

Two Decades of Predictive Biology
March 17-18, 2007



Organizing Committee

Claude Campbell
H. Eugene Stanley
Zhiping Weng

Saturday, March 17, 2007

10:15 a.m. Welcoming Remarks
Robert Brown, President, Boston University
David Campbell, Provost, Boston University

Session I **Moderator: Gyan Bhanot, Ph.D.**
Professor, Department of Biomedical Engineering and Biomaps Institute
Rutgers University

10:30 a.m. Itai Yanai, Ph.D.
Postdoctoral Fellow
Hunter Lab, Department of Molecular and Cellular Biology, Harvard University
Title: Chance and Necessity in the Manifestation of Genetic Programs

11:15 a.m. Zhiping Weng, Ph.D.
Associate Professor, Department of Biomedical Engineering, Boston University
Title: Transcription Factor Binding and Modified Histones in Human Bidirectional Promoters

12:00 p.m. Jill Mesirov, Ph.D.
Director and Chief Informatics Officer, Computational Biology and Bioinformatics
Broad Institute of MIT and Harvard
Title: Knowledge-based Paradigms for Computational Genomics

12:30 p.m. Minoru Kanehisa, Ph.D.
Director and Professor, Bioinformatics Center, Institute for Chemical Research, Kyoto University
Professor, Human Genome Center, Institute of Medical Science, University of Tokyo
Title: Linking Genomes to Biological Systems and Environments

1:00 p.m. Break

Session II **Moderator: Vladimir Brusic, Ph.D., Director of Bioinformatics, Cancer Vaccine Center, Dana-Farber**
Cancer Institute, Harvard Medical School

2:30 p.m. Micah Dembo, Ph.D.
Professor, Department of Biomedical Engineering, Boston University
Title: Dynamics of Cellular Traction Forces

3:00 p.m. Boris Shakhnovich, Ph.D.
Postdoctoral Associate
Department of Molecular and Cellular Biology, Harvard University
Title: Selective Constraints in Evolution of Gene Families and TBP-Dependent Promoters

3:30 p.m. John N. Weinstein, M.D., Ph.D.
Head, Genomics & Bioinformatics Group, LMP, CCR, NCI, NIH
Title: Integromic Molecular Profiling in Cancer Pharmacology and Therapeutics

4:00 p.m. James Collins, Ph.D.
Professor, Department of Biomedical Engineering, Boston University
Title: Predictive Biology by Design

4:30 p.m. Chris Sander
Head, Computational Biology Center, Memorial Sloan-Kettering Cancer Center
Title: Biomolecular networks: representation, perturbation and function

Sunday, March 18, 2007

Session III Moderator: Boris Shakhnovich, Ph.D., Postdoctoral Associate, Department of Molecular and Cellular Biology, Harvard University

9:30 a.m. Gyan Bhanot, Ph.D.
Professor, Department of Biomedical Engineering and Biomaps Institute
Rutgers University

10:00 a.m. Hanah Margalit, Ph.D.
Professor, Faculty of Medicine, Department of Molecular Genetics and Biotechnology
The Hebrew University of Jerusalem
Title: Role of Non-coding RNA in the Cellular Regulatory Networks

10:30 a.m. Avrum Spira, M.D.
Assistant Professor of Medicine and Assistant Professor of Pathology and Laboratory Medicine, Pulmonary
Center, Boston University Medical Center
Title: Developing Airway Epithelial Gene Expression Biomarkers for Early Lung Cancer Detection

11:00 a.m. Douglas Lauffenburger, Ph.D.
Uncas & Helen Whitaker Professor of Bioengineering & Director Biological Engineering Division,
Massachusetts Institute of Technology
Title: Cue-Signal-Response Models for Predictive Understanding of Signaling Network Control of Cell
Phenotypic Behavior

11:30 a.m. Charles Cantor, Ph.D.
Chief Scientific Officer, Sequenom
Professor, Department of Biomedical Engineering, Boston University
Title: Sensitive In Vivo Detection of Specific RNA Species

Break

**Session IV Moderator: Lubomir Chitkushev, Ph.D., Assoc. Professor and Chairman,
MET Computer Science, Boston University**

