assessment of the statistical significance of alignment solutions, makes NetAligner ideal for large-scale application in comparative interactomics experiments.

TRANSIENT COMPLEXES ALONG PROTEIN-PROTEIN ASSOCIATION

Michal Harel, Gideon Schreiber

Weizmann Institute of Science, Biological Chemistry, Rehovot, 76100, Israel

How do proteins associate quickly and specifically is a question of ongoing interest in the emerging fields of systems biology, rational protein design, medicine and more, as these reactions are involved in most biological processes. Proteins interact with variable affinities from mM to sub-pM and at rates of association and dissociation spanning over six orders of magnitude, yet their degree of specificity is not directly related to either their affinity or rate of binding. Many proteins can interact with multiple partners either simultaneously or separately, and this is one of the reasons for their ability to form complex networks. These observations are merely part of the great interest in this field of study.

The process of protein-protein association begins with their random collision, which may further develop into an encounter complex, followed by a transition state, and finally complex formation can occur. We combined experimental and computational tools in order to characterize the nature of both the transition state and the encounter complex for protein-protein association. The transition state was analyzed by double mutant cycles and computational modeling, while the encounter complex was examined by a combination of Brownian dynamics simulations and experimental examination. Two types of transition states were found: specific and diffusive, which were interchangeable by rational design. While the specific transition state can be structurally modeled, the diffusive transition state lacks preferred structures, which may result in a slower association. The encounter complex analysis indicated that encounters faraway from the interface are futile, while encounters at the vicinity of the interface are fruitful and contribute to the association reaction. Furthermore, the free energy landscape of the encounter complex can be drastically changed upon mutation. The comparison between the encounter complex and the transition state suggests an energetic funnel, in which the encounter complex is in the broader side of it, while the transition state is in the more narrow side, towards the final complex at the bottom of the funnel. Together these observations provide important insights about the protein-protein association reaction.

SYSTEMS BIOLOGY ANALYSIS OF A SMALL BACTERIUM: M. PNEUMONIAE

Luis F Serrano

Centro Regulacion Genomica, Systems Biology, Barcelona, 08003, Spain

To understand basic principles of bacterial metabolism organization and regulation, but also the impact of genome-size thereon, we systematically studied one of the smallest bacteria, Mycoplasma pneumoniae. A manually curated metabolic network of 189 reactions catalyzed by 129 enzymes allowed the design of a defined, minimal medium with 19 essential nutrients. More than 1,300 growth curves were recorded in the presence of various nutrient concentrations. Measurements of biomass indicators, metabolites and 13C-glucose provided information on directionality, fluxes and energetics; integration with transcription profiling enabled global analysis of metabolic regulation. Compared to more complex bacteria, the M. pneumoniae metabolic network has a more linear topology and contains a higher fraction of multifunctional enzymes; general features such as metabolite concentrations, cellular energetics, adaptability and global gene expression responses are similar though. In parallel we have combined strand-specific tiling arrays, complemented by transcriptome sequencing, with more than 252 spotted arrays. We detected 117 previously undescribed, mostly non-coding transcripts, 89 of them in antisense configuration to known genes. We identified 340 operons, of which 140 are polycistronic; almost half of the latter show decaying expression in a ‘staircase-like’ manner. Under various conditions, operons could be divided into 447 smaller transcriptional units, resulting in many alternative transcripts. Frequent antisense transcripts, alternative transcripts, and multiple regulators per gene imply a highly dynamic transcriptome, more similar to that of eukaryotes than previously thought. Finally we used Tandem Affinity Purification-Mass Spectrometry (TAP-MS) in a proteome-wide screen. The analysis revealed 62 homomultimeric and 116 heteromultimeric soluble protein complexes, of which the majority are novel. About a third of the heteromultimeric complexes show higher levels of proteome organization, including assembly into larger, multi-protein complex entities, suggesting sequential steps in biological processes, and extensive sharing of components implying protein multifunctionality. Incorporation of structural models for 484 proteins, single particle EM and cellular electron tomograms provided supporting structural details for this proteome organization. The dataset provides a blueprint of the minimal cellular machinery required for life.

FROM BIOLOGICAL NETWORKS TO PHYLOGENY AND DISEASE

Natasa Przulj

Imperial College London, Department of Computing, London, SW7 2AZ, United Kingdom

Genes produce thousands of different protein types that interact in complex networked ways and make cells work. It is possible that interaction network data will be as useful as the sequence data in uncovering new biology. Given the abundance of interaction data, systems-level comparisons of networks of pathogenic and non-pathogenic species could play a vital role in understanding mechanisms of pathogenicity. Also, comparing networks of healthy and disease-affected cells could deepen our understanding of disease and lead to identification of cellular parts that are candidates for therapeutic intervention. Furthermore, biological network comparison and alignment could enable transfer of knowledge between species, since we may know a lot about bio-molecules in one species and almost nothing about aligned bio-molecules in another species.

Existing network alignment methods use information external to network topology, e.g., sequence data. Since network topology provides a new and independent source of biological information, it is important to understand how much biology we can learn from it independently from any other data source. Hence, we develop mathematically rigorous ways for aligning networks based solely on their topology. Our network aligners produce by far the most complete alignments of biological networks to date, exposing large and contiguous regions of network similarity even for as distant species as yeast and human thus suggesting broad similarities in internal cellular wiring across all life on Earth. Moreover, they demonstrate that protein function and species phylogeny can be extracted solely from network topology. Analogous to reconstruction of phylogenetic trees using sequence similarities, we use our network alignment similarities to successfully reconstruct the phylogenies of protists, fungi, and herpesviruses.

In addition, we show that network topology around cancer and non-cancer genes is different and use this to predict new cancer genes. Our predictions are phenotypically validated. Next, we show that aging, cancer, and pathogen-interacting genes have topologically “central” roles in the network, as they form the “spine” of the network that connects all other network parts and is hence vital for normal cellular functioning. Also, we show that sequence and topology give insights into complementary slices of biological information and that network topology can be used as a complementary method to sequence-based homology identification.

We present evidence that topology-based analyses of biological networks provide new biological and phylogenetic insights and that they can help identify novel drug targets, hence aiding therapeutics and health care.

A PROTEIN INTERACTION WIRING OF THE HUMAN SPLICEOSOME

Anna Hegele[1], Atanas Kamburov[1], Crysovalantis Sourlis[1], Arndt Grossmann[1], Sylvia Wowro[1], Mareike Weimann[1], Vlad Pena[2], Reinhard Lührmann[2], Ulrich Stelzl[1]

[1]Max Planck Institute for Molecular Genetics (MPI-MG), OWL, Berlin, 14195, Germany, [2]Max Planck Institute of Biophysical Chemistry, Department of Cellular Biochemistry, Göttingen, 37077, Germany

Splicing of pre-messenger RNA is catalyzed by the spliceosome, which consists of the five U1, U2, U4, U6 and U5 RNAs and between 180 and 230 proteins in human. On each mRNA the spliceosome assembles and disassembles afresh. This spliceosomal assembly cycle is a highly dynamic process. Well defined large RNP complexes are formed in an ordered, stepwise manner. For example a total of ~ 65 proteins exchange during the conversion of the B to the C complex, which encompasses catalytic activation. Both, RNA-protein and protein-protein interactions are crucial for the formation, rearrangements and dissociation of the complexes.
We set out to systematically characterize human spliceosomal protein-protein interactions. A yeast two-hybrid matrix screen examining all pair wise interactions between ~ 230 spliceosomal or spliceosome associated proteins, represented by ~ 430 cDNA clones, was carried out. The screen revealed more than 500 putative protein-protein interactions. Most of the interactions are novel; however about 30% of the current PPI knowledge in the literature was recapitulated. Selected protein-protein interactions are analyzed on the domain level and a large fraction of PPIs is being validated in co-immunoprecipitation experiments in mammalian cells.
Computational analysis of the PPI network and integration with complex purification data provides novel insight into the spliceosomal assembly cycle. This includes e.g. hints to how the Prp19 complex is recruited to the B complex, the assembly of B-complex specific proteins or the interaction of step2 proteins such as hSlu7 or Prp16 with protein components of the core C complex. It also suggests important roles for some of the previously uncharacterized spliceosomal proteins.

PROTEIN COMPLEXES AND NETWORKS FOR GREEN BIOTECH

Jelle Van Leene[1], Dominique Eeckhout[1], Geert Persiau[1], Eveline Van De Slijke[1], Erwin Witters[2], Geert De Jaeger[1]

[1]VIB-Ghent University, Plant Systems Biology, Gent, B-9000, Belgium, [2]VITO-University of Antwerp, Ceproma, Antwerp, B-2020, Belgium

A central theme of todays plant biotech is the discovery of new genes or combinations of genes whose engineering can lead to crops with increased biomass production or resilience in sub-optimal environmental conditions. Building protein interaction networks forms an attractive new approach to find new knowledge for such applications. We developed a high throughput TAP technology platform for complex isolation from Arabidopsis cell suspension cultures and applied it to map the core complex machinery at the heart of the Arabidopsis cell cycle control. The resulting network offers for the first time a comprehensive view on CDK/cyclin complexes of a higher plant and contains over 100 new candidate cell cycle genes. Besides a central regulatory network of core complexes, we distinguished a peripheral network that links the core machinery to up- and downstream pathways. Furthermore, we could demonstrate that plants have evolved a combinatorial toolkit comprising at least a hundred different CDK/cyclin complex variants, which strongly underscores the functional diversification among the large family of cyclins and reflects the pivotal role of cell cycle regulation in the developmental plasticity of plants. Perturbation of a first set of the new candidate cell cycle genes lead to interesting phenotypes, showing larger biomass.

ANALYZING REGULATORY CIRCUITRY WITH EXPRESSION-PROFILE PHENOTYPES

Frank Holstege

University Medical Center Utrecht, Genomics Laboratory, Utrecht, 3508 AB, Netherlands

The availability of whole genome sequences and the parallel development of various high-throughput techniques is making it possible to analyze and understand regulatory processes in a global manner. We are systematically generating DNA microarray mRNA expression-profiles of targeted mutations in components of the signaling and transcription machinery in the yeast *S. cerevisiae*. One of the first completed sets encompasses all 150 non-essential protein kinases and phosphatases. Linear signal transduction pathways are readily identified by this approach as well as new cellular roles for previously uncharacterized kinases and in some cases direct kinase/phosphatase targets. Since one goal is to understand the entire regulatory circuitry active under a single condition, we have also systematically analyzed genetic redundancy, by profiling all kinase and/or phosphatase double-deletions that show negative synthetic genetic interactions. This reveals three different redundancy phenotypes: complete, quantitative and incongruent. Incongruency originates from partially redundant pairs that are coupled through additional regulatory links, and is the most frequently observed redundancy phenotype. Much of the redundancy involves non-homologs. This includes phosphatase-kinase redundancy, caused by phosphatase-mediated inhibitory cross-talk between kinase pathways. Mechanisms underlying incongruent phenotypes plausibly explain the evolutionary maintenance of genetically redundant pairs, based on superior combinatorial control potential involving modules of partially redundant genes linked by unidirectional repression or inhibition of each other. These pairs can be directly redundant paralogs or indirectly redundant non-homologs. Both form recurrent regulatory modules suitable for coupled and uncoupled regulation of different combinations of cellular responses, dependent on condition or context. The consequence of many such modules when wired together is the potential to deftly control a large number of completely different combinations of processes through different activity mixes of relatively few pathways. This suggests a common way in which combinatorial control is achieved through signaling.The results of profiling other sets of regulators such as chromatin remodellers and modifiers, will also be discussed.

EPISTASIS IN CONTEXT: UNDERSTANDING GENETIC INTERACTIONS THROUGH AN EXHAUSTIVE MODULAR DECOMPOSITION

Jeremy Bellay[1], Gowtham Atluri[1], Gaurav Pandey[1], Charles Boone[2], Vipin Kumar[1], Chad Myers[1]

[1]University of Minnesota, Computer Science & Engineering, Minneapolis, MN, 55455, [2]University of Toronto, Terrence Donnelly Centre for Cellular and Biomolecular Research , Toronto, M5S 3E1, Canada

Genetic interactions provide a powerful perspective into biological processes that is fundamentally different from other high-throughput genome-wide studies. Recently, Synthetic Genetic Array technology was used to produce the first genome scale map of digenic genetic interactions, which covered ~5.4 million genetic interactions or about ~30% of all possible gene pairs in yeast. This provides a first opportunity for a global, unbiased assessment of the structure of the genetic interaction network and the relationship between this structure and individual gene function. We developed a data mining approach based on association rule learning to exhaustively discover all block structures within the yeast genetic interaction network, producing a complete modular decomposition of the network. We find that genetic interaction hubs can be clearly differentiated into distinct classes of hubs based on their modular structure and that these distinctions correspond to date and party hubs previously observed in protein-protein interaction networks. Moreover, module membership provides a specific and unbiased assessment of the prevalence of multi-functionality among genes: we find that genes participate in far more functions or contexts than was previously appreciated. In addition, we find that genetic interactions contained within structured modules exhibit strikingly different functional properties relative to isolated interactions, providing insight into the evolution and functional divergence of duplicate genes. Finally, the modular structure of the genetic interaction network gives surprising evidence both for and against the traditional between-pathway and within-pathway models of genetic interactions.

FUNCTIONAL GENE NETWORK OF A PATHOGENIC BACTERIUM; *PSEUDOMONAS AERUGINOSA*

Sun-Gou Ji, Insuk Lee

Yonsei University, Biotechnology, Seoul, 120-746, South Korea

Due to their simple growth requirements and nutritional versatility, members of the genus *Pseudomonads* are ubiquitous organisms that can grow in virtually any environment and gain resistance from many antibiotics. The best-known of them - *Pseudomonas aeruginosa*, has been reported as one of the most frequent cause of nosocomial infections, as well as the predominant cause of morbidity and mortality in cystic fibrosis patients. However, the mechanisms of its pathogenesis and antibiotic resistance have not been completely understood yet.
Here, we present a genome-scale functional network of *P. aeruginosa* genes reconstructed by integrating currently available data directly derived from *P. aeruginosa*, as well as from other bacteria such as *E. coli*, *H. pylori*, etc. Novel interactions could be inferred through the guilt-by-association method using various kinds of binary associations such as functional genomics, proteomics and comparative genomics data sets. These included DNA microarray co-expression datasets downloaded from GEO by NCBI and ArrayExpress by EBI, and protein-protein interactions from small scale experiments databases such as DIP, BOND, IntAct, MPIDB, etc. Furthermore, large-scale TAP-MS, LCMS, MALDI-TOF or Yeast 2-hybrid experiments of various bacterial species from published literatures were processed and integrated into the network. In addition, interactions from phylogenetic profiles and genomic context similarities where also included, plus, associations transferred from yeast, worm, human and *A. thaliana* genes from previously constructed networks were incorporated.
Similar modeling approaches have already been successfully applied to intensively studied model organisms such as, *S. cerevisiae*, *C. elegans*, *A. thaliana*, etc, but this time, a bacterial pathogen was subjected to this approach. The system-level functional network model of the genome of *P. aeruginosa* would open new avenues to unravel the underlying interplay among genes involved in bacterial pathogenesis, and lead to a further comprehension of functional genomic interactions between a pathogen and its host.

