predictive models of complex systems

SYMPOSIUM ABSTRACTS

Saturday June 3, 2006

For Making Genetic Networks Operate Robustly, Unintelligent Non-Design Suffices
Garry Odell, Center for Cell Dynamics, University of Washington

Five years ago we (George von Dassow, Ed Munro, Eli Meir, and I) made mathematical/computer models of two ancient and famous genetic networks that act early in diverse embryos to establish spatial gene expression patterns prefiguring the body plan. Our models revealed these networks to be astonishingly robust. That is, they continue to make the correct pattern in the face of thousand-fold variations in the strengths and functional forms of interactions among participating genes. After getting over my surprise that it was even possible to design networks with such properties, I now believe only networks having this kind of robustness can be functionally heritable in polymorphic populations. What design features might endow genetic networks with the kind of extreme robustness we found? To probe for answers, I wrote a computer program that haphazardly generates randomly connected networks made from about the same number of biochemically sensible parts that constitute the segment polarity and neurogenic networks. We (Bjorn Millard, Ed Munro, and I) devised computer algorithms that discover and catalog the stable expression patterns any network can make, and, from all these, distills those patterns the network can make robustly with respect to variations of its parameters. The bottom line is that 19 out of 20 random networks that our program created (i.e. networks devoid of any purposeful design whatever) could make at least one, and usually many, complex stable spatial expression patterns with the same high robustness that the real, evolved, segment-polarity and neurogenic networks exhibit. Several of the random, non-designed networks turn out to be much more robust than either real network. Only 1 out of 20 random networks is a complete loser; it did not make any interesting pattern at all.

Our algorithms for finding patterns any network can stabilize show that it's possible to replace the network's differential equation model, which keeps track of continuous concentrations of gene products changing continuously through time, by discrete logic models with quantized far-apart concentrations. Unfortunately, there are *many* different ways to do this -- different ways for different parameter values, no way appropriate for all parameter values.

The *in silico* result that thoughtless, haphazard, non-design produces networks whose robustness seems inspired begs questioning what else unintelligent non-design might be capable of.

Sunday June 4, 2006

Network Dynamics and Cell Physiology
John J. Tyson, Katherine C. Chen & Bela Novak
Department of Biological Sciences, Virginia Polytechnic Institute & State University

Complex networks of interacting proteins control the physiological properties of a cell (metabolism, reproduction, motility, etc.). Intuitive reasoning about these networks is often sufficient to guide the next experiment, and a cartoon drawing of a network can be useful in codifying the results of hundreds of observations. But what tools are available for understanding the rich dynamical repertoire

of such control systems? How do these behaviors depend on the genetic and biochemical parameters of the system (! gene dosage, enzymatic rate constants, etc)? Using basic principles of biochemical kinetics, one can convert network diagrams into sets of ordinary differential equations and then explore their solutions by analytical and computational methods. I will illustrate this approach with mathematical models based on the regulation of cell growth and division in eukaryotes. During normal growth and division, cell size is the critical parameter that drives progression from G1 to S/G2 to M phase and back to G1. Simple diagrams, which correlate Cdk activity with cell growth, give a new way of thinking about cell cycle control, particularly the role of checkpoint pathways in arresting the cycle. The method is generally applicable to any complex gene-protein network that regulates some behavior of a living cell.

References: Tyson, Chen & Novak (2001) Nature Rev. Mol. Cell Biol. 2:908-916 Tyson, Csikasz-Nagy & Novak (2002) BioEssays 24:1095-1109 Tyson, Chen & Novak (2003) Curr. Opin. Cell Biol. 15:221-231.

How Yeast Cells Mate: Systems Properties from Protein Interactions
Andrew Murray, Matthieu Piel, Chinlin Guo, Center for Modular Biology, Harvard University

Budding yeast has sex. Haploids exist in two mating types, **a** and **α,** which can fuse to each other to form diploids. The haploids communicate with each other using peptide pheromones, which arrest the cell cycle and induce the formation of shmoos, specialized mating predictions. These reactions depend on a well characterized signaling pathway. Where cells form a shmoo and the direction in which the shmoo grows are determined by pheromone gradients. We have combined theory and real time observation of cells exposed to controlled pheromone gradients to find out how cells make only one shmoo at a time and how this process is controlled by pheromone gradients. We can exclude models based on historical marks or lateral inhibition and our data supports a class of models that we call global integration: pheromone binding leads to actin polymerization and the transport of secretory vesicles along actin filaments delivers more signaling molecules, closing a positive feedback loop. The actin cables inside cells integrate transport throughout the cell, and stochastic fluctuations in the distribution of actin can lead to self-amplifying in- homogeneities that lead to cell polarization and shmoo formation.