1:30 p.m. Jay Berzofsky, M.D., Ph.D.
Branch Chief, Center for Cancer Research Vaccine Branch, NCI, NIH
Title: Engineering Vaccines for Cancer and HIV

2:00 p.m. Matthew Pincus, M.D., Ph.D.
Professor of Pathology
Pathology Department, SUNY Downstate Medical Center
Title: Molecular Modeling of Anti-Cancer Peptides that Block Cancer but not Normal Cell Growth

2:30 p.m. Byron Goldstein, Ph.D.
Fellow, Theoretical Biology and Biophysics, Theoretical Division,
Los Alamos National Laboratory
Title: The Ubiquitous Role of Aggregation in Cell Signaling

Closing Remarks

Jay Berzofsky, M.D., Ph.D.**Engineering Vaccines for Cancer and HIV**

Most classic antiviral vaccines attempt to mimic the natural infection with an attenuated, inactivated, or subunit vaccine. However, viruses causing chronic infection, such as HIV or hepatitis C virus, or cancer, do not induce sufficient immunity to eradicate the infection or tumor, so a vaccine must induce greater immunity than does the disease itself. We are developing strategies to improve such engineered vaccines, focusing primarily on induction of cytotoxic T lymphocytes (CTL), that kill cancer cells or virus-infected cells and prevent the cancer or virus from spreading. These approaches include 1) epitope enhancement by sequence modification to improve binding of antigenic epitopes to Major Histocompatibility Complex (MHC) molecules that present antigen to T cells, 2) strategies to induce higher avidity CTL that are more effective at eradicating virus infections or cancer, 3) use of cytokines and costimulatory molecules to alter the quality as well as increase the quantity of the immune response and especially to increase CTL avidity and longevity, 4) targeting of the mucosal immune system for viruses that enter through a mucosal route or reside in the mucosa, as in the case of HIV, and 5) blockade of negative regulatory signals that dampen the immune response and inhibit tumor immunosurveillance. For the last of these, we have focused on a new regulatory axis of NKT cells that we identified. Combinations of these approaches may lead to the development of more effective vaccines for such diseases as HIV infection and cancer, in which traditional vaccine strategies have not yet been successful.

Dr. Jay Berzofsky was appointed Chief of the new Vaccine Branch, Center for Cancer Research, National Cancer Institute, in 2003, after being Chief of the Molecular Immunogenetics and Vaccine Research Section, Metabolism Branch, National Cancer Institute, NIH, since 1987. He graduated Summa cum Laude from Harvard (1967), and received a Ph.D. and M.D. from Albert Einstein College of Medicine. After interning at Massachusetts General Hospital, he joined NIH in 1974. Dr. Berzofsky's research has focused on antigen processing and presentation by MHC molecules, the structure of antigenic determinants, cytokine and regulatory cell control of T cell function and avidity, and translation to the design of vaccines for AIDS, malaria, cancer, and viruses causing cancer. He has over 390 scientific publications. Dr. Berzofsky has received a number of awards, including the U.S. Public Health Service Superior Service Award, the 31st Michael Heidelberger Award, the McLaughlin Visiting Professorship, the Australasian Society for Immunology Visiting Lectureship, and the Tadeusz J. Wiktor Memorial Lectureship. He is past President of the American Society for Clinical Investigation, and a Fellow of the American Association for the Advancement of Science, and was elected Distinguished Alumnus of the Year for 2007 by the Albert Einstein College of Medicine. He was also elected Chair-Elect of the Medical Sciences Section of the American Association for the Advancement of Science (AAAS). Drs. Berzofsky and DeLisi have been friends, colleagues and frequent collaborators for more than two decades.

Gyan Bhanot, Ph.D.