WHY WOULD A MUTATION KILL ME, BUT NOT YOU?

Ben Lehner

EMBL-CRG Systems Biology Unit and ICREA, Centre for Genomic Regulation, Barcelona, 08003, Spain

Mutations do not always have the same effect in different individuals. For example, most "disease mutations" in humans do not actually cause disease in all (or even most) of the people who carry them. What are the reasons for this? One aspect is, of course, the environment. Another aspect is, however, the "genetic environment" and we have been using *C. elegans* as a model system to understand - and to learn how to predict - how mutations combine epistatically to cause phenotypic change. Based on this work we are now also addressing the problem of non-genetic variation in mutation outcome. That is, we are dissecting the stochastic inter-individual differences that determine whether an individual is affected by a mutation or not.

SOX9-DIRECTED GENE REGULATORY NETWORK CONTROLLING MURINE EMBRYONIC SKELETAL DEVELOPMENT

Sook Peng Yap, Thomas Lufkin

Genome Institute of Singapore, Stem Cell and Developmental Biology, Singapore, 138672, Singapore

Sox9 is critical for embryonic chondrogenesis and skeleton formation, yet few direct target genes are known, limiting our understanding of this biological process. We reveal here for the first time a highly detailed gene regulatory network (GRN) anchored by Sox9. Expression profiling of wildtype, heterozygous and null Sox9 cells isolated from staged living embryos, identified 6124 differentially expressed genes as Sox9 “downstream” targets. In situ hybridization validated key targets, including Col2a1, Matn4 and Hapln1. In parallel Sox9 chromatin immunoprecipitation-sequencing (ChIP-Seq) mapped the Sox9 binding genomic loci. Along with the canonical Sox9 binding motif, we also identified 2 novel Sox9 motifs and further confirmed their direct binding and regulation by Sox9. The integration of gene expression profiling and Sox9 DNA binding data has enabled us to identify a set of Sox9 direct and in-direct target genes and successfully led to the elucidation of the Sox9-directed GRN controlling mammalian skeletogenesis.

FUNCTIONAL VS. TOPOLOGICAL PREDICTIONS OF ESSENTIAL GENES OF THE FOOD-BORNE PATHOGEN *CAMPYLOBACTER JEJUNI*

Aline Metris, Mark Reuter, Duncan J Gaskin, Arnoud H van Vliet, Jozsef Baranyi

Institute of Food Research, Foodborne Bacterial Pathogens, Norwich, NR4 7UA, United Kingdom

The emergence of network science has raised the question of whether the topology of the networks of genes or substrates in a cell could explain its functions. One such function is the determination of essential genes; those that are indispensible for growth under laboratory conditions. The highest accuracy of predicting essential genes in micro-organisms in a given environment is obtained with functional analysis such as Flux Balance Analysis (FBA), where 90% of essential genes may be predicted for the most studied micro-organisms such as *E. coli* (Feist *et al.*, 2007, Mol Systems Biology, 3:121). These predictions are only matched in terms of accuracy by the topological network analysis of the "synthetic accessibility" (Wunderlich and Mirny, 2006, Biophys J, 91:2304-11). This method may also be classified as functional analysis since it analyses the metabolic network where the purpose is the production of substrates. Topological networks, such as the Protein-Protein interaction network can point to gene essentiality though analysis of hubs and bottlenecks (Jeong *et al.*, 2001, Nature 411:41-42, Yu *et al*., 2009, PLoS Comput Biol 3:e59). However, it remains difficult to predict essential genes from such a network with accuracy. Other indicators such as the connectivity of a "molecular network" have also been investigated to determine essential genes (del Rio *et al*., 2009, Systems Biol, 3:102).
For micro-organisms other than the established paradigm model systems, reconstruction of metabolic network is more delicate and the accuracy to predict essential genes by established methods like FBA is much lower. In order to predict essential genes in the food-borne pathogen *Campylobacter jejuni*, we have reconstructed a genome scale metabolic network based on published genome annotations and on-line databases such as BioCyc and KEGG. The genome of this pathogen is small (about a third of *E. Coli's*) and presents some hypervariable sequences in genes linked to the surface structure (Parkhill *et al.*, 2000, Nature, 403:665). While its central metabolism has been studied to some extent, much remains to be known concerning the composition and biosynthesis of the cell wall and capsule surrounding the cell. The network was validated against BIOLOG experimental data (Line *et al*., 2010, J. Microbiol. Methods, 80:198-202). We have drawn evidence from FBA, analysis of the Protein-Protein interaction topology (Parish *et al*., 2007, Genome Biol 8:R130), and experimental transposon mutagenesis to make predictions about essential genes in this pathogenic organism.

PROTEIN NETWORKS REVEAL DETECTION BIAS AND SPECIES CONSISTENCY WHEN VIEWED THROUGH INFORMATION--THEORETIC GLASSES.

Franca Fraternali[1], Luis P Fernandes[1], Jens Kleinjung[2], Alessia Annibale[1], Anthonius A Coolen[1]

[1]King's College London, Randall Division, London, SE1 1UL, United Kingdom, [2]NIMR MRC London, Mathematical Biology, London, NW7 1AA, United Kingdom

Protein–protein interaction networks (PPIN) are the backbone of cellular function and allow for the identification of several biological processes, such as important pathways and macromolecular complexes. Apart from the evident biological information provided by PPIN, they also provide a series of topological properties such as degree distributions, assortativity, average degree, clustering coefficient, etc. Among many other useful purposes for the use topological information, is the direct comparison of networks either between different species or between networks of the same specie obtained by different experimental techniques. Most studies so far have fallen short of a realistic topological comparison between networks, this mainly due to the fact that only degree distribution have been used for such comparisons. We apply a series of rigorous mathematical methods, which we have recently developed, to the topological exploration of several PPIN available in the public domain. We first apply a systematic analytical framework with which to quantify the macroscopic topological structure of PPIN along with the generation of reliable 'null models' (random graphs whose topological features can be controlled precisely). We calculate entropy measures that relate to network complexity. We define complexity according to the uniqueness of a network when compared to an equivalent null model network. Alongside those calculations we also developed methods that determine differences between networks purely based on topological properties and apply those methods to perform a species to species comparison. Although the availability of datasets with adequate coverage and of high accuracy is lagging, we have now created the necessary tools for a detailed topological characterization and direct network comparisons.

Annibale A., Coolen, AAC, Fernandes, LP, Fraternali, F, Kleinjung, JK. Tailored Graph Ensembles as Proxies or Null Models for Real Networks I: Tools for Quantifying Structure.
Journal of Physics A: Mathematical and Theoretical 2009 42 485001.

REDUCING REDUNDANCY IN BIOLOGICAL NETWORKS

Mochamad Apri[1,2], Jaap Molenaar[1,2], Simon van Mourik[1], Maarten de Gee[1]

[1]Wageningen University, Biometris, Wageningen, 6708PB, Netherlands,
[2]Netherlands Consortium for Systems Biology, CWI, Amsterdam, 1090GE, Netherlands

Most biological networks are very complex and may consist of hundreds to thousands components with intricate interactions. If we describe the dynamics of such networks in terms of ordinary differential equations, we are faced with the challenge to solve a very large nonlinear system of equations. Although such detailed models may lead to accurate predictions, their complexity is a serious disadvantage in several respects, e.g, the large computing times may obstruct the efficient estimation of parameters. Moreover, the interpretation of the behavior of huge networks may be hard. Since the dynamics of networks is often dominated by only a subset of the nodes and the interactions, it is very useful to develop methods that can detect redundancy in a given network and delivers a reduced network that still captures the key features of the dynamical behavior of the original network.

Here, we present a method to systematically reduce redundancy in a biological network and thus simplifies the biological interpretation of the dynamics of the original network. The method is based on a sensitivity analysis approach that has successfully been applied in combustion engineering problems. This approach has been adjusted to account for nonlinearities that may stem from, e.g., Michaelis-Menten dynamics. The idea is to remove redundant nodes and redundant interactions from the network. As a first step, one identifies an initial set of nodes that play a crucial role in the network. Then, the sensitivities of the dynamics of these nodes to the dynamics of the rest of the nodes are evaluated. Nodes that turn out to be important for the crucial nodes are then considered as crucial and added to the set of crucial nodes. After that the sensitivities are recalculated. This step is repeated until two complementary sets of nodes are found, one of crucial and one of redundant nodes. By removing the redundant nodes a reduced network is obtained. Next, this reduced network is checked on the presence of redundant interactions by calculating the so-called “rate sensitivity matrix”. The eigenvalues and the eigenvectors of this matrix give a way to order the parameters in increasing importance and to remove non-essential interactions. Eventually, we obtain a reduced biological network that still preserves the important features of the full network.

For illustrational purposes, the method is first applied to a small artificial network. Then, we apply it to a large network model underlying the cell fate determination in the floral meristem of Arabidopsis Thaliana and show that this model allows considerable reduction.

FORM, FUNCTION AND EVOLUTION OF SPECIALIZED CILIA TRANSPORT MACHINERY: COMPARING THE FOLD ARCHITECTURES AND INTERACTIONS OF BBSOME AND IFT PROTEIN COMPLEXES

J. Fernando Bazan[1], Maxence V Nachury[2]

[1]Genentech, Depts. of Protein Engineering and Structural Biology, South San Francisco, CA, 94080, [2]Stanford University School of Medicine, Dept. of Molecular and Cellular Physiology, Stanford, CA, 94305

Nearly every cell in the human body has primary cilia, striking antenna-like projections with a microtubule core that are required for diverse sensory functions including phototransduction, olfaction, planar cell polarity and morphogen signaling. Targeting of transmembrane receptors to the tip of the cilium is critical for proper sensing and transduction of signals, and the molecular machinery responsible for their sorting and trafficking has been discovered to be a series of large protein assemblies. One of these involves an octameric complex of proteins genetically linked by mutation to Bardet-Biedl Syndrome (a group of severe human developmental disorders caused by ciliary dysfunction) and is called the BBSome. The structural anatomy of the BBSome has been explored with sensitive fold recognition, structure prediction and modeling tools, and reveals an economical architectural toolkit that is remarkably similar to COPI, COPII and clathrin coat complexes. Notably, BBS4 and BBS8 are almost entirely comprised of helical TPR-class repeats and are predicted to form rigid, rod-shaped α-solenoid structures. By contrast, we show that BBS1, BBS2, BBS7 and BBS9 have a common domain architecture distinguished by a toroidal β-propeller fold with an appendage segment similar to those found in several clathrin adaptors and in coatomer; this appendage domain is typically comprised of two protein recruitment modules, an Ig-like β-sandwich γ-adaptin ear (GAE) motif coupled to an α/β platform structure. The most conserved subunit, BBS5, contains distant pleckstrin homology (PH)-like folds that bind phosphoinositides and associate BBS5 with membranes; BBS3 is the small Arf-like GTPase Arl6 and most likely involved in the recruitment of a coat-like complex to membranes. In the present work, we contrast the fold architecture and emerging interaction map of the BBSome coat-like scaffold components with those of the two very large intraflagellar transport (IFT) multiprotein complexes A and B that are critically involved in movement of membrane ciliary cargo by engaging molecular motors that travel up and down the microtubule axoneme.The final compendium of protein folds dedicated to serve as struts and joints of coat-like assemblies (the helical solenoid rods and β-propeller toroids, respectively), or as specialized interaction and targeting modules (like the appendage and PH domains in the BBSome, or novel calponin-like domains in IFT proteins) should help clarify the evolutionary emergence and retained (or divergent) molecular functions of BBSome and IFT subunits. Key to these comparative structural studies are interaction maps that aid in the spatial packing of modular components, and also identify symmetric subcomplexes within BBSome and IFT systems that may represent both evolutionary relics and smaller functional machines active outside ciliary transport.

METABOLIC NETWORK DYNAMICS DURING SEED DEVELOPMENT IN *ARABIDOPSIS*

Raju Datla[1], Xiang Daoquan[1], Prakash Venglat[1], Chabane Tibiche[2], Gopalan Selvaraj[1], Mathieu Cloutier[2], Edwin Wang[2]

[1]National Research Council of Canada, Plant Biotechnology Institute, Saskatoon, S7N 0W9, Canada, [2]National Research Council of Canada, Biotechnology Research Institute, Montreal, H4P 2R2, Canada

Seed development is coordinated with synthesis and deposition of storage products in many plant species. The complex interconnected regulatory programs involving the embryo development and metabolism that define qualitative and quantitative aspects of seeds are largely unknown. To address this critical gap, we have used integrated systems approach and generated large genomic, proteomic and metabolomic datasets for key stages of embryo development in model plant *Arabidopsis*. Analysis of these datasets provided critical frame work to develop metabolic network models. To test the biological implications of these models, two putative regulatory nodes associated with fatty acid and carbohydrate synthesis are selected and the associated genes encoded by "*ACC1*" and "*Epimerase*" have been investigated using loss and gain of function studies. Genetic lesions in both loci independently lead to embryo lethality with significant down regulation of genes connected to the affected nodes. Further, in *acc1* mutant line, the production of a number of fatty acids including long chain is severely affected and interestingly increased the levels of some of the carbohydrates. Similarly, in epimerase mutant line, the production of some of the carbohydrates significantly reduced with increase in the synthesis of some fatty acids. Consistent with these observations, *Epimerase* over expression in the seeds showed reduction in the overall oil content. Together, these studies provide important regulatory connections between the synthesis and deposition of carbohydrates and lipids, the two major storage products in the seed. Thus, our results provide direct evidence and insights into metabolic network models in *Arabidopsis* seed.