Global Patterning from Local Signals: the planar cell polarity signaling network
Jeff Axelrod, Claire Tomlin, Keith Amonlirdviman, Dali Ma, Wei-Shen Chen, Narmada Khare, David Tree, Stanford University, School of Medicine, Dept. of Pathology

We would like to derive a detailed understanding of the generation of developmental patterning that emerges from the properties of local signaling events. I will describe combined biological experimentation and mathematical modeling that has yielded important insights into the function of the Planar Cell Polarity (PCP) signaling network. Planar cell polarity (PCP) signaling generates subcellular asymmetry along a tissue axis orthogonal to the apical-basal axis. PCP is governed by a two-tiered system: a global directional cue orients the function of a local Frizzled (Fz) dependent feedback mechanism that aligns cells' polarity with that of their neighbors. The resulting cell polarization is evident in the asymmetric distribution of Fz and other signaling components at the apical adherens junctions. Cell clones mutant for some PCP signaling components, including some, but not all alleles of the receptor Fz, cause polarity disruptions of neighboring, wild-type cells, a phenomenon referred to as domineering non-autonomy. We have demonstrated that our hypothesized local signaling network, combined with a global directional cue, is sufficient to explain all of the characteristic PCP phenotypes without invoking a diffusible though elusive "factor X." Furthermore, we were able to propose and experimentally verify a model explaining why some but not other alleles of fz produce domineering nonautonomy, in which clones of mutant cells perturb polarity of neighboring wild-type cells. In a subsequent set of experiments, we have shown that the local

feedback signal that propagates polarity is sensitive to cell geometry, thereby accounting for the apparently stochastic polarity errors when the global signal is disabled. The mathematical modeling tools developed in the course of this work will be applicable to a broad class of problems in both biological and engineered systems.

Systems Pharmacology and the Treatment of Cancer
Peter Sorger, Center for Cell Decision Processes, Massachusetts Institute of Technology

The overall goal of the MIT Center for Cell Decision Processes (CDP Center) is to understand mammalian signal transduction from a systems perspective in which information on specific molecular mechanisms is linked to cellular physiology through data-driven mathematical models. The CDP Center applies a *modify-measure-mine-model* paradigm and focuses on the study of receptor-mediated death and survival signaling in human cells. In pursuit of its scientific goal, the CDP Center has three programs: (i) interdisciplinary research (ii) training and education at undergraduate, graduate and postdoctoral levels (iii) outreach to a broad audience, particularly in disadvantaged and minority communities.

The cornerstones of CDP research are data-driven mathematical models that range from physico-chemical to statistical (Figure 1c). All rely on rich data sets describing changes in protein and mRNA levels, biochemical activities, post-translational modifications, protein localization in live cells etc. as determined from multiple cells and tissues using a variety of techniques (Figure 1d). The primary foci of modeling efforts are pro-death networks downstream of TRAIL, TNF and FAS receptors and pro-survival networks downstream of EGF, IGF and insulin receptors. To date, we have constructed both large-scale statistical models using Dependent Partial Least Squares Regression (PLSR) that describe interactions among receptor families as well as ODE-based mechanistic models that describe the detailed activities of individual receptor pathways. Data collection centers on a pipe-line of commercial and in-house high throughput technologies, as well as significant efforts in the areas of imaging[1] and Microsystems research[2, 3].

PLSR modeling of crosstalk between death and survival signals in human epithelial cells treated with combinations of TNF, EGF and insulin[4] has revealed that cells respond to TNF directly, via activated TNF receptor and indirectly, via the sequential released of three cytokines: TGF-a, IL-1a, and IL1 receptor antagonist (IL1ra)[5, 6]. This contingent and time-varying series of signals constitutes an autocrine cascade linking cellular responses to the local environment and making death cues self-limiting [6]. A practical insight is the close link between drugs that affect pro-inflammatory signaling by TNF and IL1 (e.g. Remicade, Enbrel, Kineret) and those that inhibit ErbB1 (e.g. Iressa and Erbitux).

A potential weakness of principal components analysis (PCA), of which PLSR is a variant, is that information on which dimensions contribute to principal components (PCs) is usually not informative. We were therefore very surprised to learn that our PLSR model was interpretable in this fashion: one PC was composed of activated versions of proteins in the Akt kinase pathway and the other of activated Jnk1, MK2 and caspase 8 [5]. When the TNF+insulin and TNF+EGF treatments were plotted against the PCs, the response vectors were nearly orthogonal showing that EGF and insulin oppose TNF-mediated death in very different ways. Surprising, a PLSR model developed for TNF, insulin and EGF was shown to perform well in a completely different biological context: predicting cell death following infection with adenovirus, suggesting that statistical models may be valuable outside the scope of the data on which they were trained[7].