Gyan Bhanot received his PhD in theoretical Physics from Cornell, and did post-doctoral research with Claudio Rebbe. He subsequently moved to Thinking Machines, then to IBM and is now a Professor - BioMaPS and Biomedical Engineering, Rutgers having made substantial contributions to biomedical science during the past 5 years. His lab is currently focused on two research areas. The first is computational biology related to cancer. Specifically, he uses microarray, mass spec and SNP polymorphism data to identify, quantify and explain cancer initiation, progression, and metastasis. The goal is to develop protocols for clinical use, participate in clinical trials and identify therapeutic strategies. This research involves collaborations with local clinicians and researchers at BME, CINJ, UMDNJ, CABM and BioMaPS as well as collaborators from the Institute for Advanced Study, Princeton, Broad Institute, IBM Research, Tokyo Gerontological Institute and Boston University.

His second focus area is Evolutionary Genetics including human migration, phylogeny, disease association studies and patterns of mutations correlated with longevity and complex disease phenotypes.

Robert Brown, Ph.D.

Robert A. Brown, Ph.D., 10th president of Boston University, is a distinguished scholar of chemical engineering and an innovative leader in higher education. Dr. Brown, 54, a Texas native, earned a B.S. and an M.S. in chemical engineering at the University of Texas at Austin. He earned a Ph.D. at the University of Minnesota.

Prior to his appointment at Boston University, Dr. Brown was provost and Warren K. Lewis Professor of Engineering at the Massachusetts Institute of Technology. He is a member of the National Academy of Sciences, the American Academy of Arts and Sciences, the National Academy of Engineering, and numerous other prestigious professional societies.

Dr. Brown has published approximately 250 papers in areas related to mathematical modeling of phenomena associated with materials processing, fluid mechanics of viscoelastic fluids, interface morphology, and modeling of semiconductor processing. In February 2006, President George W. Bush appointed Dr. Brown to the President's Council of Advisors on Science and Technology (PCAST), a panel established to maintain a steady stream of expert advice from the private sector and the academic community on a wide range of scientific and technical matters.

David Campbell, Ph.D.

Dr. David K. Campbell was named Provost of Boston University in September 2005. He had served as provost ad interim since July 2004 and Dean of the Boston University College of Engineering since September 2000.

Campbell came to Boston University from the University of Illinois, where he served as professor and head of the Department of Physics. A theoretical physicist who specializes in nonlinear phenomena and condensed matter physics, he holds degrees from Harvard and Cambridge universities. He and Charles DeLisi arrived at Los Alamos National Laboratory at approximately the same time, and were colleagues in the Theoretical Division for 2 years. Dr. Campbell spent nearly 20 extraordinarily productive years at LANL, including eight years as director of its Center for Nonlinear Studies.

Charles Cantor, Ph.D.

Sensitive in vivo detection of specific RNA species

The use of gene fusions incorporating fluorescent proteins like EGFP has revolutionized our ability to study protein localization in living cells and organisms. Similar fusions to RNA binding proteins have been employed to localize cellular RNAs, but these have seen very limited applicability because of high background from unbound proteins. We have used protein complementation of split EGFP to pairs of RNA binding proteins to overcome this difficulty and now have a high contrast method for specifically labeling any RNA in any cell. Results with this system show unexpected behavior when specific RNA transcripts in E.coli are visualized.

Dr. Cantor received his A.B. in Chemistry from Columbia in 1963, and after receiving his Ph.D. from Berkeley, rejoined the Columbia Chemistry Department, where he was appointed full Professor in 1970. In 1982 he was appointed Chairman of Genetics and Development at the Columbia College of Physicians and Surgeons, and in 1987 became first Director of the Human Genome Center, U.C. Berkeley. He joined the BU College of Engineering in 1992, where he remains Professor of Biomedical Engineering. In 1998 he joined Sequenom as Chief Scientific Officer and Chairman of the Scientific Advisory Board. In May 2000, A world-renowned scientist, Dr. Cantor is a consultant to more than 16 biotech firms, has published more than 400 peer reviewed articles, been granted 54 US patents, and co-authored a three-volume textbook on Biophysical Chemistry. He published the first textbook on genomics entitled, Genomics: The Science and Technology of the Human Genome Project. He is recipient of numerous awards including election to the National Academy of Sciences in 1984. He has been a friend, colleague and collaborator of Dr. DeLisi for more than two decades.

James Collins, Ph.D.