SLIMFINDER: A WEB SERVER TO FIND NOVEL, SIGNIFICANTLY OVER-REPRESENTED, SHORT PROTEIN MOTIFS

Norman E Davey[1], Niall J Haslam[2], Denis C Shields[2], Richard J Edwards[3]

[1]European Molecular Biology Laboratory, Structural and Computational Biology Unit, Heidelberg, 69117, Germany, [2]University College Dublin, School of Medicine and Medical Sciences, Dublin, D4, Republic Of Ireland, [3]University of Southampton, School of Biological Sciences, Southampton, SO16 7PX, United Kingdom

Short, linear motifs (SLiMs) play a critical role in many biological processes, particularly in protein–protein interactions. The SLiMFinder (Short, Linear Motif Finder) web server is a *de novo* motif discovery tool that identifies statistically over-represented motifs in a set of protein sequences, accounting for the evolutionary relationships between them. Motifs are returned with an intuitive *p*-value that greatly reduces the problem of False Positives and is accessible to biologists of all disciplines. Input can be uploaded by the user or extracted directly from UniProt. Numerous masking options give the user great control over the contextual information to be included in the analyses, including evolutionary conservation and protein disorder prediction. The SLiMFinder server combines these with user-friendly output and visualisations of motif context to allow the user to quickly gain insight into the validity of a putatively functional motif. These visualisations include alignments of motif occurrences, alignments of motifs across homologous sequences and a visual schematic of the top-ranked motifs. Returned motifs can also be compared to known SLiMs from the literature using CompariMotif. All results are available for download.

In addition to the main SLiMFinder webserver, additional recent developments of methods and resources for SLiM discovery will be outlined.

The SLiMFinder server is available at:
http://bioware.ucd.ie/slimfinder.html.

FUNCTIONAL ANALYSIS OF TISSUE-SPECIFIC PROTEIN INTERACTIONS BASED ON RNA-SEQUENCING DATA

Dorothea Emig, Mario Albrecht

Max Planck Institute for Informatics, Computational Biology and Applied Algorithmics, Saarbrücken, 66123, Germany

Proteins are essential for all kinds of biological processes. They usually act in concert by forming pairwise interactions or complexes through functional subunits such as protein domains. Tissue-specific gene expression can lead to the presence or absence of proteins, resulting in the gain or loss of protein interactions and complexes in certain tissues. Over the last years, several studies have been performed aiming at the identification of universal and tissue-specific gene expression and their functional implications. These studies were commonly based on extensive microarray data such as the Novartis Gene Atlas. Microarrays, however, have the drawback that they cannot accurately measure gene expression at low levels. This may lead to the misclassification of low expressed genes as absent, which introduces a bias towards high expressed genes into molecular studies.
With the advent of second-generation RNA-sequencing techniques, it is now possible to sequence complete transcriptomes and measure gene expression with higher accuracy. Therefore, we used publicly available RNA-sequencing data for fifteen tissues and cell lines to identify universal and tissue-specific expression of proteins, protein interactions, protein domains, and complexes. Our results show that by far fewer protein interactions are tissue-specific than previously thought. Our Gene Ontology analysis indicates that tissue-specific interactions are highly enriched in receptor and transmembrane activities, with substantial differences in the numbers of specific interactions among all tissues and cell lines.
Furthermore, we investigated the tissue-specificity of protein domains. We collected all genes encoding a certain domain and computed the average number of tissues per domain according to the gene expression. Interestingly, our results show that only few domains are tissue-specific. Furthermore, our results indicates that a high percentage of these tissue-specific domains mediate protein-DNA interactions rather than protein-protein interactions, which is also confirmed by GO enrichment analysis. In addition, we find that most known protein complexes obtained from CORUM and PDB are completely expressed (i.e. all co-complexed proteins are expressed) in all tissue and cell line samples, regardless of the complex size. In contrast, a closer examination of only tissue-specific complexes suggests that tissue-specific protein complex formation is frequently controlled by one or few tissue-specific proteins, while the majority of co-complexed proteins are universally expressed. Overall, our new insights are of great relevance not only for studying gene expression, protein functions, and protein interactions, but also for analyzing diseases and possible drug targets.

INFERRING GENE REGULATORY NETWORKS FROM EXPRESSION DATA: PRINCIPLES AND ANALYSIS

Frank Emmert-Streib, Gokmen Altay

Queen's University Belfast, Center for Cancer Research and Cell Biology, Belfast, BT9 7BL, United Kingdom

The inference of causal interactions from high-throughput data is important because the analysis of the resulting networks may provide novel insights into the functioning of the underlying molecular system. Despite its importance, the systematic analysis of statistical and computational approaches is still in its infancy. Our contribution consists of two parts. First, we address the question of experimental design and discuss various evaluation measures. Specifically, we study network-based evaluation measures in combination with ensemble data. Second, we discuss a novel inference methods, C3NET, and present results from a comparative study.

A NETWORK-BASED APPROACH FOR ESTABLISHING CONNECTIONS BETWEEN DISEASES AND DRUGS

Hai Fang, Owen Rackham, Julian Gough

Department of Computer Science, University of Bristol, Bristol, BS8 1UB, United Kingdom

Background: It is well known that disease-causing genes (and drug-targeted genes) are more likely to interact with each other in the interactome network. However, there is lack of an analysable framework to systematically exploit connections between diseases and drugs.
Methodology: Here, we report a network-oriented approach to identify all the possible components of the connections between diseases and drugs. We reason that these high-level relationships among diseases and drugs are principally encoded in the gene-level topological structure of the compiled human gene network. The strength of relationships can be measured by the observed links against the expected links under the binomial distribution-based model. Specifically, we first classified diseases from OMIM database into 22 main disease categories based on the organ system affected, and drugs from DrugBank database into 14 main drug groups according to the Anatomical Therapeutic Chemical classification system. Then, we quantified the topological relationships between disease genes and drug targets in the human gene network, resulting in interconnectedness matrix between/within OMIM diseasome categories and DrugBank drug groups.
Results and Discussion: As expected, genes causing the same type of disease or targeted by same group of drug are significantly linked together in the network. Inspections of the disease-disease block show that high interconnectivity is biased towards the most common diseases (i.e., respiratory, hematological, endocrine, cancer, and immunological disease classes), forming a fully interconnected clique. Others are less interconnected. Co-existence of the disease clique and less-interconnected disorders reflects the underlying evolutionary constraints on pathogenesis. Further inspections of the drug-drug block show that all of drugs are highly interconnected to form a drug clique, partially indicating the current practices of using off-target drugs to combat the diseases. Interestingly, the drug clique tends to interconnect the disease clique, further indicating that the drug clique probably occurs during the efforts against these most common diseases in disease clique. A closer inspection of connections between drug and disease, both associated with the same category, shows that the successful on-target efforts in disease clique compromise the off-target consequences. These observations call for the more rational drug design to maximize the efficiency while minimizing the side effects.

Keywords: Network; interconnectedness; diseases; drugs

PATHWAYS TO COGNITIVE DEFICITS IN DOWN SYNDROME

Katheleen J Gardiner[1], Krzysztof J Cios[2]

[1]University of Colorado Denver, Pediatrics, Aurora, CO, 80045, [2]Virginia Commonwealth University, Computer Science, Richmond, VA, 23284

Down syndrome (DS), trisomy of human chromosome 21 (HSA21), is characterized by an average IQ of 45, and a constellation of learning/memory deficits implicating hippocampal dysfunction, in particular spatial learning. The phenotype is caused by the increased expression, due to gene dosage, of human chromosome 21 (HSA21) genes that include 160 protein coding genes and an additional 350 genes of unassigned function. To address this complexity, with the goal of identifying potential targets for pharmacoptherapies, we have implemented comprehensive pathway analysis, combining information on protein-protein interactions, genetics of intellectual disability (ID), drug-induced gene expression responses, and experimental analysis of mouse models of DS. Curated pathways from multiple databases (KEGG, Reactome, PID) were scored based on the number of HSA21 components, the number of HSA21 primary interactors (obtained from HPRD, BIND, BioGRID), and on the number of additional components known to cause ID when mutated (based on evidence in OMIM and the literature). The highest scoring pathways are those most likely to be perturbed in DS and to contribute to ID in DS. Predictions of pathway perturbations are tested in cell lines and brains of mouse models of DS that are trisomic for partially overlapping subsets of HSA21 orthologs and that display learning/memory deficits relevant to DS. We use the moderate through-put technique of Reverse Phase Protein Arrays, to generate protein profiles of HSA21 proteins and candidate phosphorylated and non-phosphorylated pathway components. Drugs that alter the expression of HSA21 genes and that impact high scoring pathways were identified from the Connectivity Map (The Broad Institute) and PubChem. Mouse models are being treated with candidate drugs and tested for corrections of abnormalities in the protein profiles and behavior. Computational methods are used to predict behavior from protein profiles and to identify the most significant protein abnormalities in DS models. We have identified perturbations in pathways involving NMDA receptor, calcineurin and MAP kinase signaling, oxidative stress and apoptosis. Machine Learning Methods have successfully predicted locomotor responses to MK-801 and memantine has corrected subsets of protein profile abnormalities in the major mouse model of DS. Pathway analysis can help to overcome the limitations of mouse models of DS and improve success rates in selecting drugs for preclinical evaluations.

SYSTEM-LEVEL CONTROL PROPERTIES OF THE EUKARYOTIC MRNA TRANSLATION PATHWAY

Shichina Kannambath, Helena Firczuk*, Jürgen Pahle, Hans Westerhoff, Pedro Mendes, John McCarthy

University of Manchester, Manchester Interdisciplinary Biocentre, Manchester, M1 7DN, United Kingdom

Precise control of the translation machinery underpins the capacity for the growth and selective competitiveness of living organisms, and therefore the quantitative properties of the overall system and the evolution of its molecular components are tightly coupled. Here, we identify the system-level control features of the entire eukaryotic translation pathway, thus identifying the contributions of all of the known yeast translation factors and defining the quantitative meaning of translational control. The highly non-stoichiometric in vivo complement of translation components is found to form an exquisitely balanced system in which no single factor or step is uniquely limiting or in excess of optimal system requirements. To summarize the key aspects of control, we have also incorporated our data into an in silico mathematical model of translation covering initiation, elongation and termination. A quantitative measure of system-specificity reveals how these proteins share their commitment across multiple cellular pathways. The comprehensive quantitative control map obtained for this modular gene expression system opens a new window on the adaptive consequences of collective molecular evolution.

IMPLEMENTING PHYLOGENETIC STATISTICAL ANALYSIS TO ESTABLISH GENE REGULATORY NETWORKS IN BACULOVIRUSES

Hagit Katzov-Eckert[1], Juliana Velasco-Oliveira[1], Luis Diambra[2], Paolo Zanotto[1]

[1]University of São Paulo, Laboratory of Molecular Evolution and Bioinformatics, Department of Microbiology, São Paulo, 05508-000 , Brazil, [2]University of São Paulo, Laboratory of System Biology, Department of Physics, São Carlos, 13560-970, Brazil

The Baculoviruses are important bio-control agents and biological vectors. Despite insights into their genome structure from sequencing studies, our understanding of gene regulation, interaction and function is limited. We use an evolutionary approach to reconstruct gene regulatory networks (GRNs) in baculovirses. We test the hypothesis that baculovirus gene function can be predicted by combining phylogenetic data and gene regulatory modules. All baculoviruses share a core set of twenty-nine genes involved in replication and transcription, suggesting common pathways for these vital functions. Gene expression in baculoviruses is also highly regulated and occurs in successive phases divided into early, late, and very late genes. In addition, phylogenetic analyses have demonstrated that the process of gene acquisition and gene loss are important factors in baculovirus evolution. We assess functional categories and analyze the dynamics of gene gain and loss along the phylogenetic tree of individual baculoviruses. Gene expression data was applied to reconstruct a phylogeny tree based on clusters of late expression genes of the Autographa californica multinucleocapsid virus (AcMPNV). We applied a maximum parsimony method that produces a phylogenetic classification that recognizes delimited clades shared by derived gene expressions of AcMPNV. We then established a phylogenetic gene cluster of expression genes to infer a topological relationship within AcMPNV genes. The derived tree showed one large inclusive clade that encompassed all of the late expression genes. The results provide the information necessary to predict functional gene links and contribute to our understanding on the regulators of late transcription. The GRN provides an informative predictive model for the identification of baculovirus gene function using phylogeny and expression profiles.

HIGH-THROUGHPUT SCREENING FOR VIRUS-HOST INTERACTIONS.

Manfred Kögl[1], Frank Schwarz[2], Haas Jürgen[3], Korn Bernhard[4]

[1]German Cancer Res. Center, Core Facilities Genomics and Proteomics, Heidelberg, 69120, Germany, [2]German Cancer Res. enter, Core Facilities Genomics and Proteomics, Heidelberg, 69120, Germany, [3]Max-von-Pewttenkofer Institut. LMU München, Lehrstuhl Virologie, München, 80336, Germany, [4] German Cancer Res. Center, Core Facilities Genomics and Proteomics, Heidelberg, 69120, Germany

We have systematically mapped the interactions of several influenza virus strains with host cells and integrated these results with data from genome-wide siRNA screens. Host proteins binding to viral proteins were identified by large-scale yeast two-hybrid (Y2H) screening. 111 protein-protein interactions were identified. 46 of the interactions were re-tested in LUMIER assays with a confirmation rate of 42%. Interactions were highly similar between viral strains, and overlapped to a small but significant extent with published interactions. Recently, five papers reported the results of genome-wide siRNA screens for proteins required for influenza virus replication. The extent of overlap between the host cell proteins identified in our yeast two hybrid data with those found in the siRNA is highly significant, and comparable to the overlap of the siRNA screens amongst each other. GO-term analysis shows that host cell proteins identified by siRNA screens and protein interaction studies are in part enriched in the same processes, namely protein folding, intracellular transport, RNA processing, cell cycle regulation and signal transduction. Several of the newly reported interactions provide molecular explanations for the identification of particular genes in siRNA screens. For example, siRNA screens identified ARCN1 as a host cell factor necessary for viral infection. ARNC1 is part of the COPI complex involved in early to late endosome transport vehicle formation. siRNAs vs ARCN1 and COPG blocked infection. The observation that ARCN1 binds to the viral protein M1 provides a molecular basis for this genetic interaction.

GENETIC NETWORKS THAT UNDERLY QUANTITATIVE AND QUALITATIVE ASPECTS OF
THE RNA TRANSCRIPTION IN BXH/HXB RATS.

Marieke Simonis, Vyacheslav Koval, Victor Guryev, Edwin Cuppen

Hubrecht Institute, Universitair Medisch Centrum Utrecht, Genome biology, Utrecht, 3584 CT, Netherlands

One of the challenges of modern functional genomics is to understand genetic
determinants and networks that affect quantitative and qualitative aspects of the RNA transcription.We use the BXH/HXB rats - one of the most popular models for human hypertension. The advent of next generation sequencing enables us to generate genome-wide maps of genomic variants and expression
profiles at base-pair resolution.

The RNA-Seq expression profiling was performed in liver and brain tissues using replicates from 3 animals each. Both normotensive BN-Lx and hypertensive SHR were investigated for differences in expression change. We
scored differences in gene expression, exon usage and splice forms between these rats.