Recent research in CDP has focused on the construction of ~300 to 500 species ODE models describing particular receptor subsystems involved in the death and proliferation of human cells. These data are trained on quantitative time-course data and validated using RNAi, drugs and other

means of perturbing cell behavior. Models of TNF and TRAIL-mediated signaling show the value of single cell measurement in dissecting a snap-action, variable delay switch.[8] We have established that proteins hitherto thought to be redundant for death, such as SMAC and Apaf1, are in fact individually essential for correct switching. We speculate that mutations in apoptotic proteins found in human cancer might lead to abnormal states of partial cell death
Our most intensive modeling effort has been devoted to ErbB-mediated signaling in tumor cells. Our current model correctly predicts differences in Iressa responsiveness of human non-small-cell lung cancer cells carrying wild-type ErbB1 receptor and ErbB1 with drug-sensitizing mutations. Using a combination of systematic measurement, modeling and reconstitution in transgenic mice we have shown that ErbB1 mutations have at least three effects: increasing Km for ATP, decreasing internalizing rate and changing the spectrum of receptor hetero-oligomers on the cell surface. We believe that further analysis of these effects will influence the design and use anti-ErbB therapeutics. Based on this work we are interesting in extending systems biology concepts to both theoretical and data-driven analysis of pharmacological action[9]. We refer to this emerging area as "computational and systems pharmacology" and envision it as a bridge between genomic medicine and drug discovery.

Citations

1. Goldberg, I.G., et al., The Open Microscopy Environment (OME) Data Model and XML file: open tools for informatics and quantitative analysis in biological imaging. Genome Biol, 2005. 6(5): p. R47.
2. El-Ali, J., P.K. Sorger, and K. Jensen, Cells on Chips. Nat Insight, 2006. in press.
3. Burg, T.P., et al., Vacuum-Packaged Suspended Microchannel Resonant Mass Sensor for Biomolecular Detection. J. Microelemechanical Systems (MEMS), 2006. in press.
4. Gaudet, S., et al., A compendium of signals and responses triggered by prodeath and prosurvival cytokines. Mol Cell Proteomics, 2005. 4(10): p. 1569-90.
5. Janes, K.A., et al., A systems model of signaling identifies a molecular basis set for cytokine-induced apoptosis. Science, 2005. 310(5754): p. 1646-53.
6. Janes, K.A., et al., The response of human epithelial cells to TNF involves an inducible autocrine cascade. Cell, 2006. 124(6): p. 1225-39.
7. Miller-Jensen, A., K.A. Janes, and D.A. Lauffenburger, Cell-specific signalling and apoptosis are linked by a common effector-processing mechanism. 2006. submitted.
8. Albeck, J.G., et al., Mathematical modeling and single-cell analysis of a snap-action feed-forward switch controlling receptor-mediated apoptosis in preparation, 2006.
9. Fitzgerland, J., et al., Systems Biology and Combination Therapy in the Quest for Clinical Efficacy. Nature Chemical Biology, 2006. submitted.

Please note, all CDP manuscripts are available on line and cdpcenter.org

Predictive Models of Infectious Diseases
Irene Eckstrand, NIH/NIGMS

The Models of Infectious Disease Agent Study (MIDAS) program is a collaborative network of scientists who conduct research on the use of computational and mathematical models that will prepare the nation to respond to outbreaks of infectious diseases. MIDAS consists of eight collaborative agreements – seven research groups and one informatics groups – currently focused on models of pandemic influenza. Model results are one source of information used by U.S. government and by the World Health Organization to formulate policies to contain and mitigate outbreaks. More information about MIDAS is available at the MIDAS Portal at:
https://www.epimodels.org/midas/about.do.

The Mysterious Meanderings of Myxobacteria
George Oster, University of California

The life cycle of myxobacteria resembles in many respects that of the well-studied slime mold *Dictyostelium discoideum*. These rod-shaped cells propel themselves by a mechanism called "gliding" and aggregate into giant swarms that hunt by secreting exoenzymes that lure and digest prey. When food is scarce, they coalesce into fruiting bodies containing the spores that will seed the next generation. During this aggregation they pass through a developmental stage called the "ripple phase" characterized by elaborate patterns of waves that propagate over the colony surface. These waves generate the same kinds of patterns observed in *Dictyostelium* aggregations, including bulls-eye and spiral patterns. However, these patterns are unlike those in *D. discoideum* in several crucial respects: (i) they can persist for long periods in the absence of mass transport, (ii) colliding waves appear to "pass through" one another, analogous to soliton waves in water, whereas D. discoideum waves, like those in chemical wave systems, annihilate one another when they meet. (iii) the spatial patterns in D. discoideum are organized by relaying diffusible morphogens, whereas myxobacteria communicate by direct cell contact only. Subsequent to, or coincident with, the ripple phase, cells commence aggregating into fruiting bodies that can take on quite elaborate shapes. I will present a model for the wave and aggregation patterns in myxobacteria that quantitatively explains many of their observed characteristics. Aside from the novel mechanism that generates these patterns, the model shows how the patterns can be used as a probe of the intercellular signal transduction mechanism, and provides an alternate system for studying multicellular pattern formation based on direct cell contact rather than diffusible morphogens.

