



Network Biology 2.0

Connecting Genomes to Disease Progression and Drug Response

SYMPOSIUM PROGRAM, SCHEDULE AND DETAILS

SCHEDULE OF EVENTS

- 8:30 AM *Breakfast provided by GNS*
- 9:15 **Iya Khalil:** Welcome & Opening Remarks
- 9:30 **Eric Schadt, Morning Keynote:** “Elucidating the Complexity of Cancer Networks”
- 10:30 **Paul McDonagh:** “Causal Modeling using Network Ensemble Simulations Predicts Novel Lipid Metabolism Genes”
- 11:30 **Gustavo Stolovitzky:** “Modeling and Simulation of Biological Networks”
- 12:30 PM *Lunch Break*
- 1:30 **Jim Collins:** “Bacterial network biology”
- 2:30 **Gurinder Atwal:** “Population Genetics of the Human MDM4 Oncogene: Evolution and Cancer Risk”
- 3:30 **Dana Pe’er:** “Driving Mutations: Lessons from Yeast and Cancer”
- 4:30 **Mark Boguski:** “Customized Care 2020”
- 5:30 **George Church, Evening Keynote**
- 6:30 *Reception, with light fare and drinks provided by GNS*
- 7:30 **Integrative Biology Roundtable Discussion**
- 8:30 *Conference close.*

LOCATION:

The Broad Institute, 7 Cambridge Center, Cambridge, MA 02142.

The Broad Institute is within easy walking distance of Kendall Square/MIT Station on the Red Line. The symposium sessions will occur in The Broad Auditorium.

HOTELS & ACCOMMODATIONS:

If you wish to stay nearby to the GNS offices, we recommend the Hotel Marlowe, <http://www.hotelmarlowe.com>, 1-800-825-7140. The Boston Marriott Cambridge is also right next door to the Broad Institute, the site of the Symposium: <http://www.marriott.com/hotels/travel/boscb-boston-marriott-cambridge>; 1-617-494-6600.

CONTACT:

Website registration: <http://netbio.eventbrite.com>

Email contact: netbio@gnsbiotech.com

GNS Office Phone: 617 494 0492

PARKING:

Parking in East Cambridge is tight and The Broad does not have space in our lot to accommodate our guests. We do not recommend you park on the street; the meter monitors are extremely vigilant. The nearest parking garage is at 7 Cambridge Center with an entrance off of Ames Street, which is a left turn off of Broadway. For a map and more details, see <http://www.broad.mit.edu/node/1238>

MEALS:

Breakfast will be continental baked goods, juices and coffee. Coffee service will be available throughout the afternoon. The evening reception will feature light appetizers and a selection of beverages.

Eric Schadt

Department of Genetics, Rosetta Inpharmatics, LLC/Merck Research Labs & Founder, SAGE.

MORNING KEYNOTE: ELUCIDATING THE COMPLEXITY OF CANCER NETWORKS

Integrative genomics approaches provide a path to identifying gene networks that underlie common human diseases like obesity, diabetes, heart disease and cancer. We have previously demonstrated how integrating genotypic, gene expression, and clinical data can lead to networks that are predictive of disease associated traits. However, applying such an approach in cancer cohorts comes with a number of challenges given the complexity of the cancer genome and the impact radical genomic rearrangements in the tumor tissue have on molecular states.

Here we review integrative genomics approaches, motivate their utility, and then apply such an approach to a cancer cohort in which tumor and adjacent normal tissues were collected. We characterize changes in the network connectivity structure between the normal and tumor tissues and highlight the extensive association between copy number variation (CNV) and gene expression in the tumor samples. Changes in the correlation structure among expression traits induced by CNVs appear to be non-random and include unidirectional CNV (amplification or deletion) for commonly changed loci. Expression traits that are differentially connected between the normal and tumor tissues are enriched for association to CNVs in cis and in trans. Further, key hub nodes in the normal tissue expression network are enriched for association to CNVs in cis, leading to a loss of connectivity in the tumor tissue network and as a result a loss of coherence in biological processes that are key to normal tissue function.

A comparison of connectivity patterns among tumor stage reveals a progressive increase in connectivity changes that are increasingly associated with CNVs at loci that are also associated with tumor stage, suggesting that changes in the network architecture are the result of random generation of variance followed by selection of combinations that are advantageous to tumor growth and progression. We demonstrate how causal associations among expression traits can be inferred using the CNVs as a perturbation source. These analyses highlight novel treatment strategies and biomarkers to assess disease risk and progression.

Paul McDonagh

Vice President of Discovery Biology, Gene Network Sciences

CAUSAL MODELING USING NETWORK ENSEMBLE SIMULATIONS PREDICTS NOVEL LIPID METABOLISM GENES

We describe a novel simulation strategy for predicting drug targets and functional analysis of genetically defined associations. This strategy, which integrates multiple raw experimental data types, utilizes Bayesian model-averaging over a sample of parameterized networks to achieve robust predictions while accounting for uncertainty of the network structure. We illustrate this method's power by analyzing liver gene expression, serum lipid profiles and body weight measured on 120 male mice from a mouse intercross population. An ensemble of 1024 networks gave accurate predictions of animals that were not part of the training data and explained almost twice the variance compared to quantitative trait loci alone. Further *in silico* experiments identified an additional 38 transcripts – many of them unexpected -- that are predicted to play a significant role in controlling high density lipoprotein and free triglycerides plasma concentrations.