Based on correlation of expression data we have constructed regulatory gene
networks for BN and SHR rats. Comparative analysis of these will help to detect genomic variants and biological pathways that determine phenotypic differences between the strains used and may provide insight in mechanisms
underlying common human disease.

COMBINED USE OF METABOLOMICS, LIPIDOMICS, FLUXOMICS AND BIOENERGENICS TO DETERMINE EARLIEST EFFECTS OF PPARΓ STIMULATION ON 3T3-L1 DIFFERENTIATION

Gustavo Palacios, Bhavapriya Vaitheesvaran, Haitao Lu, Irwin Kurland

Albert Einstein College of Medicine, Medicine/Endocrinology, Bronx, NY, 10461

Proteomic studies have shown that thiazolidinedione (TZD) PPARγ agents stimulate changes in mitochondrial mass and morphology of 3T3-L1adipocytes. Using an integrated physiology approach focusing on metabolic flux assessments and lipidomic profiling, we examined glycolytic, pentose and TCA metabolic pathways during 3T3-L1 differentiation to adipocyte, and how this was potentiated by the administration of the TZD rosiglitazone (rosi). 3T3-L1 cells were treated by the protocol from the Orphan nuclear Receptor Consortium, with induction for 2 days, and then differentiation medium with insulin ± rosi. Morphologic changes resulting from rosi were seen only from day 6. Using 2H2O assessment of protein synthesis, insulin + rosi significantly increased the rate by day 3 and the lipogenic rate by day 5. Measurements using the Seahorse XF96 Extracellular Flux analyzer showed rosi + insulin induced higher metabolic rates compared to insulin alone, with 40-50% increases in OCR by day 3 and in ECAR at day 4. This suggested increased mitochondrial function in the TZD treated 3T3-L1 adipocytes. Lipidomic studies, using a Thermo LTQ linear ion trap, showed that cardiolipin, an exclusive mitochondrial lipid, increased significantly by day 3 only in the rosi treated cells, suggesting that TZDs potentiate both mitochondrial mass and function. After induction, [1,2-13C]glucose flux studies showed that TCA glucose oxidation, as the proportion of anapleurotic TCA cycle flux, decreased in insulin + rosi treated cells, on days 3, 4 and 5 and was coincident with the increase in 2H2O assessment of the lipogenic rate in that interval. TCA glucose oxidation was increased on days 6-10 with rosiglitazone treatment, while TG content remained constant. Triacylglycerol species (TGAs) accumulated also increased notably under rosiglitazone treatment, from 30-100 fold by day 3. The increased TGAs, despite increased accumulation of ceramides, suggest the stimulation of insulin action by rosi. In conclusion, the combined use of metabolomics, lipidomics, fluxomics and bioenergenics can detect a series of metabolic effects induced by rosiglitazone much greater, and sooner, than significant morphological effects are seen. TZDs may alter mitochondrial energetics to foster biosynthesis by using mitochondrial fuel switching to maximize TG storage. This combined -omic approach may be useful for screening drug agents designed to affect bioenergenics or fat storage.

TRANSCRIPTIONAL ANALYSIS OF SALT TOLERANCE GENES IN TUNISIAN GRAPEVINES.

Ahmed MLIKI[1], Samia DALDOUL[1], Michael HOEFER[2]

[1]Centre de Biotechnologie de Borj-Cedria, Lab. Plant Molecular Physiology, Hamma-lif, 2050, Tunisia, [2]RLP AgroScience GmbH, RLP AgroScience GmbH, Neustadt a.d. Weinstrasse, 67435 , Germany

In order to understand the molecular basis of salt-stress response in grapevine plants, suppression subtractive hybridization (SSH) and microarray based screening approaches were combined. Two leaf-specific subtractive cDNA libraries were constructed from grapevine plants subjected to salt stress treatment. SSH were performed 6 and 24 h after addition of 100 mM NaCl in the nutritional solution using cDNAs prepared from leaves of a salt tolerant cultivar (Razegui) as testers and cDNAs from unstressed leaves as drivers. A pre-screened subset of cDNA clones from these SSH libraries were used to construct a Vitis vinifera cDNA array, in order to monitor the expression changes of the genes upon salt treatment. Expression profiles were compared between the salt-tolerant and a susceptible cultivar (Syrah) under both control conditions and after salt stress treatment. Seven cDNA clones were identified which were up-regulated by salt stress and confirmed by RNA blot analysis. The transcript expression patterns of the selected genes differed between the contrasting grapevines cultivars were tested with respect to stress-regulation.

PRIORITIZING CANDIDATE GENES BY NETWORK ANALYSIS OF DIFFERENTIAL EXPRESSION USING MACHINE LEARNING APPROACHES.

Daniela Nitsch, Yves Moreau

K.U. Leuven, ESAT-SCD, Leuven, 3001, Belgium

Discovering novel disease genes is still challenging for diseases for which no prior knowledge - such as known disease genes or disease-related pathways - is available. Performing genetic studies frequently result in large lists of candidate genes of which only few can be followed up for further investigation. We have recently developed a computational method that identifies the most promising candidate genes by replacing prior knowledge by experimental data of differential gene expression between affected and healthy individuals (Nitsch et al. 2009, PlosOne 4(5): e5526).
To improve the performance of our prioritization strategy, we have extended our previous work by applying different machine learning approaches that identify promising candidate genes by determining whether a gene is surrounded by highly differentially expressed genes in a functional association or protein-protein interaction network.
Our novel ranking strategies, scoring disease candidate genes, rely on network-based machine learning approaches, such as kernel ridge regression, heat kernel, and Arnoldi kernel approximation. For comparison purposes, a local measure based on the expression of the direct neighbors is also computed. We have benchmarked these strategies on 40 publicly available knockout experiments in mice, and performance was assessed against results obtained using a standard procedure in genetics that ranks candidate genes based solely on their differential expression level. Our results showed that our four strategies could outperform this standard procedure and that the best results were obtained using the heat kernel approach leading to an average ranking position of 8 out of 100 genes, an AUC value of 92.3% and an error reduction of 52.8% relative to the standard procedure approach which ranked the knockout gene in average at position 17 with an AUC value of 83.7%.

Conclusion: In this study we could identify promising candidate genes using network-based machine learning approaches even if no knowledge is available about the disease or phenotype.

STROLLING INSIDE THE COMPREHENSIVE HIV-1 HOST CELL INTERACTOME

Marcela Nunez[1], Petra Tafelmeyer[1], Jean-Christophe Rain[1], Fanny Moisant[1], Philippe Leclerc[1], Vincent Collura[1], Richard Benarous[2], Etienne Formstecher[1]

[1]Hybrigenics, Service, Paris, 75014, France, [2]CellVir SA, Biochemistry, Paris, 75014, France

More than 25 years after its discovery, the Human Immunodeficiency Virus 1 (HIV-1) is still one of the major threats on human life. Most anti-HIV drugs currently in use inhibit the catalytic activities of viral enzymes. However, a promising alternative is to target the interactions between viral proteins and host cells. In order to understand the complex interplay between HIV and its host, we have performed a comprehensive large-scale yeast-two hybrid study with all HIV-1 proteins as baits, using highly complex oligo_dT and random-primed cDNA libraries from the CEMC7 T cells line. A total of 34 screens resulted in the identification of 367 different host proteins. The interaction network was further densified by rebound screens with human proteins against a random-primed HIV-1 library. Here, we present an in depth analysis of our comprehensive HIV - human interactome and compare this network with the literature protein interaction data. In addition, we will demonstrate our sophisticated PIMRider software that allows straightforward analysis and navigation inside complex networks combining interaction, genetic and expression data from different sources.

PATHWAY RECONSTRUCTION WITH CELL-TYPE-SPECIFIC CO-EXPRESSION PROFILES IN *ARABIDOPSIS* ROOT

Taeyun Oh, Insuk Lee

Yonsei University, Life Science and Biotechnology, Seoul, 120-749, South Korea

How plant adapt to ever changing environment is a major question in plant biology. For better understanding of transcriptional regulations relate to phenotypic change, a variety of high-throughput genome-wide gene expression profiling had been conducted.

Multicellular organisms, however, reveal different transcriptional expression patterns in each cell-type. This cell-type specific transcriptional variation leads to the differential regulation of specific biological functions in subsets of cell layers, several of which correspond to recognizable physiological changes.

If we profile gene expression in whole tissue containing heterogeneous cell types, only cell-type independent expression response will be recognized but cell-type specific expression response could be buried by confounding effect derived from different cell-types.

With high-resolution root transcriptome profiling with various abiotic stress conditions and its tissue-level transcriptome control, we may compare effectiveness of pathway reconstruction between by whole-tissue expression profiles and by cell-type specific expression profiles.

ANALYSIS AND PREDICTION OF PROTEIN-MRNA INTERACTIONS IN YEAST

Vera Pancaldi, Jürg Bähler

University College London, Dep. Genetics, Evolution and Environment, London, WC1E 6BT, United Kingdom

In the past few years, we have learnt a lot about regulatory networks within the cell. While the entire physical and genetic interactomes have been extensively explored, both experimentally and computationally, relatively little experimental data is available for protein-mRNA interactions. Although the complex network of RNA-binding proteins and their targets plays an important role in gene expression control, interactions are known for only approximately 40 proteins in *S. cerevisiae*.

In this study, we first perform an analysis of the available data, showing that RNA-binding proteins tend to bind the cognate mRNAs of their protein interaction partners and that proteins that share an mRNA target are more likely to physically interact. We then assemble a database of features of both the proteins and their mRNA targets to examine their relationships. We include functional characterization through Gene Ontology terms as well as properties at the mRNA level (such as UTR characteristics, mRNA structure, translational properties, expression levels) and features of the protein (such as amino acid composition and physico-chemical properties). Besides numerous correlations amongst features of a single gene, we identify relationships between features of the protein and of its target, with the main finding that proteins with a higher isoelectric point tend to bind longer transcripts. Using the available protein-mRNA interaction data and an assembled randomized negative set as training, we combine the collected features in a Support Vector Machine to predict protein-mRNA interactions in budding yeast.

We achieve an estimated accuracy of around 70% in predicting targets of proteins for which some other mRNA targets are known, encountering a limitation of the method where there is no previous knowledge available on any of the protein's targets. Finally, we extend the method to predictions of fission yeast protein-mRNA interactions.

A REPRODUCIBLE SYSTEMATIC APPROACH FOR THE MAPPING OF HUMAN KINASES INTERACTOME

Roberto Sacco, Keiryn Bennett, Jacques Colinge, Giulio Superti-Furga

CeMM, Vienna, 1090, Austria

Kinases have a fundamental role in a great number of biological processes. However, a complete picture of the molecular networks associated to this class of proteins is still missing. Although the use of large scale proteomic approaches is of great value to draw such a map, one main criticism is that data reproducibility is generally low and too much dependent on the variability of certain parameters, such as the experimental and analytical set-ups established in each laboratory. We have started to systematically study the human kinome interaction network in the cellular environment, by applying a systems level approach that is not only focused on the generation of an exhaustive functional annotation, but that primarily aims at the production of highly reliable data. The results obtained on a subset of purified complexes generated through this approach allow us to suggest new standards concerning data quality in mass spectrometry analysis, and constitute a first valuable fraction of the final kinome interaction map.

COMBINATORIAL DEPLETION ANALYSIS TO ASSEMBLE THE NETWORK ARCHITECTURE OF THE SAGA AND ADA CHROMATIN REMODELING COMPLEXES

Mihaela E Sardiu, Kenneth Lee, Laurence Florens, Jerry Workman, Michael P Washburn

Stowers Institute for Medical Research, Proteomics, Kansas City, MO, 64110

At a higher level of structure, proteins within a cell interact dynamically with specific partners to form functional modules and consequently macromolecular assemblies which are responsible for particular processes in a cell. To date it remains a key challenge to discover and understand how proteins interact and are spatially arranged to form these macromolecular functional modules. The high-resolution characterization of multi-protein assemblies at physiological conditions is generally hard to achieve using traditional methods such as X-ray crystallography or NMR. However, even if it is not feasible to determine the structure of protein complexes at atomic or amino acid levels, methods predicting lower resolution macromolecular models explaining the position of proteins and their interrelationships within complexes still improve our understanding of these protein assemblies and the cellular processes they are involved in.
As a starting point for predicting a lower resolution description of protein complexes we developed a method combining computational approaches and quantitative proteomics data generated from a systematic collection of wild-type and different deletion strain purifications. In order to test this approach we generated and analyzed a dataset aiming to gain novel insights intothe Saccharomyces cerevisiae SAGA (Spt-Ada-Gcn5) histone acetyltransferase (HAT) complex. Our unique method allowed us to refine the architectural organization of both complexes by separating its distinct functional modules, which allows to explain the shared and unique functions of either complex, and also helps to characterize functional consequences of perturbations to the system. The identification of the different modules is also a crucial step for defining the direct subunit-subunits contacts within protein complexes. Furthermore, this approach also led to the identification of a novel integral subunit of the ADA complex, which we termed Ahc2, and to the characterization of Sgf29 as an ADA family protein which is present in all Gcn5 HAT complexes.
Finally, we could demonstrate that this unique approach also allows the prediction of novel functional associations within the SAGA complex and provides mechanistical insights into phenotypical observations in SAGA mutants.

ELUCIDATING COMPOUNDS' MODE OF ACTION BY MEANS OF WEIGHTED PROTEIN INTERACTIONS

Ramona Schmid*[1], Patrick Baum*[1], Carina Ittrich[2], Katrin Fundel-Clemens[1], Bärbel Lämmle[1], Fabian Birzele[1], Andreas Weith[1], Benedikt Brors[3], Roland Eils[3,4], Detlev Mennerich[5], Karsten Quast[1]

[1]Boehringer Ingelheim Pharma GmbH & Co. KG, Pulmonary Diseases Research - Genomics, Biberach/Riss, 88397, Germany, [2]Boehringer Ingelheim Pharma GmbH & Co. KG, Medical Data Services - Biostatistics, Biberach/Riss, 88397, Germany, [3]German Cancer Research Center (DKFZ), Theoretical Bioinformatics Department, Heidelberg, 69120 , Germany, [4]University of Heidelberg, Department of Bioinformatics and Functional Genomics, IPMB, BioQuant, Heidelberg, 69120 , Germany, [5]Boehringer Ingelheim Pharmaceuticals Inc., Immunology & Inflammation, Ridgefield, CT, 06877

Motivation: Clarification of compounds' mode of action is a common objective in scientific research both for academia and especially for the pharmaceutical industry. Elucidating the mode allows for detailed analyses of on as well as off target effects. It is mediated by binding of compounds to proteins subsequently influencing related regulatory networks within the organism. As more and more data from diverse sources is available, the integration of this knowledge is an important step to get a deeper insight into biology.
Results: Here, we have integrated for the first time comprehensive information about the relatedness of proteins. In order to resolve mode of action, protein interactions are taken from iRefIndex and enriched with information about transcription factor binding sites, Gene Ontology annotation as well as gene expression measurements. Finally, the accumulated information is translated to edge weighted graphs and subnetworks are extracted based on protein interactions exhibiting high contextual relatedness. Results are quantitatively evaluated by a comparison to randomized networks. Finally the biological context of the identified protein interactions is elucidated by conducting Fisher's exact test on gene sets as for example determined by Reactome.