Virtual Cell Project
Leslie Loew, University of Connecticut Health Center

Cells and their subcellular organelles can have complex structures that provide a framework for the spatial distribution of signaling molecules. But how this cellular architecture shapes and controls the response of cells to their environment is a question that is not often addressed, perhaps because of the lack of appropriate systems biology modeling tools. It facilitates the organization of experimental date into quantitative analysis of data obtained from fluorescent probe experiments in live cells. A key feature of the Virtual Cell is that it permits the incorporation of experimental microscope images within full 3D spatial models of signal transduction networks. The operation of the Virtual Cell will be illustrated with PDE models of several important cell signaling and membrane transport processes.

Simulations Indicate that Cell Shape is Sufficient to Initiate Polarity in Fission Yeast
Francois Nedelec, Dietrich Foethke, Damian Brunner, EMBL-Heidelberg

The fission yeast *Schizosaccharomyces pombe* maintains a cylindrical shape by alternating polarized growth at the cell ends with divisions in the middle. The localized growth is established during interphase by four to six bundles of microtubules that deposit growth factors specifically at cell ends. The bundles are anti-parallel and align along the main cell axis. Furthermore, they are attached to the nucleus, and position it near the cell-center. We model the motion of the nucleus and microtubules by simulated Langevin dynamics in three dimensions. We find that proper localization of microtubule catastrophes is essential, and that a model in which microtubules are affected by force when pushing on the cortex fits most of the experimentally measured cell features. This model does not include any positional information, other than the shape of the cell. It suggests therefore that a proper shape is sufficient to establish cell polarity.

Monday June 5, 2006

Parameter Estimation for Large-Scale, Under-Determined Systems with Application to Dynamic Models of Cellular Metabolism
Daniela Calvetti, Center for Modeling Integrated Metabolic Systems,
Case Western Reserve University

The estimation of the parameters which specify mathematical models for metabolism at the cellular and subcellular level is a challenging and difficult problem, because of the complexity of the models and the scarcity of available data. In general, since dynamic metabolic models consist of large systems of coupled nonlinear differential equations and the data of a few values of biochemical species at a few observation times, the estimation problems are severely underdetermined and ill-conditioned. Therefore, a successful solution requires additional information about the parameters to be supplied. In this talk we show how the Bayesian paradigm is a natural framework for the introduction of available *a priori* information of various nature into the numerical methods for the parameter estimation. The power of this approach is illustrated with applications to the estimation of the parameters for dynamic models of cellular metabolism for cardiac and skeletal muscles.

Uncovering Chemotaxis: Where are we headed?
Pablo Iglesias, Department of Electrical & Computer Engineering, Johns Hopkins University

Chemotaxing cells, such as *Dictyostelium* and mammalian neutrophils, sense shallow chemoattractant gradients and respond with highly polarized changes in cell morphology and motility. Uniform chemoattractant stimulation induces the transient translocations of several second messengers, including PI3K, PTEN, and PI(3,4,5)P3. In contrast, static spatial chemoattractant gradients elicit the persistent, amplified localization of these molecules. We have proposed a model in which the response to chemoattractant is regulated by a balance of a local excitation and a global inhibition (LEGI), both which are controlled by receptor occupancy. The LEGI model can account for both the transient and spatial responses to chemoattractants, but alone does not amplify the external gradient. In this talk we show how parallel LEGI mechanisms induce an amplified $PI(3,4,5)P_3$ response that agrees quantitatively with experimentally obtained PH-GFP distributions. In doing so, we also highlight how coordinated modeling and experimentation can be used to uncover complex biological behavior.

Dynamic Protein Clusters in Bacterial Chemotaxis
Karen Lipkow, University of Cambridge

The interior of cells is a complicated jungle of macromolecules: some highly organised as large complexes or as part of the cytoskeleton, others roughly clustered together or simply crowding the space. I am interested in how the cellular architecture comes about and how it affects and is affected by signalling. I am exploring these questions, focusing mainly on the well-studied bacterial chemotaxis pathway of *Escherichia coli*, which enables the bacterium to swim to the most beneficial environment. Signals from the environment are detected by transmembrane receptor-kinase complexes, which are mostly clustered at the cell pole, and then transmitted to the randomly positioned flagellar motors by diffusion of the phosphorylated CheYp protein. A component of the pathway that promotes dephosphorylation of CheYp, the protein CheZ, has been shown by fluorescent methods to be distributed between the cytoplasm and the receptor cluster.