Gustavo Stolovitzky

Manager, Functional Genomics and Systems Biology, IBM Computational Biology Center, T.J. Watson Research Center

MODELING AND SIMULATION OF BIOLOGICAL NETWORKS

The recent explosion of data brought about by high throughput technology has allowed systems biologists to pose a question that previous researchers couldn't pose: What is the map of the biological interactions that confers the cell its behavior and function? In this talk I will discuss a project called DREAM, the Dialogue on Reverse Engineering Assessment and Methods. The DREAM project is fostering a concerted effort by computational and experimental biologists to understand the limitations and enhance the strengths of the efforts to reverse engineer cellular networks from high throughput data. We do this by challenging the community to infer networks from high throughput data. I will highlight the strategies that have achieved the better inference results and discuss the state of the art in Reverse Engineering, as well as some of the challenges and opportunities awaiting us.

Having these networks at hand allows us to investigate their global structure/function organization. One effort in this direction that has recently given some insight has been to analyze these networks to discover topological motifs. I will discuss our study of characterization of a special type of motifs: cyclic motifs, one instance of which are feedback loops. Using tools from statistical physics we developed a theoretical framework for characterizing the ensemble of cyclic motifs. This approach led us to uncover a generic property of real networks, namely an *anti-ferromagnetic* organization of cyclic motifs. We find that biological networks show a depletion of feedback loops, where the number of nodes affected by feedback loops seems to be at a local minimum for real networks compared with surrogate networks. I will present evidence that this topological property allows for a more stable behavior of the concentration of molecular species.

Jim Collins

Professor of Biomedical Engineering & Co-Director of the Center for BioDynamics at Boston University

BACTERIAL NETWORK BIOLOGY

Many fundamental cellular processes are governed by genetic programs which employ protein-DNA interactions in regulating function. Owing to recent technological advances, it is now possible to design synthetic gene regulatory networks, and the stage is set for the notion of engineered cellular control at the DNA level. In this talk, we describe how techniques from nonlinear dynamics and molecular biology can be utilized to model, design and construct synthetic gene regulatory networks. We present examples in which we integrate the development of a theoretical model with the construction of an experimental system. We also discuss the implications of synthetic gene networks for biotechnology, biomedicine and biocomputing.

Additionally, we present integrated computational-experimental approaches that enable construction of quantitative models of gene-protein regulatory networks using expression measurements and no prior information on the network structure or function. We discuss how the reverse-engineered network models, coupled to experiments, can be used to gain insight into the regulatory role of individual genes and proteins in the network, and identify the pathways and gene products targeted by pharmaceutical compounds.

Gurinder “Mickey” Atwal

Member, Institute for Advanced Study

POPULATION GENETICS OF THE HUMAN MDM4 ONCOGENE: EVOLUTION AND CANCER RISK

A large body of evidence strongly suggests that the p53 tumor suppressor pathway is central in reducing cancer frequency in vertebrates. The protein product of the haploinsufficient mouse double minute 2 (MDM2) oncogene binds to and inhibits the p53 protein. Recent studies of human genetic variants in p53 and MDM2 have shown that single nucleotide polymorphisms (SNPs) can affect p53 signaling, confer cancer risk, and suggest that the pathway is under evolutionary selective pressure.

We analyzed the haplotype structure of MDM4, a structural homolog of MDM2, in several different human populations. A Bayesian information-theoretic analysis of linkage disequilibrium in the haplotype distribution of MDM4 indicated the presence of candidate SNPs that may also modify the efficacy of the p53 pathway. Association studies in 5 different patient populations revealed that these SNPs in MDM4 confer an increased risk for, or early onset of, human breast and ovarian cancers in Ashkenazi Jewish and European cohorts, respectively. This study implicates MDM4 as a key regulator of tumorigenesis in the human breast and ovary, but also exploits, for the first time, evolutionary driven linkage disequilibrium as a means to select SNPs of p53 pathway genes that might be clinically relevant.

Dana Pe'er

Assistant Professor of Biology & Computer Science, Columbia University

POPULATION GENETICS OF THE HUMAN MDM4 ONCOGENE: EVOLUTION AND CANCER RISK

We will discuss methods that harness gene expression to identify genetic variants that influence a trait of interest. Our premise is that much of the influence of genotype on phenotype is mediated by changes in the regulatory network and these can be inferred using gene expression. We will demonstrate two such methods: Camelot, an algorithm method that integrates genotype and gene expression collected in a reference condition (un-drugged) and phenotype data to predict complex quantitative phenotypes in entirely different conditions (drug response) and identify causal genes that influence these traits.

We systematically applied our algorithm to a collection of yeast segregants to predict the response to 87/94 drugs and experimentally confirmed 22/24 gene-drug interactions. Our second method, Conexic, a novel Bayesian Network-based framework to integrate chromosomal copy number and gene expression data to detect genetic alterations in tumors that drive proliferation, and to model how these alterations perturb normal cell growth/survival. The underlying assumption to our approach is that significantly recurring copy number change, coinciding with its ability to predict the expression patterns varying across tumors, strengthens the evidence of a gene's causative role in cancer.

We applied Conexic to a melanoma dataset comprising 62 tumor samples and correctly identified most known 'driver' events, while also connecting these to their known targets (e.g. MITF). In addition, our analysis suggests a number of novel drivers, including a number of genes involved in regulation of protein trafficking and endosome biology in this malignancy. Preliminary experimental validation supports several of these findings.

Mark Boguski

Harvard Medical School & Department of Pathology, Beth Israel Deaconess Medical Center

CUSTOMIZED CARE 2020

One of the “disruptive” changes in medical care will be the widespread adoption of precision diagnostics to guide therapeutics. The medical specialty of clinical pathology will play a central role in this new era of personalized medicine. Enabled by advances in both medical re-sequencing technologies and disease pathway modeling and simulation, pathology departments of the future will provide diagnoses of unprecedented specificity. Data-driven reverse engineering of disease processes will identify the underlying molecular pathways that are most relevant in individual patients. These pathways will then be forward-simulated in the presence of virtual drug combinations to predict which therapies will be most effective for individual patients.