CELL-TYPE-SPECIFIC NETWORKS: A MINING METHOD AND THEIR RELATIONS TO DISORDERS AND EVOLUTION

Jun Sese

Ochanomizu University, Dept. of Computer Science, Tokyo, 1128610, Japan

It is important to determine cell-type-specific gene expressions in order to understand cellular mechanisms. Genes expressed only in the same multiple cell types may share common functional components. Particularly in higher species, these genes are associated with specific tissues, developmentally close cells, or histologically similar cells; furthermore, defects in their associated genes may engender diseases in several related tissues and may cause serious damage during developmental processes.

To identify these biologically important genes, we developed a novel computational method that extracts functional gene networks associated with plural cell types by combining gene expressions with protein-protein interactions. We applied this method to human gene expressions observed in 79 tissues with large protein-protein interaction networks. We termed the extracted networks “tissue-sharing networks” (TSNs). Predicted TSNs reveal networks related to multiple tissues that are developmentally or histologically similar to one another, and the tissue-specific expressions of the TSN genes are conserved in mouse tissues. Many genes in TSNs are responsible for the occurrence of diseases in more than 1 tissue. However, the genes in TSNs do not overlap greatly with essential genes. Our results confirm that our method is a potential tool for the discovery of genes involved in developmental biology, pathology, and genetics.

GENE NETWORK COMPARISON BASED ON MULTIPLE TIME-COURSE MICROARRAY DATA

Teppei Shimamura, Seiya Imoto, Rui Yamaguchi, Masao Nagasaki, Satoru Miyano

University of Tokyo, Human Genome Center, Institute of Medical Science, Tokyo, 108-8639, Japan

Elucidating the differences between cellular responses to various biological conditions or external stimuli is an important challenge in systems biology. Many approaches have been developed to reverse engineer a cellular system, called gene network, from time series microarray data in order to understand a transcriptomic response under a condition of interest. Comparative topological analysis has also been applied based on the gene networks inferred independently from each of the multiple time series datasets under varying conditions to find critical differences between these networks. However, these comparisons often lead to misleading results, because each network contains considerable noise due to the limited length of the time series. We propose an integrated approach for inferring multiple gene networks from time series expression data under multiple biological conditions and present a new parameter estimation method, relevance-weighted recursive elastic net, for providing higher precision and recall than existing reverse-engineering methods. We analyze experimental data of MCF-7 human breast cancer cells stimulated by epidermal growth factor or heregulin with several doses and provide novel biological hypotheses through network comparison.

CONSTRUCTION OF A MODEL OF MITOCHONDRIAL METABOLISM AND ITS USE IN INVESTIGATING DISEASE.

Anthony C Smith, Alan J Robinson

MRC - Mitochondrial Biology Unit, Bioinformatics, Cambridge, CB2 0XY, United Kingdom

Mitochondria are a vital component of eukaryotic cells, with functions that extend beyond energy production to include metabolism, signalling, cell growth and apoptosis. Their dysfunction is implicated in a large number of metabolic, degenerative and age-related human diseases. Therefore, it is important to gain an understanding of mitochondrial metabolism as this may provide an insight into these disorders. To this end a new manually curated metabolic model of the mitochondrion was created. The framework for its construction was MitoMiner, a new mitochondrial proteome database. The model was built by evaluating the localisation of 2187 reactions that had either experimental evidence for mitochondrial localisation or were annotated as mitochondrial, using the principle of metabolite availability. The resultant model contains 242 reactions and is the most comprehensive of its type available to date. Flux balance analysis was used to simulate metabolism and showed the fluxes representing core metabolism compared well with experimentally measured figures, suggesting the model is biologically relevant. The model was then used to simulate several metabolic disorders including fumarase deficiency and pyruvate dehydrogenase deficiency. The results of these simulations corresponded well with phenotypic data from the primary literature and provided some insight into the complicated and unintuitive phenotypes of these disorders, allowing several predictions to be made.

NETWORK MODELING SUGGESTS NEW MECHANISTIC DETAILS UNDERLYING ALZHEIMER'S DISEASE

Montserrat Soler-Lopez[1], Andreas Zanzoni[1], Ricart Lluis[1], Ulrich Stelzl[2], Patrick Aloy[1,3]

[1]Institute for Research in Biomedicine, Joint IRB-BSC program in Computational Biology, Barcelona, 08028, Spain, [2]Max Plank Institute for Molecular Genetics, Otto-Warburg-Laboratories, Berlin, 14195, Germany, [3]Institució Catalana de Recerca i Estudis Avançats, ICREA, Barcelona, 08010, Spain

Alzheimer's disease (AD) is a devastating neurodegenerative disorder characterized neuropathologically by the extracellular accumulation of amyloid-β (Aβ) plaques and the intracellular accumulation of hyperphosphorylated Tau protein in the form of neurofibrillary tangles (NFTs). Although highly heritable, AD is associated with multiple genetic defects either mutational or of susceptibility, making genetic analysis difficult. This is particularly relevant for the sporadic late-onset AD, which accounts for more than 90% of AD cases.

In order to uncover novel genetic risk factors whose susceptibility implication in AD still remains unclear, we have applied an integrative approach between network modelling and an AD disease-oriented focus to generate a stringent protein interaction network for AD causative and susceptibility genes.

Our phenotype-based network revealed four novel interactions among known AD causing genes and, notably, identified 170 novel interactions with AD related genes. Bioinformatics analysis showed functional categories among AD susceptibility genes to common cellular pathways like inflammation and oxidative stress that might lead to mitochondrial dysfunction and ultimately to neuronal death. Finally, this protein interaction network provides 50 candidate genes for AD whose genetic defect has not yet been identified.

BISTABLE RLR-IFNβ SIGNALING SYSTEM DETERMINES APOPTOTIC CELL FATE

Sun-Young Hwang[1], Kye-Yeon Hur[2], Seung-Hwan Kim[2,3], Jeong-Rae Kim[4], Kwang-Hyun Cho[4], Joo-Yeon Yoo[1,2]

[1]POSTECH, Life Science, Pohang, 790-784, South Korea, [2]POSTECH, 2School of Interdisciplinary Bioscience and Bioengineering , Pohang, 790-784, South Korea, [3]POSTECH, Physics, Pohang, 790-784, South Korea, [4] KAIST, Bio and Brain Engineering, Daejeon, 305-701, South Korea

IFNs play key roles in the fate decisions of virus-infected cells. Cytosolic RLRs sense foreign RNA and activate signaling cascade to produce IFNβ, which modulates its own synthesis through multiple feedback loops. Using mathematical modeling and biochemical experimentation, we have characterized the RLR-IFNβ production signaling system. The cooperative production of RLR and the intrinsic synthesis rates of RLR are analyzed as the primary factors influencing the bistability of the system. Using IFNβp-GFP-IFNβUTR reporter cells transfected with Cy5-polyI:C, characteristics of biphasic IFNβ production have been analyzed in the mammalian cellular system. Our findings suggest that the RLR-IFNβ production system is intrinsically bistable, mainly due to genetically wired regulatory circuits and cellular feedback regulation, which determine the dynamics of IFNβ production and the fate of virus-infected cells.

A NETWORK MEDICINE APPROACH TO HUMAN DISEASE

Patrick Aloy

Institute for Research in Biomedicine, Barcelona, 08028, Spain

High-throughput interaction discovery initiatives are providing thousands of novel protein interactions which are unveiling many unexpected links between apparently unrelated biological processes. In particular, analyses of the first draft human interactomes highlight a strong association between protein network connectivity and disease. Indeed, recent exciting studies have exploited the information contained within protein networks to disclose some of the molecular mechanisms underlying complex pathological processes. These findings suggest that both protein-protein interactions and the networks themselves could emerge as a new class of targetable entities, boosting the quest for novel therapeutic strategies. In this talk, I will summarize our work towards the characterization and modelling of the protein-interaction network underlying Alzheimer´s disease, together with our most recent attempts to decipher complex cell networks to the point of being able to predict how the perturbation of a node might affect the system as a whole.

NUCLEAR RECEPTORS, TRANSCRIPTIONAL NETWORKS AND CANCER.

Kevin P White

The University of Chicago, Institute for Genomics & Systems Biology, Chicago, IL, 60637

I will present a discussion of the genetic basis for different subtypes of breast cancer. Using transcriptome and exome sequencing we have identified novel isoforms predictive of breast cancer progression, and we have determined genetic markers predictive of triple negative (basal type) vs. Estrogen Receptor positive breast cancers. These results are very promising for early identification and treatment of breast cancers that differ in their prognosis, but also suggest that exome sequencing may generally be a very powerful tool for dissecting the genetic basis of complex diseases.

NETWORK-BASED CLASSIFIERS FOR IDENTIFYING DRIVER-MUTATION-BEARING CANCER GENES

ARA CHO, JungEun Shim, Insuk Lee

Yonsei University, Department of Biotechnology, Seoul, 120-749, South Korea

It is well known that all cancers accumulate somatic mutations, and these mutated genes, called cancer genes, are the cause of cancer. There has been a lot of efforts to identify these cancer genes that are causative of cancer when mutated, and high-throughput sequencing technologies accelerate these studies by finding out the somatic mutations in tumour samples at genomic level. However, not all somatic mutations in cancer genome are involved in oncogenesis. Some of the mutations known as 'driver mutations ' confer growth advantages to the cells while the other mutations called 'passenger mutations' do not contribute cancer development. Therefore, distinguishing the genes which carry the driver mutations from the genes that have passenger mutation is very important to understand oncogenesis. Here we made a network-based classifier, to distinguish cancer genes from passenger-mutated genes by scoring somatically mutated genes in cancer genomes. The idea is that the highly ranked genes that have high scores by the classifier are more likely to have driver mutation and ,therefore, regarded as cancer genes. On the other hand, low-scored genes are not regarded as cancer genes, and the mutation they have are expected to be a passenger mutation. We made a 4 kinds of Network-based classifiers by using Human Network and Worm(*C.elegans*) Network : Degree centrality, GBA(Guilt by association), geneMANIA algorithm, and integrated version of these 3 classifier. Also, we collect 2972 somatically mutated genes from unbiased sequencing data of breast, colorectal, and pancreatic cancer and glioblastoma multiforme(GBM), and made a candidate cancer gene list. We evaluated those classifier by training set(384 cancer genes) from Cancer Gene Census by Welcome Trust Sanger Institute. We then map those candidate genes to *C.elegans* for further biological validation.
The networks that we used are functional networks, which include many kinds of information such as PPI, co-expression, and many other information. Therefore, we expect that the classifier based on the these functional networks are reliable, and suggest new approach to distinguish driver mutations from passenger mutations.

INTEGRATIVE SYSTEMS BIOLOGY ANALYSIS AND VISUALIZATION OF RELEVANT SUBNETWORK MODULES AND PATHWAYS IN PHENOTYPIC TRAITS AND DISEASES

Tanja Kunej[1], Peter Dovc[1], Jernej Ogorevc[1], Eva Ceh[1], Minja Zorc[2], Irena Godnic[1], Dasa Jevsinek Skok[1], George A Calin[3], Jana Ferdin[1,3], Crt Gorup[4], Carlo V Cannistraci[5,6]

[1]University of Ljubljana, Biotechnical Faculty, Department of Animal Science, Domzale, 1230, Slovenia, [2]UL, Medical Faculty, IBMI, Ljubljana, 1000, Slovenia, [3]University of Texas, MDACC, Houston, TX, 77030, [4]UL, Faculty of Computer and Information Science, Ljubljana, 1000, Slovenia, [5]San Raffaele Scientific Institute, Proteome Biochemistry, Milano, 20132, Italy, [6]KAUST, Chemical and Life Sciences and Engineering Division, Thuwal, 23955-6900, Saudi Arabia

The use of high throughput technology in molecular biology is rapidly increasing the amount and the number of different types of 'omics' data. The fast accumulation of diversified biological information generates fragmented, and sometimes contradictory, scenery which emerges from cross-comparison of data from different sources (publications and databases). Therefore, systematic assembling, organization and integration of this information is an urgent need in computational biology and will allow complementation of different pieces of evidence based on a holistic-network-driven approach. The single-gene-biomarker paradigm is now replaced by a network-based approach to biomarker discovery and disease prediction, which merges the omics perspective at different levels: from single-nucleotide polymorphism (SNP) (1) to regulatory networks (2).
We tested applicability and robustness of the systems biology approach, in search for relevant subnetwok modules and pathway-based biomarkers for diverse group of phenotypic traits and diseases including: mammary gland/milk traits, mastitis, fat deposition, male reproduction, cryptorchidism and chronic lymphocytic leukemia. The following procedure was employed:
1) extracting diverse omics data from the literature and databases,
2) data organization, identification and visualization of overlapping genetic loci,
3) systems biology analysis including visualization of relevant microRNA-mRNA interactions, protein-protein interactions and pathways,
4) development of network/pathway based biomarkers,
5) experimental confirmation.
We are refining a robust protocol for pathway based biomarker development which could be used for other fields of research.

1. Hannum G, Srivas R, Guenole A, van Attikum H, Krogan NJ, Karp RM, Ideker T. Genome-wide association data reveal a global map of genetic interactions among protein complexes. PLoS Genetics 2009;5:e1000782.

2. Ravasi T, Cannistraci CV, Katayama S, Bajic VB, Tan K, Akalin A, Schmeier S, et al. An atlas of combinatorial transcriptional regulation in mouse and man. Cell 2010;140:744-752.