Predictive Biology by Design

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. Theoretically, the biochemistry of the feedback loops associated with protein-DNA interactions often leads to nonlinear equations, and the tools of nonlinear analysis become invaluable. 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. In addition, we present integrated computational-experimental approaches that enable construction of first-order quantitative models of gene-protein regulatory networks using only steady-state 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: (1) to gain insight into the regulatory role of individual genes and proteins in the network, (2) to identify the pathways and gene products targeted by pharmaceutical compounds, and (3) to identify the genetic mediators of different diseases.

Dr. Collins is a Professor of Biomedical Engineering, University Professor and co-Director of Boston University's BioDynamics Center. Dr. Collins arrived at BU in 1990, after 3 years as a Rhodes scholar at Oxford. He has done pioneering research at almost every level of human physiology. His early research on noise-based sensory prostheses led to the development of vibrating gel shoe innersoles that have been shown to improve the balance of elderly people and those who have suffered from strokes and diabetes. His work on human locomotion has been recognized by numerous awards including the McArthur "Genius" Award and the MIT TR100 award. More recently Dr. Collins has been seminal to the development of a field that has come to be known as synthetic biology. His "genetic toggle switch" promises to create programmable cells to fight disease or warn against environmental toxins, and his work in systems biology may lead to a new understanding of the causes of cancer.

Micah Dembo, Ph.D.**Dynamics Of Cellular Traction Forces**

Cells exert forces on their external environment for a variety of reasons; for example to crawl, to rearrange than deform the extracellular matrix, and to send and receive information. A longstanding goal of cell biology is to measure and quantify these very tiny forces, in situ, while perturbing the cell as little as possible.

I will describe a methodology we have developed for reaching this goal.

It involves plating a cell on an elastic material of known mechanical properties, observing the way the substrate deforms, and then utilizing detailed knowledge of substrate mechanics as a basis for deducing the exact magnitude and placement of the cellular traction forces. This last step typically requires the solution of an ill-conditioned integral equation subject to various nonlinear constraints. We will describe the way this problem can be solved numerically and we will then go on to discuss a few practical applications that combine this method with other more classical techniques to yield new insights into the mechanics and control of cell motion and into the molecular biology, physiology and pathology, of force production at the cellular level.

Dr. Dembo received his doctorate in Biomathematics from Cornell University. He has over 75 journal publications to his credit, in addition to the numerous books to which he has contributed chapters and a video. He served on the editorial board for Biophysical Journal for six years and was the associate editor for Comments on Theoretical Biology for another six years. He lectures both nationally and internationally. He currently serves as the Director for the Whitaker Center of Computational Biology at Boston University in addition to being a professor in the Department of Biomedical Engineering at Boston University. He is also a member of the American Institute for Medical and Biological Engineering. Drs. Dembo and DeLisi have been colleagues for 3 decades, sharing an early interest in the cellular mechanisms underlying desensitization and activation of the immune system in response to allergens.

Dr. Dembo's current research is in the area of cellular mechanics. His work includes measurement of the forces and stresses produced by individual cells during active locomotion, finite element modeling of cytokinesis, finite element modeling of amoeboid locomotion, and the role of the cytoskeleton in oncogenic transformation.

Byron Goldstein, Ph.D.**The Ubiquitous Role of Aggregation in Cell Signaling**

Ligand-induced receptor aggregation is a well-known mechanism for the initiation of intracellular signals, but recent evidence also indicates that oligomerization of signaling molecules not directly associated with the receptor may be required for signal propagation. We review some of the models of ligand-induced receptor aggregation and then look at the early events initiated by the aggregation of immune response receptors. We discuss how the oligomerization of a key scaffolding protein in T cell and mast cell signaling arises and the possibility of "valence switching" as a way to cause a dramatic increase in oligomerization.

Dr. Byron Goldstein received his Ph.D. in Physics from New York University, where he and Charles DeLisi met and published their first paper together. From 1970-1975 he was associate Professor of Physics at Farleigh Dickenson University, and since 1975 he has been in the Theoretical Division at Los Alamos National Laboratory, arriving there about the same time as David Campbell. Dr. Goldstein is currently Laboratory Fellow, a title conferred by LANL only rarely. His research is focused on receptors that play key roles in the immune response. He is currently working on building a detailed predictive model of the early cell signaling events mediated by the high affinity receptor for IgE, a major participant in allergic reactions of the immediate type.