With the aid of a computer program that can simulate the movement and interaction of a large number of individual molecules in a structured environment [1], we constructed a three-dimensional model of an *E. coli* cell. We examined the generation and diffusion of CheYp through the cell under control conditions and in response to attractant and repellent stimuli. The results agree well with experimental

observations but allow an analysis at much higher detail, such as the calculation of diffusion traces and lifetimes of individual molecules. Exploring the effects of cellular architecture, macromolecular crowding and positioning of the CheZ phosphatase, we identified conditions for the formation of gradients of phosphorylated CheY [2].

Supported by analytical methods and simulations, we demonstrate that intracellular gradients can have an unexpectedly complicated form. This will occur, for example, if one of the phospho-states binds to a large or immobile structure and needs to be taken into account when measuring gradients experimentally, for example via FRET probes [3]. I formulated a model in which CheZ dynamically changes location and self-organises into oligomeric clusters of higher activity at the pole depending on stimulus level. Our simulations suggest that the changing location of CheZ will sharpen responses of the cell, make adaptation more precise, and increase the range of detectable ligand concentrations. They introduce an unprecedented level of sophistication into what is usually considered a simple signalling pathway [4].

Collaborating with experimentalists who were puzzled by their data, we recreated the movement of cytosolic molecules in bacteria and of membrane proteins in mammalian auditory hair cells [5,6]. These simulations enabled us to understand the recovery, clearance and noise of the fluorescently labelled proteins, hinting at the underlying intracellular structure and dynamics. I will present some preliminary simulation results on the effects of macromolecular crowding and binding affinities.

1. Andrews, S. S., and D. Bray. 2004. Stochastic simulation of chemical reactions with spatial resolution and single molecule detail. Phys Biol 1:137-51.
2. Lipkow, K., S. S. Andrews, and D. Bray. 2005. Simulated diffusion of phosphorylated CheY through the cytoplasm of *Escherichia coli*. J Bacteriol 187:45-53.
3. Lipkow, K. and D. Odde. 2004. Mathematical and computational model for protein concentration gradients in the cytoplasm driven by ATP and GTP hydrolysis. Presented at the American Society for Cell Biology Annual Meeting, Washington, DC. *In preparation*.
4. Lipkow, K. 2006. Changing cellular location of CheZ predicted by molecular simulations. PLoS Comput Biol 2:e39.
5. DePristo, M. A., L. Chang, K. Lipkow, S. Khan, and R. D. Vale. 2004. A stable chemosensory machine in Escherichia coli revealed by FRAP. Presented at the American Society for Cell Biology Annual Meeting, Washington, DC. *unpublished*.
6. Grati, M., M. Schneider, K. Lipkow, E. Strehler, R. J. Wenthold, and B. Kachar. 2006. Rapid turnover of stereocilia membrane proteins: evidence from the trafficking and mobility of plasma membrane Ca2+-ATPase 2. J Neurosci*, in press*.

Towards a Comprehensive Understanding of Cellular Metabolism
Joshua Rabinowitz, Center for Quantitative Biology, Princeton University

A key current challenge in biology is relating the behaviors of cellular systems to the molecular properties of the underlying biochemical networks. Metabolism provides a uniquely well mapped and medically important network to study in this context. Systems level experimental investigation of metabolism has been hampered, however, by the absence of tools for comprehensive metabolite measurement. Here we apply liquid chromatography – tandem mass spectrometry to quantify ~ 100 different known small molecule metabolites in parallel, and to enable dynamic measurements of the influx of isotope labeled nutrients into these compounds. Application of these metabolite and flux measurement tools to study the global response of /E. coli/ and /S. cerevisiae /to nutrient starvation reveals a highly conserved program of starvation-induced metabolome remodeling, with homeostasis maintained for only a small subset of metabolites. Due to the strong conservation of the metabolic

network between bacteria, yeast, and humans, similar techniques should also provide insight into metabolic alterations occurring in human disease.

Health Care of the 21st Century: Predictive, Preventive, Personalized and Participatory (P4) Medicine
Leroy Hood, Center for Systems Biology, Institute of Systems Biology