MOLECULAR NETWORKS IN LEUKEMIA AND INNATE IMMUNITY

Giulio Superti-Furga, Florian Grebien, Jacques Colinge, Oliver Hantschel, Tilmann Bürckstümmer, Roberto Sacco, Evren Karayel, Adriana Goncalves, Christoph Baumann, Andreas Pichlmair, Alexei Stukalov, Andre Müller, Keiryn L Bennett

CeMM, Center for Molecular Medicine of the Austrian Academy of Sciences, Vienna, 1090, Austria

Physiology relies on the concerted action of a number of molecular interactions of gene products and metabolites operationally organized in so-called pathways and in yet larger molecular networks. Through integrated approaches using proteomics as central “glue” it is possible to obtain physical, functional and “knowledge” maps of human pathways. We use affinity proteomics using TAP-mass spectrometry to chart protein complexes and larger molecular networks and use drugs to establish target profiles. The combination of these approaches allows not only to map drug/ligand- target relationships but also to position these elements onto molecular pathways. We mapped pathways involved in innate immunity and in chronic myelogenous leukemia. We also compared five drugs against CML and investigated the Bcr-Abl protein network. The data suggest: 1. It is possible to map entire human signaling pathways using mass spectrometry-based proteomics, 2. Even “modern” targeted drugs are promiscuous, 3. Drug targets may generally be part of larger protein complexes, 4. Binding partners of targets may influence drug action on the complex and in turn be affected by it, 5. In light of cellular “proteostasis”, complex components may become available after drug treatment and redistribute to affect other signaling pathways, 6. Drugs are likely to be understood as systems perturbers and not “erasers” of protein activity (with some gain-of-function effects matching loss-of-function). A better understanding of these perturbations should form the basis for an informed combination-type of therapy.

MOLECULAR MECHANISMS IN SCHIZOPHRENIA UNCOVERED WITH SYSTEMS GENETICS

Jacob J Michaelson, Andreas Beyer

Cellular Networks and Systems Biology, Biotechnology Center, TU Dresden, Dresden, 01307, Germany

Schizophrenia is a mental disorder with a significant impact on both the personal and societal level. It is highly heritable, and approximately 1% of the population manifests the disease. In spite of its prevalence and strong genetic component, the etiology of schizophrenia remains elusive because of its nature as a complex trait. In this work, we used a variety of data sources and novel computational techniques to investigate whether disparate risk factors and associated genes are connected by a common molecular mechanism. As a first step in this analysis, we integrated gene expression data sets derived from brain tissue of several cohorts and used Random Forests to classify subjects as schizophrenic or control, based on gene expression. We used the variable importance measures resulting from this classification procedure to learn which combinations of genes were crucial for distinguishing schizophrenia patients and controls. These importance measures were combined with text mining evidence to build a set of high-confidence schizophrenia-associated genes. To further explore the regulatory context of these genes, we exploited systems genetics data from a recombinant inbred panel of mice that vary widely in their neurological phenotypes. Specifically, we examined three types of regulatory relationships to give context to these risk genes: upstream regulators, downstream targets, and lateral (epistatic) co-regulators. Examining these complex relationships was made possible by a novel extension of Random Forests that utilizes latent information in the forest structure. The insight gained through this work demonstrates the value of a systems approach when deconstructing the etiology of schizophrenia, and the presented framework also has wider applicability to complex diseases in general.

PROTEOMICS AND SYSTEMS ANALYSIS OF ESTROGEN-INDUCED CELL GROWTH OR CELL DEATH IN BREAST CANCER CELLS

Zhang-Zhi Hu, Benjamin L Kagan, Lihua Zhang, V. Craig Jordan, Anna T Riegel, Anton Wellstein

Georgetown University, Lombardi Cancer Center, Washington, DC, 20057

Estrogen is known to induce growth in estrogen receptor (ER)-positive breast cancer cells. Paradoxically, in breast cancer cells that have been chronically deprived of estrogen stimulation, re-introduction of the hormone can induce apoptosis. Here, we sought to identify signaling networks that are triggered by estradiol (E2) in isogenic MCF-7 breast cancer cells that undergo apoptosis (MCF-7:5C) versus cells that proliferate upon exposure to E2 (MCF-7). One focus was on proteins interacting with the nuclear receptor co-activator AIB1 (Amplified in Breast Cancer-1) the other on complexes containg tyrosine phosphorylated proteins. AIB1 is known to be rate-limiting for E2-induced cell survival responses in MCF-7 cells and was found here to also be required for the induction of apoptosis by E2 in the isogenic MCF-7:5C cells. Proteins that interact with AIB1 as well as complexes that contain tyrosine phosphorylated proteins were isolated by immunoprecipitation and identified by mass spectrometry (MS) at baseline and after a two hour exposure to E2. Bioinformatic network analyses of the identified protein interactions were then used to analyze E2 signaling pathways that trigger apoptosis versus survival. Comparison of MS data with a computationally-predicted AIB1 interaction network showed that 22 proteins identified in this study are within this network, and are involved in signal transduction, transcription, cell cycle regulation and protein degradation. G-protein-coupled receptors, PI3 kinase, Wnt and Notch signaling pathways were most strongly associated with E2-induced proliferation or apoptosis and are integrated here into a global AIB1 signaling network that controls qualitatively distinct responses to estrogen.

REWIRING OF INFLAMMATION AND TUMOR SUSCEPTIBILITY NETWORKS IN SKIN CANCER

David A Quigley[1], Minh D To[1,2], Il-Jin Kim[1], Donna G Albertson[1,3], Jonas Sjolund[1], Allan Balmain[1]

[1]University of California San Francisco, Hellen Diller Family Comprehensive Cancer Center, San Francisco, CA, 93143, [2]University of California San Francisco, Thoracic Oncology Program, San Francisco, CA, 93143, [3]University of California San Francisco, Department of Laboratory Medicine, San Francisco, CA, 94143

Germline polymorphisms influence gene expression networks in normal mammalian tissues. Common genetic variants have also been shown to affect many complex traits including cancer susceptibility. However, factors responsible for most of the expected heritable risk of cancer development have not yet been identified. Finding these alleles and isolating the causal polymorphisms is challenging because the heritable component of susceptibility is influenced by many alleles exerting modest effects which may be pleiotropic, epistatic, or context-dependent. We have previously identified alleles affecting the expression of networks of genes which influence complex traits including inflammation and cancer susceptibility. Analysis of the genetic architecture of gene expression in normal skin from a Mus spretus / Mus musculus backcross identified expression Quantitative Trait Loci (eQTL) that influence both structural and functional phenotypes including hair follicle development, inflammation and tumor susceptibility. Changes in germline influence on gene expression during the development of benign and malignant tumors have not been investigated systemically. We now document major changes in germline control of gene expression during skin tumor development, likely reflecting cell selection, somatic genetic events, and changes in tumor microenvironment. Numerous genes, including the immune response genes Interleukin 18 and Granzyme E, are under germline control in malignant tumors but not normal skin. Gene expression networks linked to tumor susceptibility and hair follicle stem cell markers in normal skin undergo significant reorganization during tumor progression. Our data highlight opposing roles of Interleukin-1 signaling networks in tumor susceptibility and tumor progression and have implications for the development of chemopreventive strategies to reduce cancer incidence.

GENE REGULATORY NETWORKS IN SYSTEMS PHYSIOLOGY

A.J. Marian Walhout

University of Massachusetts, Program in Gene Function and Expression, Worcester, MA, 01605

Gene regulatory networks (GRNs) provide insights into the mechanisms of differential gene expression at a systems level. GRNs that relate to metazoan development have been studied extensively. However, little is still known about the design principles, organization and functionality of GRNs that control physiological processes such as metabolism, homeostasis and responses to environmental cues. I will present the first experimentally mapped metazoan GRN of Caenorhabditis elegans metabolic genes. This network is enriched for nuclear hormone receptors (NHRs). The NHR family has greatly expanded in nematodes: humans have 48 NHRs, but C. elegans has 284, most of which are uncharacterized. We find that the C. elegans metabolic GRN is highly modular and that two GRN modules predominantly consist of NHRs.
Network modularity has been proposed to facilitate a rapid response to different cues. As NHRs are metabolic sensors that are poised to respond to ligands, this suggests that C. elegans GRNs evolved to enable rapid and adaptive responses to different cues by a concurrence of NHR family expansion and modular GRN wiring.

ROBUST PATTERNING IN ARABIDOPSIS FLOWERS: MADS SCIENCE

Simon van Mourik, Richard Immink, Susan Urbanus, Jaap Molenaar

Plant Research International, Plant Sciences Group, Droevendaalsesteeg, Wageningen, 6700AC, Netherlands

The Arabidopsis Thaliana flower consists of four types of organs, that grow in four concentric whorls on the floral meristem. In each whorl, the identity of the cells is determined by the concentrations of five types of MADS proteins.
In this presentation we investigate the sensitivity and robustness of a model describing the dynamics of the gene regulatory network that is responsible for flower organ identity [1]. This recently developed model incorporates transcription regulation via Michaelis-Menten kinetics, decay, dimer interactions. It also includes so-called trigger mechanisms, short boosts of protein production, that turn out to ultimately lead to four types of cell differentiation. This model generates realistic gene expression patterns in the different whorls, and has been validated by successful predictions of mutant phenotypes.
In this model no cell-cell interactions are yet included, such as protein transport between cells. In [2] is was shown that the MADS proteins can easily travel through cell membranes, giving rise to a large diffusive effect. This diffusion was not yet included in the model in [1]. This raises the question how robust this model is against inclusion of diffusive effects. For example, how are the four steady states of the model, each corresponding to one of the whorls, affected by an extension of the model with diffusion. In this talk the biological and mathematical implications of this question will be treated.

[1] Continuous-time modeling of cell fate determination in Arabidopsis flowers. Simon van Mourik, Aalt-Jan van Dijk, Maarten de Gee, Richard Immink, Kerstin Kaufmann, Gerco Angenent, Roeland van Ham, and Jaap Molenaar. Submitted to BMC Systems Biology, 2010.
[2] Intercellular transport of epidermis-expressed MADS domain transcription factors and their effect on plant morphology and floral transition. Susan L. Urbanus, Adriana P. Martinelli, Q.D. (Peter) Dinh, Lilian C.B. Aizza, Marcelo, C. Dornelas, Gerco C. Angenent, and Richard G.H. Immink. Accepted by Plant Journal, 2010.

A TAXONOMY OF NETWORKS

JP Onnela[1], D Fenn[2], S Reid[1], P Mucha[3], M Porter[2], N Jones[1]

[1]Oxford, Physics, Oxford, OX1 3PU, United Kingdom, [2]Oxford, Maths, Oxford, Ox1 3PU, United Kingdom, [3]UNC, Maths, Chapel Hill, NC, 27599-3250

When is a biological network functionally typical or atypical, pathological or healthy? To answer this we may need to ask: how should we compare biological networks? Our means of comparison should indicate when networks with similar functions are indeed similar. Unfortunately, a good understanding of network function requires knowledge of the processes on the network, but this is something we often lack. Some good news is that despite this problem, some structural features of networks, the modular or community structure of the network, can put a strong constraint on many plausible types of network dynamics. As a result there might indeed be a connection between network structure and system function.

We create a characteristic signature of the modular structure of networks by probing them at multiple resolutions. We use this signature to cluster a set of networks from across the sciences and find that we group networks of similar functions together. This allows us to consider a taxonomy of networks. We also consider a case study of a comparison of the structure of a large set of fungal networks and also find structural similarities between some protein interaction data and networks of scientific collaborations.

This method can be used to detect anomalous networks and identify connections between apparently disparate types of systems.

PROTEIN INTERACTION NETWORKS: THE EDGE OF REGULATION

Guilhem Chalancon[1,2], M. Madan Babu[1]

[1]MRC Laboratory of Molecular Biology, Structural Studies, Cambridge, CB10QH, United Kingdom, [2]Ecole Normale Superieure de Cachan, Biology department, Cachan, 94230, France

Background
Molecular crowding and proteome diversity permanently require cells to adjust the quantities of the proteins they produce. Arguably, protein abundance must be balanced in an complex manner to enhance functional protein-protein interactions (PPI) while limiting non-specific associations. Despite this constraint, cells have strikingly evolved more than six orders of magnitude in the binding affinities of their PPI, from transient to stable interactions. How is this possible? Addressing this problem requires to understand how a cell regulates the flux of synthesis and degradation of proteins involved either in transient or stable interactions. We hypothesized that distinct strategies exist to regulate stable or transient proteins interactions.

Results
We integrated multiple large scale datasets describing most of the regulatory steps influencing protein abundance, and used statistical network analysis to investigate the interplay between regulatory properties and topologies in 3 S.cerevisiae PPI networks. We present evidence for distinct constraints in the regulation of stable and transient PPI, in the form of significantly more/less enhanced use of a particular regulatory step during protein synthesis. In particular, distinct constraints are found at the transcriptional, post-transcriptional and post-translational levels, as well as for protein noise (single-cell variations of protein abundance).

Conclusion
We provide a comprehensive analysis of protein abundance regulation in three distinct protein interaction networks. We show that stable and transient PPI are under distinct patterns of regulation. This likely provides several degrees of freedom for the cell to dynamically fine-tune abundance of proteins involved in stable or transient interactions. Our findings have implications for a system-level understanding of the dynamics of protein interactions and for synthetic biology experiments aimed at engineering functional interactions.

SINGLETON AND DUPLICATED HUBS IN THE EVOLUTION OF PROTEIN INTERACTION NETWORKS

Matteo D'Antonio, Francesca D Ciccarelli

European Institute of Oncology (IEO), Experimental Oncology, Milan, 20139, Italy

In several organisms from prokaryotes to yeast and fly, modifications in the dosage of essential genes, genes coding for protein complexes, and network hubs are deleterious and often lethal. This has been interpreted as a sign of fragility of these genes towards perturbations, because dosage modifications of essential and highly connected proteins may cause the simultaneous disruption of tightly regulated biological processes. Recently, our and other labs have shown that mammalian networks behave differently from other species. In mouse and human, the majority of genes coding for hubs have paralogs and essentiality is not always anticorrelated to duplicability. In this context, genes that are somatically modified in cancer (i.e. cancer genes) constitute an exception: they are mostly singletons but they encode proteins that are highly connected. This finding suggests that cancer genes constitute a point of fragility in the human interactome and that they have evolved differently from other singleton genes, gaining peculiar genetic and network properties.
Here we present the newest results of a study that aimed at clarifying when the differences between the interaction networks of mammals and other species arose during evolution and how they affected the evolution of cancer genes. We specifically focused on four species (human, fly, yeast, and E.coli) that represent major transitions towards the progressive increase in complexity during evolution and for which good quality genomic and primary interaction data are available. In all these species we compared network properties of each node with conservation and duplicability of the corresponding gene. We discovered that in all species, including human, very ancient and conserved genes encode singleton and central hubs. These genes usually accomplish very basic functions that deal with the survival of the cell. A second group of hubs can be found only in human because they appeared later in evolution and duplicated through the two rounds of vertebrate-specific whole genome duplication. These recent hubs are involved in more complex functions, such as signaling, cell-to-cell communications, and regulation of biological processes. We also discovered that the dosage of these genes is regulated through alternative and more sophisticated mechanisms. Interestingly, cancer proteins are mostly ancient and they therefore result enriched in singleton hubs.