Minoru Kanehisa, Ph.D.**Linking Genomes to Biological Systems and Environments**

Since the completion of the Human Genome Project, high-throughput experimental projects have been initiated for uncovering genomic information in an extended sense, including transcriptome and proteome information, as well as metabolome, glycome, and other genome-encoded information. Together with traditional genome sequencing for an increasing number of organisms from bacteria to higher eukaryotes, we are beginning to understand the genomic space of possible genes and proteins that make up the biological system. In contrast, we have very limited knowledge about the chemical space of possible chemical substances that exists as an interface between the biological world and the natural world. Chemical genomics is an emerging discipline for systematic analysis of the chemical space. Experimentally, this is being achieved by high-throughput screening of large-scale chemical compound libraries with various biological assays at the molecular, cellular, and organism levels. In order to best utilize the bioassay data being generated, bioinformatics methods have to be developed to extract the information encoded in the chemical structures, and to understand the information in the context of molecular interactions and reactions involving proteins and other biomolecules. This would eventually lead to basic understanding of the chemical environment that interacts with and drives evolution of the biological system.

KEGG (<http://www.genome.jp/kegg/>) is a database of biological systems, consisting of genetic building blocks of genes and proteins (KEGG GENES), chemical building blocks of both endogenous and exogenous substances (KEGG LIGAND), molecular wiring diagrams of interaction and reaction networks (KEGG PATHWAY), and hierarchies and relationships of various biological objects (KEGG BRITE). KEGG provides a reference knowledge base for linking genomes to biological systems by the process of network reconstruction, which is to map, for example, a genomic or transcriptomic content of genes to KEGG reference pathways to infer higher order functions of the cell or the organism. In addition, KEGG now provides a reference knowledge base for linking biological systems and environments, such as for analysis of drug-target relationships and integrated analysis of genetic and environmental factors of diseases. This is based on the new development of KEGG BRITE, a collection of hierarchically structured vocabularies representing our knowledge on various aspects of biological systems. In contrast to KEGG PATHWAY, which is limited to molecular interactions and reactions, KEGG BRITE incorporates many different types of relationships involving, for example, cells, tissues, organs, diseases, and drugs. Thus, the mapping of genomic data to KEGG BRITE supplements the KEGG PATHWAY mapping, both in terms of the extended repertoire of gene/protein families and the links made from genes/proteins to other biological objects.

Dr. Kanehisa is professor in the Institute for Chemical Research and director of the Bioinformatics Center at Kyoto University. He has worked in the Johns Hopkins University School of Medicine, the Los Alamos National Laboratory (where he was one of the cofounders of GenBank), and in the National Cancer Institute of the National Institutes of Health, where he and Dr. DeLisi collaborated for several years. His other activities include: concurrent professorship at the Human Genome Center, Institute of Medical Science, University of Tokyo (1991-1995 and 2002-Present), principal investigator with the Japanese Genome Informatics Project (1991-2000), and president of the Japanese Society for Bioinformatics (1999-Present). Prof. Kanehisa has a D.Sc. in physics from the University of Tokyo (1976), and has been a collaborator of Dr. DeLisi for more than 2 decades.

Douglas Lauffenburger, Ph.D.

Multivariate Cue/Signal/Response Analysis of Cell Decision Processes

Cell behavioral functions are governed by biomolecular networks which translate stimulatory cues (e.g., ligand/receptor binding interactions, mechanical stresses, physical or chemical insults) into intracellular signals which then influence transcriptional and post-transcriptional, metabolic, and cytoskeletal processes that integrate to effect proximal and ultimate cell responses. While there is an increasingly effective body of work enhancing our understanding of how intracellular signals are generated by stimulatory cues, an exceptionally difficult challenge is to understand how these signals operate to influence cell behavioral responses. We are undertaking an effort address this question by means of a combination of quantitative, dynamic protein-centric experimental manipulations and measurement with a spectrum of computational mining and modeling approaches. Particular application problems include cell migration, differentiation, and death. This talk will present some overview and a specific example vignette outlining our efforts.