The Human Genome Project has catalyzed fundamental changes in the practice of biology and medicine. One of these has been to generate a genetics parts list that includes all human genes (and by inference all human proteins). Analysis of the information from the Human Genome Project has also catalyzed the view that "biology is an informational science." Together, these advances have promoted the idea of systems biology—the view that biology can only be understood through an analysis of biological machines and networks. The corresponding systems view of disease has catalyzed revolutionary changes in how to think about diagnosis, therapy and even prevention. Fueled by dramatic changes in in vitro and in vivo measurement technologies, systems medicine is pushing toward a revolution in health care—one that over the next 5-20 years will lead away from the current reactive medicine (wait until one gets sick before treating) to a medicine that is predictive, preventive, personalized and participatory. Increasingly, the focus will be on wellness rather than disease. P4 medicine will require billions of measurements on each patient and the means of reducing these measurements to coherent hypotheses about the health and disease of individual patients. How the acquisition, storage, mining, integration (of different data types), modeling and eventually distribution of the data and the resulting inferences will occur will be one of the grand challenges of this future in health care. Security and access will be critical considerations. This P4 medicine will catalyze fundamental changes in virtually every aspect of the healthcare system and it will require rethinking the educational requirements for physicians. It will lead to the digitalization of medicine with changes even more profound than the digitalization of information technologies and communications. It will also lead to a turn around in the inexorably increasing healthcare costs—with the possibility of bringing developed world medicine to the developing world. Medicine will truly become an informational discipline—with the enormous potential for individuals to take an active role in helping to guide their future health choices. The digitalization of medicine will transform the computational requirements of health care in ways that we can only begin to imagine.

Co-Adaptated Alleles and the Evolution of Complex Traits
Gary Churchill, Center for Genome Dynamics, The Jackson Laboratory

Why are the most common disease related traits genetically complex? In this talk, I will consider complex trait genetics from an evolutionary perspective. I will argue that the genetic factors determining complex disease inheritance may be fundamentally different from those involved in Mendelian disorders. Adherence to the concept of disease susceptiblility loci may limit our ability to discern the genetic loci and mechanisms that are responsible for variation in common disease-related traits. Allelic co-adaptation is a mechanism that may be implicated in many instances of complex disease inheritance. I will provide examples from QTL mapping in the mouse to illustrate this point. Our studies of a complex type II diabetes model provide an example in which a mechanism of co-adaptation can be described at the molecular level. I will present evidence that loci with co-adapted alleles may be widespread throughout the laboratory mouse genome. The mixing of diverse subspecies and intense selection during inbreeding have produced a genome wide signature of selection for co-adapted alleles. These observations suggest that our efforts to identify the "bad" alleles responsible for complex disease traits should be augmented with approaches that consider allelic combinations at multiple loci as the genetic cause of many complex disease traits.

NHLBI Exploratory Program in Systems Biology
Jennie Larkin, National Heart, Lung, and Blood Institute Systems Biology Program

NHLBI has two new programs in systems biology. The Exploratory Programs in Systems Biology is an RFA that will support collaborative, multiple-PI research that applies systems biology approaches to heart, lung, blood, and sleep research. Each award will be a cluster of at least two individual grants to computational and biomedical PIs, who will work collaboratively to both develop quantitative, predicative models and to validate these models with experimental data. The second NHLBI program supports short courses to introduce biomedical researchers to the tools and application of computational techniques in their area of research (Computational Modeling for Heart, Lung, Blood, and Sleep Biologists: Introductory Courses).

Dynamics of Cell Polarization in the Early C. elegans Embyo
Ed Munro, Chris Schoff, Adriana Dawes, Eliana Hechter and Howard Clarke,
Center for Cell Dynamics, University of Washington

The ability to polarize is fundamental to all cellular life. Here, I describe our attempts to comprehend how the interplay between PAR proteins - conserved biochemical regulators of polarity – and the acto-myosin cytoskeletal mechanics, orchestrates the establishment and maintenance of cellular polarities in the early *C. elegans* embryo. Polarization begins when an asymmetrical acto-myosin-based contraction drives cortical flows away from a localized sperm cue, causing a redistribution of the PAR proteins, which in turn regulate one another and modulate cortical contractility and flows, leading to the formation and maintenance of cortical PAR domains. While many of the key components of this system have been characterized, the complexity and distributed nature of their interactions makes it impossible to say if and how they add up to a robust polarization mechanism. We have developed a computational model that integrates known PAR protein biochemistry and conserved cytoskeletal mechanics to predict spatiotemporal patterns of cortical deformation, flow and protein distribution. Here we focus on key insights, predictions, and experimental conformations. Analysis of the model shows that local cycles of acto-myosin assembly, contraction and disassembly observed during polarization in living embryos represent contractile instabilities that emerge as a robust consequence of the conserved mechanochemistry of acto-myosin networks. The model predicts that tuning e.g. F-actin assembly/disassembly rates or myosin activity can promote or defeat local contractile instabilities and we have verified these predictions experimentally. In the model, contractile instability acts as a "mechanical shunt" that lowers cortical tension and thus limits the extent of cortical flow and PAR protein segregation in response to a localized cue. This leads to the counterintuitive prediction that lowering myosin activity in the embryo to defeat contractile instability will enhance cortical flows and produce hyper-segregation of the PAR proteins, which we have confirmed experimentally. Finally, when we added known interactions among anterior and posterior PAR proteins, and between PAR proteins and the acto-myosin cytoskeleton, the model failed initially to produce sustained cortical flow and stable polarity from transient local cue as seen in the embryo. Analysis of this failure pointed to key defects in our original assumptions about the organization of the cortical acto-myosin cytoskeleton. A more detailed experimental analysis verified these defects and suggested revised assumptions, which yielded the "correct" model behavior. These results illustrate how a close interplay between computer simulations and experiment can yield deeper insights into complex biological mechanisms than either alone.