IMMUNE SIGNALING LANDSCAPE: INTERROGATING SINGLE CELLS IN 30 DIMENSIONS

Rachel Melamed[1,2], Elad David-Amir[1,2], Sean C Bendall[3], Erin Simonds[3], Karen Sachs[3], Garry Nolan[3], Dana Pe'er[1,2]

[1]Columbia University, Biological Sciences, New York , NY, 10027, [2]Columbia University, C2B2, New York, NY, 10032, [3]Stanford University, Department of Microbiology and Immunology, Palo Alto, CA, 94305

We employ mass cytometry, a recently developed technology, to measure the abundance of over thirty surface and signaling proteins at single cells resolution in complete bone marrow samples. High dimensional clustering was used to identify immune cell types. Clustering finds many of the known cell types, but also finds a more complex landscape of cell types than previously appreciated. We mapped signaling events onto this immune landscape gaining a comparative and comprehensive view signaling networks in the immune system.

WIDESPREAD INDUCTION OF GENETIC NETWORKS IN RESPONSE TO DNA DAMAGE

Sourav Bandyopadhyay[1], Monika Mehta[2], Ruedi Aebersold[3], Michael-Christopher Keogh[2], Nevan J Krogan[4], Trey Ideker[1]

[1]University of California, San Diego, Medicine, La Jolla, CA, 92093, [2]Albert Einstein College of Medicine, Cell Biology, Bronx, NY, 10461, [3]ETH Zurich, Institute of Molecular Systems Biology, Zurich, 8093, Switzerland, [4] University of California, San Francisco, Cellular and Molecular Pharmacology, San Francisco, CA, 94158

Genetic networks are central to pathway structure and function but the effects of different environmental conditions on these networks and the pathways they represent are poorly understood on a global scale. We report the generation of genetic interaction networks among all yeast signaling and transcriptional machinery before and after exposure to DNA damage. This perturbation induces a distinct genetic network that is greater in size than the untreated network and whose interactions are largely DNA-damage specific. In addition, the inducible network has guided the discovery of new DNA damage response pathways involving the phosphatase Pph3 and histone variant Htz1. Thus, inducible genetic networks are a valuable tool for mapping pathways stimulated by specific conditions and probing an uncharted space of the genetic interactome.

PEPTIDE-MEDIATED INTERACTIONS IN HIGH-RESOLUTION 3D STRUCTURES

Amelie Stein[1,2], Patrick Aloy[1,2,3]

[1]Institute for Research in Biomedicine (IRB) Barcelona, Structural and Computational Biology, Barcelona, 08028, Spain, [2]Barcelona Supercomputing Center (BSC), Life Sciences, Barcelona, 08034, Spain, [3]Institució Catalana de Recerca i Estudis Avançats (ICREA), Life & Medical Sciences , Barcelona, 08010, Spain

Many biological responses to intra- and extracellular stimuli are regulated through complex networks of transient protein interactions where a globular domain in one protein recognizes a linear peptide from another, creating a relatively small contact interface. These peptide stretches are often found in unstructured regions of proteins, and contain a consensus motif complementary to the interaction surface displayed by their binding partners. While most current methods for the *de novo* discovery of such motifs exploit their tendency to occur in disordered regions, our work here focuses on another observation: upon binding to their partner domain, motifs adopt a well-defined structure. Indeed, through the analysis of all peptide-mediated interactions of known high-resolution three-dimensional (3D) structure, we found that the structure of the peptide may be as characteristic as the consensus motif, and help identify target peptides even though they do not match the established patterns. Our analyses of the structural features of known motifs reveal that they tend to have a particular stretched and elongated structure, unlike most other peptides of the same length. Accordingly, we have implemented a strategy based on a Support Vector Machine that uses this features, along with other structure-encoded information about binding interfaces, to search the set of protein interactions of known 3D structure and to identify unnoticed peptide-mediated interactions among them. We have also derived consensus patterns for these interactions, whenever enough information was available, and compared our results with established linear motif patterns and their binding domains. Finally, to cross-validate our identification strategy, we scanned interactome networks from four model organisms with our newly derived patterns to see if any of them occurred more often than expected. Indeed, we found significant over-representations for 64 domain-motif interactions, 46 of which had not been described before, involving over 6,000 interactions in total for which we could suggest the molecular details determining the binding.

COMPARATIVE LOGICAL MODELS OF SIGNALING NETWORKS IN NORMAL AND TRANSFORMED HEPATOCYTES

Julio Saez Rodriguez[1,2], Leonidas Alexopoulos[1,2], Melody K Morris[2], Dan Kirouac[2], Jonathan Epperlein[1], Regina Samaga[3], Steffen Klamt[3], Douglas A Lauffenburger[2], Peter K Sorger[1,2]

[1]Harvard Medical School, Systems Biology, Boston, MA, 02115, [2]M.I.T., Biological Engineering, Cambridge, MA, 02139, [3]Max-Planck-Institute, Systems Biology, Magdeburg, 33106, Germany

Protein interaction networks (PINs or interactomes), protein signaling networks (PSNs) and gene regulatory networks have successfully been used to classify drug-target interactions, identify master transcriptional regulators and uncover new disease genes. Unfortunately, PINs and PSNs are rarely cell-type specific and do not encode the input-output relationships required for analyzing receptor-mediated signaling cascades and the drugs that target them. Conversely, traditional approaches to studying cell signaling do not make use of the wealth of information that is now encoded in PINs and PSNs. Here we describe a hybrid method to convert PSNs into logical models (either discrete Boolean models or quantitative Fuzzy-logic models) that can be trained against data in which cells are exposed to-combinations of ligands and drugs followed by multiplex biochemical measurement of intracellular responses, and its implementation it in the toolboxes CellNetOptimizer (Mol. Sys. Biol., 5:331, 2009) and DataRail (Bioinformatics, 15(24):840, 2008). Our approach can accommodate network information from different sources (manual curation, text mining, and high-throughput experiments). We apply the method to distinguishing the topologies of immediate early signaling networks in primary human hepatocytes and four hepatocellular carcinoma (HCC) cell lines. We show that five distinct models cluster topologically into normal and diseased sets, revealing functional differences between normal and diseased cells that involve activation of growth factor receptors and intracellular kinase cascades. In a proof-of-principle experiment we also infer a target for an I-kappa B kinase inhibitor developed to treat arthritis and airway inflammation.

A SYSTEMS BIOLOGY APPROACH TO LIMB BUD DEVELOPMENT: TOOLS AND MODULES

Rolf Zeller

University of Basel, Biomedicine, Basel, CH-4058, Switzerland

My group studies the development of mouse limb buds as a paradigm for vertebrate organogenesis. In particular, we are interested in identifying and understanding the signalling systems that control the coordinate proliferative expansion and specification of the mesenchymal progenitors destined to form the chondrogenic elements of the future limb skeleton. We have recently uncovered a robust, but highly dynamic system of interlinked signalling feedback loops that controls progression of limb bud development in a largely self-regulatory manner. The challenge is now to integrate both spatial and temporal information into our mathematical simulations starting from quantitative and functional genetic data. To this aim we are developing novel genetic tools that aid our *in vivo* quantitative analysis of the molecular systems that control initiation and progression of mouse limb bud development. We have already identified several signalling modules that are molecularly interlinked such that limb bud axes development is temporally and spatially coordinated. I will present our ongoing functional analysis of the dynamic functions of two main gatekeepers in the BMP and SHH signalling pathways, the Smad4 and Gli3 transcriptional regulators. Finally , I will discuss the importance of combining genetic and quantitative analysis with data-driven mathematical simulations to achieve a comprehensive understanding of the complex morpho-regulatory interactions (in collaboration with Prof. D. Iber and Dr. A. Zuniga).

SYSTEM WIDE ANALYSIS OF THE HUMAN UBIQUITIN E3-RING INTERACTION NETWORK

Jonathan D Woodsmith, Chris M Sanderson

University of Liverpool, Physiology, Liverpool, L69 3BX, United Kingdom

Really Interesting New Genes (RINGs) comprise over half of the predicted E3 ligases encoded within the human genome. As such they play a large part in controlling the specificity and functionality within the ubiquitin system. Whilst catalytically inert, they act as molecular scaffolds which bring specific E2 conjugating enzymes and substrate into close proximity to allow ubiquitination to occur. However, RING proteins have also been shown to homo- and hetero- dimerise, with some family members forming higher order multimeric complexes. Several studies have investigated the functional implications of selective oligomerisation events however the full extent or importance of RING-RING interactions has not been systematically studied.

In this study two differing vectors systems were utilised in high throughput yeast-two-hybrid screens to investigate binary protein-protein interactions of a subset of the human RING proteins. In both screens against E2 conjugating enzymes and E3 RING proteins different vector systems provided complimentary datasets which increase coverage of binary interactions within the ubiquitin network.

Site directed mutagenesis and in vitro ubiquitination assays were used to verify E2/E3-RING interactions, while co-expression and co-precipitation studies were performed to investigate putative 3-RING/E3-RING interactions and provided new insights into the extent, specificity and functional consequences of E3-RING oligomerisation networks.

Interactions between 125 soluble E3-RING proteins were tested revealing 228 high confidence binary interactions. Of the interactions re-tested in secondary assays > 71% of positive and 100% of negative interactions were reconfirmed indicating a high degree of confidence in the Y2H derived network.

Data from these studies were combined with known interaction data to generate a human E3 RING network containing 4345 binary interactions. The canonical E1/E2/E3-RING enzymatic cascade has been investigated heavily and the combinatorial complexity of E2/E3-RING interactions within this system has been reported. However, the frequency and consequences of E3-RING oligomerisation remain less well documented. Here we present a network of binary E3-RING/E3-RING interactions within the context of the human ubiquitome. This data suggests that E3-RING/E3-RING interactions may confer a further level of fidelity within the ubiquitination cascade. In particular, E3-RING binding preferences may control the specificity and architecture of substrate ubiquitination through the selective recruitment of different E2-conjugating enzymes.

FUNCTIONAL GENOMICS STRATEGIES TO UNRAVEL PHOSPHORYLATION NETWORKS THAT CONTROL THE DNA DAMAGE STRESS RESPONSE IN ES CELLS*

J. Carreras Puigvert[1], L. von Stechow[1], R. Siddappa[1], A. Pines[2], L. H.F. Mullenders[2], B. van de Water[1], E. H.J. Danen[1]

[1]Leiden University,LACDR, Division of Toxicology, Leiden, 2333CC, Netherlands, [2]Leiden University Medical Center, Department of Toxicogenetics, Leiden, 2333ZC , Netherlands

Exposure of cells to DNA-damaging agents activates signal transduction cascades that affect DNA repair, cell cycle progression, and cell survival. Different cellular strategies exist to cope with DNA damage: damage is tolerated (with the risk that cells acquire malignant properties), damage is repaired, or cells with damaged DNA are removed from the tissue. The latter strategy involves a complex of signal transduction pathways that can induce various types of cell death, including apoptosis. In a highly simplified model DNA-damage causes activation of the sensor kinases (ATM, ATR, DNA-PK) that activate p53 and the checkpoint kinases (Chk1, Chk2), which in turn regulate effector pathways that control repair, cell cycle and cell survival. In reality, the process is much more complex with many additional regulators, extensive cross-talk, and multiple positive and negative feedback loops. Posttranslational modifications play a crucial role in the transmission of the signal during the DNA-damage response. In this project, we make use of gene family short interference (si)RNA libraries to identify kinases and phosphatases that are involved in the apoptotic response to various DNA-damaging drugs in ES cells. High throughput screening for cell viability has identified potential regulators of the sensitivity of ES cells to the DNA damaging drug Cisplatin. In addition, automated high content live cell imaging analysis is used to qualitatively and quantitatively follow the onset of apoptotic cell death, providing kinetic information of the process. From ~900 genes screened, 86 have been identified as potential hits, and have entered the validation process. In parallel, phosphoproteomics (SILAC) data have been obtained for exposure of ES cells to Cisplatin. Bioinformatics analysis of these two datasets is used to link our hits to posttranscriptional modifications of their substrates. We aim to build functional signaling networks involved in the response to Cisplatin-induced DNA damage in ES cells. These will be tested for multiple compounds and cell types, including cancer cells. In summary, this strategy will allow us to define potential biomarkers and new drugable targets in order to improve chemotherapeutic treatments.

*Supported by the Netherlands Toxicogenomics Center.

EVOLUTION OF PHOSPHOREGULATION: FROM INTERACTIONS TO FUNCTION

Pedro Beltrao[1], Jonathan C Trinidad[2], Wendell A Lim[1], Alma L Burlingame[2], Nevan J Krogan[1]

[1]University of San Francisco, Department of Cellular and Molecular Pharmacology, San Francisco, CA, 94158, [2]University of San Francisco, Department of Pharmaceutical Chemistry, San Francisco, CA, 94158

Protein phosphorylation is one of the most well characterized post-translational modifications. Yet, very little is known about the evolutionary dynamics of kinase-substrate interactions and its functional consequences. We have used a mass-spectrometry approach to characterize the in-vivo phosphoproteomes of three fungal species (S. cerevisiae, S. pombe, C. albicans) and have combined this data with cross-species genetic interaction data for comparative studies. We observed that kinase-substrate interactions change at a fast rate although the average level of phosphorylation of functional groups (i.e. complexes, signaling pathways) is well conserved. In order to address this apparent contradiction we have developed methods to predict the function of known phosphorylation sites using structural information as well as data on other post-translational modifications. This analysis indicates that the positional conservation of phosphorylation sites significantly under-predicts the conservation of function. The observation that phosphorylation sites can change during evolution without affecting the function suggest that neutral evolution plays a significant role in shaping kinase-substrate networks.

FUNCTION DICTATES TOPOLOGY IN BIOCHEMICAL NETWORKS

Chao Tang

University of California San Francisco, Bioengineering and Therapeutic Sciences, San Francisco, CA, 94158

A diverse set of biological functions in a diverse set of cells and organisms are performed by a myriad of biochemical networks. A major challenge in systems biology is to seek design principles or general rules that link network function and topology. I will present two case studies in which the network function limits the choice of the network topology to a very small set. One case concerns with the biochemical adaptation. We show that there are only two core topologies that can achieve perfect adaptation. The other case is the parasegments boundary formation in fruit fly development. The network topology for this patterning function has to be one of the few variants of a core topology. A function-topology map would help us comprehend natural biochemical networks, identifying key or missing nodes in the network and suggesting the network's function. It can also provide us with a manual to synthesize artificial networks.

EXTENDING THE STRUCTURAL RECONSTRUCTION OF THE METABOLIC NETWORK OF *T. MARITMA*.