Cytoskeletal Regulation by the Adenomatous Polyposis Coli Protein in Interphase and Mitosis
Inke Näthke, University of Dundee

Mutations in the adenomatous polyposis coli protein (APC) are common to most colonic tumors. One well-characterized function of the APC protein is its ability to support the assembly of a protein complex that regulates the degradation of α-catenin in a Wnt-regulated manner. Accumulation of α-

catenin leads to changes in the activity of TCF/Lef transcription factors and this defect has been implicated in the transformation produced by APC mutations. In addition, APC is an important regulator of the cytoskeleton. Inactivating APC in cells leads to changes in cell migration, but also disrupts mitotic spindles and causes defects in cell division. Thus mutations in APC, which occur extremely early during tumoragenesis, lead to the accumulation of cells in an inappropriate environment, changes in their differentiation and an in increase in genetic instability. I will describe our work in a number of experimental systems, including early chicken embryos, gut tissue, and cultured cells, that investigates the effect of APC on the cytoskeleton.

The Secret Life of Actin
R. Dyche Mullins[1], Ethan Garner, Chris Campbell, Lea Trichet, Jonathan Alberts[2]
UC, San Francisco School of Medicine, [2]Center for Cell Dynamics, University of Washington

The mechanisms that provide force to segregate bacterial chromosomes are still mysterious. We do, however, understand one important example of bacterial DNA segregation in molecular detail – segregation of the R1 and R100 drug-resistance plasmids. These large (100kb), low-copy plasmids encode genes for antibiotic and heavy-metal resistance and have been isolated from many pathogens. To ensure inheritance by both daughter cells during division, the R1 par operon constructs a simple DNA-segregating machine from three components. One of these components, ParM, is related to eukaryotic actins and assembly of ParM into actin-like filaments appears to drive plasmid segregation directly. We find that this simple prokaryotic cytoskeleton exhibits a remarkable collection of activities usually associated with eukaryotic cytoskeletons, including: dynamic instability, processive capping, insertional polymerization, and the ability to generate force. We also find that R1 plasmid segregation is a dynamic process in which assembly of unstable ParM filaments induces plasmids to oscillate rapidly from pole to pole of the cell producing a dynamic rather than static bipolar distribution. Computer modeling of plasmid capture and segregation indicates that the assembly dynamics of prokaryotic cytoskeletal systems are important to their cellular function and that the prokaryotic systems also make use of mechanisms similar to those of the eukaryotic cytoskeleton to establish long-range order and to move intracellular cargo.

Tuesday June 6, 2006

Genomic Approaches to Common Chronic Disease
Andrew G. Clark[1], Eric Boerwinkle[2], James E. Hixson, Alan R. Templeton[3] and Charles F. Sing[4]
[1]Cornell University, [2]University of Texas, Houston, [3]Washington University and [4]University of Michigan

An ultimate goal of genomics is to improve human health. This necessitates spanning the gap from model systems and controlled experiments to populations of humans. There are great difficulties and challenges in making this transition, but also great opportunities. First, human populations have high levels of genetic diversity in most regions of the genome. Thus, the genetic state space is massive and usually has to be reduced in some way to allow practical searches for genotype-phenotype associations. Second, the amounts and frequencies of genetic variants are influenced by past mutations, historical events and patterns of gene flow in natural populations. Artifacts of genotype-phenotype associations can be created by sampling that ignores evolutionary history and population structure. Third, most of the medically relevant phenotypes are complex in that they arise from multiple interacting components, involving both interactions among gene effects and interactions of genotypic effects with environmental effects. The frequencies of both genotypes and environments are not controlled in natural populations. Hence, our ability to detect interaction effects is strongly influenced by the frequencies of the interacting elements. Consequently, the focus in human genetics is on statistical modeling for prediction rather than deterministic modeling of etiology. Investigators

in the **MICORTEX (**MIchigan**, COR**nell**, TEX**as**)** project are (1) testing assumptions that the biomedical research community has accepted when analyzing the relationship between genomic variation and variation in the onset, progression and severity of a common chronic human disease and (2) developing alternative analytical strategies for researching the most important challenges facing genetic studies. These challenges include:

- *How do we reduce the genetic state space so as to retain the most relevant biological signals?*
- *What analytical strategy best combines data mining to select models with the evaluation of the predictability of the models selected?*
- *What is the nature of hidden population substructure and how should it be dealt with in the analyses of genotype-phenotype relationships?*
- *What is the nature of the mutational process and how does it constrain our search for genotype-phenotype associations?*
- *Can genetic variation that influences risk factor variation also improve prediction of disease outcome?*
- *Is there genetic variation that influences risk of disease beyond that captured in the traditional risk factor variation?*
- *What is the appropriate dimensionality to consider for gene-gene and gene-environment interactions affecting risk of a complex chronic disease?*

Answering these questions by modeling the true causal agents and developing an understanding of the processing of those agents is not practical at present in studies of natural populations. Accordingly, our focus is on identifying which DNA sequence variations, in which genes, in which individuals, in which populations and in which environmental strata within a population replicably *predict* variation in disease risk. We will present our experience with three strategies to select models and evaluate the predictability of the models selected. Each strategy will be illustrated using human data collected in population-based studies.

First, we have developed a Combinatorial Partitioning Method (CPM) that does not require a pre-specified genetic model, and therefore is free of the assumption that causation is attributable entirely to the actions of one or two variable sites. The CPM simultaneously considers multiple genes, each containing multiple variable sites, to identify partitions of multi-site genotypes that predict inter-individual phenotypic variation. Our work suggests that traditional methods of building multi-site models that rely on statistically significant single-site effects can fail to identify combinations of sites that best predict trait variability.

Second, we have developed a Markov Chain Monte Carlo (MCMC) method for identifying combinations of SNPs that define genotypic classes that predict trait variation. An expression is obtained for the posterior probability of a configuration of risk sets, and an MCMC algorithm is set up to sample across the possible configurations of genotypic classes. An application of the method to human cardiovascular risk data identified interactions involving up to four different factors. Based on simulations for various genetic architectures, the MCMC method has more power than conventional linear models to detect deleterious combinations of alleles when multiple SNPs and interactions are present. The method can easily accommodate environmental factors and a test for genotype by environment interaction even in a high-dimensional setting.

Third, we have developed tree scanning that uses the evolutionary origins of currently existing haplotypes to efficiently search for genotype-phenotype associations. We have shown that incorporating evolutionary history into our analysis increases power beyond that possible with single SNP associations and provides a better explanatory framework for formulating new hypotheses. We have also extended tree scanning to investigate epistasis and pleiotropy. Supported by NIGMS grant GM065509.

Noisy cellular decision-making: from temporal to spatial choices
Alexander van Oudenaarden, Massachusetts Institute of Technology, Department of Physics

Feedback regulation is often used in gene and protein networks to define a stable decision switch. I will discuss two example of networks that use this feedback strategy in both a temporal and spatial sense.

First, I will discuss the galactose signaling pathway in budding yeast. This network contains multiple nested feedback loops. From the two positive feedback loops only the Gal3p-mediated loop is able to generate two stable expression states with a persistent memory of previous galactose consumption states. The parallel, Gal2p-mediated loop only increases the expression difference between the two states. A negative feedback through Gal80p reduces the strength of the core positive feedback. Despite this fact, a constitutive increase of the Gal80p concentration tunes the system from having destabilized memory to persistent memory. A model reveals that fluctuations are trapped more efficiently at elevated Gal80p levels. Indeed, the rate at which single cells randomly switch back-and-forth between expression states, was reduced.

Second, I will discuss a protein network that utilizes feedback regulation to control spontaneous cell polarization in budding yeast. Cellular polarization is often a response to distinct extracellular or intracellular cues, such as nutrient gradients or cortical landmarks. However, in the absence of such cues, some cells can still select a polarization axis at random. Positive feedback loops promoting localized activation of the GTPase Cdc42p are central to this process. I will discuss spontaneous polarization during bud site selection in yeast cells that lack functional landmarks. We find that these cells do not select a single random polarization axis, but continuously change this axis during the G1 phase of the cell cycle. This is reflected in traveling waves of activated Cdc42p which randomly explore the cell periphery. Our integrated computational and in vivo analyses of these waves reveal a negative feedback loop that competes with the positive feedback loops to regulate Cdc42p activity and confer dynamic responsiveness on the robust initiation of cell polarization.

Dynamics of Probabilistic and Transient Differentiation
Gurol Suel, Jordi Garcia-Ojalvo, Louisa Liberman, and Michael Elowitz, California Institute of Technology

Certain types of cellular differentiation are probabilistic and transient. In such systems individual cells can switch to an alternative state and after some time switch back again. In Bacillus subtilis, competence is such a transiently differentiated state associated with the capability for DNA uptake from the environment. Individual genes and proteins underlying differentiation into the competent state have been identified, but it has been unclear how these genes interact dynamically in individual cells to control both spontaneous entry into competence and return to vegetative growth. Here we show, through modeling and experiments, that this behavior can be understood in terms of excitability in the underlying genetic circuit.