Ying Zhang, Andrei Osterman, Adam Godzik

Sanford-Burnham Medical Research Institute, Bioinformatics and Systems Biology, La Jolla, CA, 92037

Sequencing of a genome provides a foundation for the reconstruction and modeling of cell- and organism-level networks, which allow us to understand and eventually simulate entire cells and organisms. Knowledge of three-dimensional structures of proteins that compose such networks extends our understanding of the underlying processes to the atomic level. We developed a reconstruction of a central metabolic network for a hyperthermophilic bacterium, *Thermotoga maritima*, and followed it with experimental and computational determination of three dimensional structures of all proteins in the reconstruction. Combination of these two perspectives gave us the ability to analyze the network in novel ways. However, our reconstruction covered only the central metabolism i.e. less than 25% of the entire genome. Now, we attempt to extend the reconstruction to the entire *T. maritima* genome and in particular to include enzymes and probable enzymes that are specific to *T. maritima*.

INFORMATION PROCESSING BY BIOCHEMICAL NETWORKS: A DYNAMIC APPROACH

Clive Bowsher

University of Cambridge, Centre for Mathematical Sciences, Cambridge, CB3 0WB, United Kingdom

Background: Understanding how information is encoded and transferred by biochemical networks is of fundamental importance in cellular biology. This requires analysis of the relationships between the stochastic trajectories of the constituent molecular species of the network. We will discuss how to identify conditional independences between the trajectories or time courses of groups of species. These are robust network properties that provide important insight into how information is processed.

Objectives: The problem of how to decompose an entire network exactly into modules on informational grounds is considered: all information relevant to a particular module's trajectory should be encoded by those species in the region where the module overlaps with the rest of the network. In the context of signalling networks, the aim is to identify the route and species involved in sequential information processing between input and output modules.

Methods: The MIDIA algorithm is presented which allows automated identification of exact decompositions for large networks and visualisation using a tree structure that encodes the conditional independences. The algorithm relies on construction of the Kinetic Independence Graph for the network and uses only stoichiometric information. No knowledge of rate parameters or use of simulation is required.

Results: A bespoke version of the algorithm designed for signalling networks identifies the process of sequential encoding and visualises its structure using the collection of input-output paths in the tree. Together with the input-output path matrix introduced, this provides the first powerful, visual tool of its kind for investigating information processing by signalling networks. Application of the techniques to the Toll-like Receptor signalling network reveals that inputs can be informative for signalling outputs in ways unanticipated by steady-state analyses, that the information processing structure is not well described as a bow-tie, and that encoding for the interferon response is unusually sparse compared to other outputs of this innate immune system.

Conclusion: The dynamic conditional independence-based techniques introduced require only stoichiometric information but have the strength of fully incorporating both stochastic and non-steady state dynamics. We believe they constitute an important step in the development of a quantitative theory of cellular information processing.

DEAF PLANTS, BLEEDING YEAST, AND OTHER SURPRISING DISEASE MODELS FROM DEEPLY CONSERVED GENE NETWORKS

Edward M Marcotte[1], Kriston L McGary[1], Tae Joo Park[1,2], Hye Ji Cha[1], John O Woods[1], Martin Blom[1], John B Wallingford[1,2,3]

[1]University of Texas, Center for Systems and Synthetic Biology, Austin, TX, 78712, [2]University of Texas, Dept. of Molecular, Cell, and Developmental Biology, Austin, TX, 78712, [3]Howard Hughes Medical Institute, Chevy Chase, MD, 20815

Systems biology has shown great promise in providing a better understanding of human disease, and in identifying new disease targets. Nonetheless, it remains extraordinarily difficult to identify causal genes in most genetic diseases, in particular highly polygenic disorders, such as, for example, coronary artery disease, diabetes, and autism, for which current approaches are most limited. I'll discuss recent progress on this problem, including an unusual approach for identifying candidate genes for polygenic diseases, based on mapping mutational phenotypes between organisms as distant as humans and plants or yeast, revealing new models of disease and candidate disease genes (the "phenolog" hypothesis). This method suggests a yeast model for angiogenesis defects, worm models for breast cancer and neural tube birth defects, and a plant model for the neural crest defects associated with Waardenburg syndrome, among others, and reveals functionally coherent, evolutionarily conserved gene networks—many predating the plant-animal divergence—capable of identifying candidate disease genes.

Participant List

Dr. Mario Albrecht
Max Planck Institute for Informatics
mario.albrecht@mpi-inf.mpg.de

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

Dr. Mohammed Alshalalfa
University of Calgary
msalshal@ucalgary.ca

Mr. Mochamad Apri
Wageningen University
mochamad.apri@wur.nl

Dr. Sabine Asser-Kaiser
German Cancer Research Center
s.asser-kaiser@dkfz.de

Dr. Charles Bailey
Global Institute for Scientific Thinking
cbailey1@cfl.rr.com

Dr. Sourav Bandyopadhyay
University of California, San Diego
sourav@gmail.com

Dr. J. Fernando Bazan
Genentech
jfbazan@gene.com

Dr. Jeremy Bellay
University of Minnesota
bellay@cs.umn.edu

Dr. Pedro Beltrao
UCSF - Lim lab and Krogan Lab
pedro.beltrao@ucsf.edu

Dr. Charlie Boone
University of Toronto
charlie.boone@utoronto.ca

Ms. Susanne Bornelöv
Uppsala University
susanne.bornelov@gmail.com

Dr. Clive Bowsher
University of Cambridge
C.Bowsher@statslab.cam.ac.uk

Dr. Pascal Braun
Dana-Farber Cancer Institute
pascal_braun@dfci.harvard.edu

Dr. Roger Brown
University of Edinburgh
Roger.Brown@ed.ac.uk

Mr. Jordi Carreras Puigvert
Leiden University
j.c.puigvert@chem.leidenuniv.nl

Mr. Guilhem Chalancon
MRC Laboratory of Molecular Biology
guilhem@mrc-lmb.cam.ac.uk

Mr. Tamir Chandra
Cancer Research UK/ Cambridge University
tamir.chandra@cancer.org.uk

Dr. Vivian Cheung
University of Pennsylvania
vcheung@mail.med.upenn.edu

Ms. Ara Cho
Yonsei University
araaraara@naver.com

Dr. Francesca Ciccarelli
IEO - European Institute of Oncology
francesca.ciccarelli@ifom-ieo-campus.it

Dr. Krzysztof Cios
Virginia commonwealth University
kcios@vcu.edu

Dr. Jacques Colinge
Head of Bioinformatics
office@cemm.oeaw.ac.at

Dr. Raju Datla
National Research Council
patricia.eilersen@nrc-cnrc.gc.ca

Prof. Geert De Jaeger
VIB-Ghent University
gejae@psb.ugent.be

Dr. Natalie De Souza
Nature Methods
n.desouza@natureny.com

Dr. Richard Edwards
University of Southampton
r.edwards@southampton.ac.uk

Ms. Dorothea Emig
Max Planck Institute for Informatics
demig@mpi-inf.mpg.de

Dr. Frank Emmert-Streib
Queen's University Belfast
v@bio-complexity.com

Dr. Hai Fang
University of Bristol
hfang@cs.bris.ac.uk

Mr. Mike Fletcher
University of Cambridge
mike.fletcher@cancer.org.uk

Dr. Franca Fraternali
King's College London UK
franca.fraternali@kcl.ac.uk

Dr. Katheleen Gardiner
University of Colorado Denver
katheleen.gardiner@ucdenver.edu

Dr. Anne-Claude Gavin
EMBL
gavin@embl.de

Dr. Michael Gilchrist
National Institute for Medical Research
m.gilchrist@nimr.mrc.ac.uk

Dr. Adam Godzik
Burnham Institute for Medical Research
adam@burnham.org

Dr. Andrey Grigoriev
Rutgers University
agrigoriev@camden.rutgers.edu

Ms. Michal Harel
Weizmann Institute of Science
mic.harel@weizmann.ac.il

Dr. Frank Holstege
Genomics Laboratory, UMC Utrecht
f.c.p.holstege@med.uu.nl

Dr. Trey Ideker
University of California, San Diego
trey@bioeng.ucsd.edu

Mr. Sun-Gou Ji
Yonsei University
ji.sungou@gmail.com

Dr. Nick Jones
Oxford University
nick.jones@physics.ox.ac.uk

Dr. Allan Jones
Allen Institute for Brain Science
allanj@alleninstitute.org

Ms. Shichina Kannambath
University of Manchester
shichina.k@postgrad.manchester.ac.uk

Dr. Hagit Katzov-Eckert
University of São Paulo
kehagit@gmail.com

Mr. Gaurav Kaushik
Postgraduate Institute of Medical Education & Rese
gauravkaushik21@gmail.com

Mr. Samuel Kerrien
EMBL-EBI
skerrien@ebi.ac.uk

Dr. Manfred Kögl
German Cancer Research Center
m.koegl@dkfz.de

Mr. Vyacheslav Koval
Hubrecht Institute
s.koval@hubrecht.eu

Mr. Marcin Kruczyk
Uppsala University
kruczyk@gmail.com

Dr. Tanja Kunej
University of Ljubljana, Biotechnical Faculty
tanja.kunej@bf.uni-lj.si

Dr. Irwin Kurland
Albert Einstein College of Medicine
irwin.kurland@einstein.yu.edu

Mr. Tak Lee
Yonsei University
helpmeout@naver.com

Dr. Ben Lehner
Centre for Genomic Regulation
ben.lehner@crg.es

Dr. Thomas Lufkin
Genome Institute of Singapore
lufkin@gis.a-star.edu.sg

Dr. Edward Marcotte
University of Texas
marcotte@intron.icmb.utexas.edu

Prof. Hanah Margalit
The Hebrew University of Jerusalem
hanahm@ekmd.huji.ac.il

Dr. Florian Markowetz
Cancer Research UK
florian.markowetz@cancer.org.uk

Mr. John May
EMBL-EBI
john@john-may.com

Dr. William McCombie
Cold Spring Harbor Laboratory
mccombie@cshl.edu

Dr. Aline Metris
Institute of Food Research
aline.metris@bbsrc.ac.uk

Mr. Jacob Michaelson
Biotechnology Center, TU Dresden
jacob.michaelson@biotec.tu-dresden.de

Prof. Ahmed Mliki
Centre de Biotechnologie de Borj-Cédria
ahmed.mliki@cbbc.rnrt.tn

Mr. Pablo Moreno
EMBL-EBI
pmoreno@ebi.ac.uk

Mr. Neelanjan Mukherjee
Duke University
nm29@duke.edu

Dr. Atsushi Niida
University of Tokyo
aniida@ims.u-tokyo.ac.jp

Ms. Daniela Nitsch
Katholieke Universiteit Leuven
Daniela.Nitsch@esat.kuleuven.be

Mr. Quentin Nunes
University of Liverpool / Liverpool NIHR PBRU
Quentin.Nunes@liverpool.ac.uk

Dr. Marcela Nunez
Hybrigenics
mnunez@hybrigenics.com

Dr. Taeyun Oh
Yonsei University
terra19@gmail.com

Mr. Roland Pache
IRB Barcelona
roland.pache@irbbarcelona.org

Dr. Vera Pancaldi
UCL
v.pancaldi@ucl.ac.uk

Dr. Dana Pe'er
Columbia University
dpeer@biology.columbia.edu

Ms. Beatriz Penalver Bernabe
Northwestern University
bpenalver@u.northwestern.edu

Dr. Christian Probst
Carlos Chagas Institute
cprobst@fiocruz.br

Dr. Natasa Przulj
Imperial College London
natasha@imperial.ac.uk

Mr. David Quigley
University of California, San Francisco
dquigley@cc.ucsf.edu

Mr. Owen Rackham
University of Bristol
owen.rackham@bristol.ac.uk

Dr. Indika Rajapakse
Fred Hutchinson Cancer Research Center
irajapak@fhcrc.org

Dr. Anna Riegel
Georgetown University
ariege01@georgetown.edu

Dr. Roberto Sacco
CeMM Gmbh
office@cemm.oeaw.ac.at

Dr. Julio Saez-Rodriguez
Harvard Medical School
saezrodriguez@gmail.com

Dr. Christopher Sanderson
University of Liverpool
c.sanderson@liv.ac.uk

Dr. Mihaela Sardiu
Stowers Institute for Medical Research
mis@stowers.org

Ms. Ramona Schmid
Boehringer Ingelheim Pharma GmbH & Co. KG
ramona.schmid@boehringer-ingelheim.com

Dr. Luis Serrano
Centre for Genomic Regulation
luis.serrano@crg.es

Dr. Jun Sese
Ochanomizu University
sesejun@is.ocha.ac.jp

Dr. Teppei Shimamura
University of Tokyo
shima@ims.u-tokyo.ac.jp

Dr. Ramakrishnaiah Siddappa
Leiden University/LACDR
r.siddappa@lacdr.leidenuniv.nl

Mr. Anthony Smith
Medical Research Council (UK)
acs@mrc-mbu.cam.ac.uk

Dr. Montse Soler
IRB Barcelona
montse.soler@irbbarcelona.org

Ms. Amelie Stein
IRB Barcelona
amelie.stein@irbbarcelona.org

Dr. Ulrich Stelzl
Max-Planck Institute for Molecular Genetics
stelzl@molgen.mpg.de

Dr. Alexey Stukalov
CeMM - Research Center for Molecular Medicine GmbH
office@cemm.oeaw.ac.at

Dr. Giulio Superti-Furga
Austrian Academy of Sciences
giulio.superti-furga@cemm.oeaw.ac.at

Dr. Chao Tang
University of California, San Francisco
chao.tang@ucsf.edu

Dr. Andrew Teschendorff
University College London
a.teschendorff@ucl.ac.uk

Dr. Simon van Mourik
Wageningen University and Research Center
simon.vanmourik@wur.nl

Dr. Tim van Opijnen
Tufts University, School of Medicine
tim.van_opijnen@tufts.edu

Dr. Jan Veldink
UMC Utrecht
j.h.veldink@umcutrecht.nl

Dr. Marian Walhout
UMass Medical School
marian.walhout@umassmed.edu

Dr. Anton Wellstein
Georgetown University
wellstea@georgetown.edu

Prof. Kevin White
The University of Chicago
kpwhite@uchicago.edu

Mr. Jonathan Woodsmith
University of Liveprool
jdw@liv.ac.uk

Mr. Peter Woollard
GlaxoSmithKline
peter.m.woollard@gsk.com

Dr. Joo-Yeon Yoo
POSTECH
jyoo@postech.ac.kr

Dr. Rolf Zeller
University of Basel Medical School
rolf.zeller@unibas.ch

Dr. Heng Zhu
Johns Hopkins University
hzhu4@jhmi.edu