Professor Lauffenburger received his Ph.D. in Chemical Engineering from University of Minnesota in 1974, and was appointed assistant professor at U Penn in 1975. He is currently Uncas & Helen Whitaker Professor of Bioengineering in the Biological Engineering [BE] Division, Biology Department, and Chemical Engineering Department, and is a Member of the Center for Cancer Research, Center for Biomedical Engineering, and Biotechnology Process Engineering Center at MIT. He serves as Director of the BE Division and on the Executive Committee of the MIT Computational & Systems Biology Initiative. Dr. Lauffenburger's major research interests are in receptor-mediated cell communication and intracellular signal transduction for cell cue/signal/response relationships important in pathophysiology with application to drug discovery and development. Professor Lauffenburger has served as a consultant or scientific advisory board member for numerous pharmaceutical and biotechnology companies. He is a member of the National Academy of Engineering and of the American Academy of Arts & Sciences, and has served as President of the Biomedical Engineering Society and on the Advisory Council for the National Institute for General Medical Sciences at the NIH. Drs. Lauffenburger and DeLisi have been friends and colleagues for more than 3 decades, sharing an interest in cell biophysics, and collaborating on an early paper in the field.

Hanah Margalit, Ph.D.

Role of non-coding RNA in the cellular regulatory networks

The importance of small non-coding RNA molecules in genetic regulatory networks has recently been recognized both in pro- and eukaryotes. Small non-coding RNAs provide another layer of post-transcriptional regulation in addition to that mediated by proteins. It is intriguing to examine the cellular functions of these molecules and their integration with other regulatory mechanisms. Using dynamical simulations we compare the properties of regulation by small RNAs (sRNAs) in bacteria with those of transcriptional regulation and protein-protein interaction. We show that regulation by sRNAs provides an effective regulation mechanism, which can regulate a few dozens of genes at a time. Regulation of a larger number of genes results in weakening of the regulation. The response time of sRNA regulation is faster than in transcriptional regulation and most often slower than in protein-protein interaction. We identify mixed regulatory patterns, such as feed-forward loops consisting of both transcriptional regulation and post-transcriptional regulation by sRNAs. Our analysis demonstrates that such feed-forward loops guarantee effective shutdown of the target gene, suitable for fast responses that are needed under changing environments.

Dr. Margalit received her doctorate from The Hebrew University of Jerusalem and moved to NIH in 1985 to join Charles DeLisi's research group. In 1990 she returned to Israel as an Assistant Professor at Hadassah Medical School in Jerusalem, and is currently a full professor at that institution. She has been a prolific contributor to immunology and bioinformatics. Her current work focuses on molecular control and recognition mechanisms, genome-wide analysis of regulatory elements, and sequence-structure-function relationship in nucleic acids and proteins.

Jill Mesirov, Ph.D.

Knowledge-based Paradigms for Computational Genomics

DNA microarrays now make it possible to capture the expression pattern of all the genes in the genome in a single experiment. Genome-wide expression analysis is at the heart of global genomic approaches to biomedical research and appears in approximately 6000 published papers a year. The challenge that now faces us is not obtaining these molecular profiles, but interpreting them to gain a better understanding of underlying biological processes. We will describe how prior biological knowledge can be incorporated into robust, reproducible, quantitative approaches to analyzing mRNA profile data from cancer and infectious disease.

Dr. Mesirov is the Associate Director, Chief Informatics Officer & Director of Bioinformatics & Computational Biology Programs, The Eli and Edythe L. Broad Institute, Massachusetts Institute of Technology and Harvard University. She is also Adjunct Professor of Bioinformatics at Boston University. Mesirov spent many years working in the area of high performance computing and developing parallel algorithms relevant to problems which arise in science, engineering, and business applications. Her current research interest is the study and development of algorithms for computational biology in such areas as molecular pattern recognition and discovery for cancer genomics, genome analysis and interpretation, protein structure prediction and classification, molecular dynamics, and inverse protein folding. Mesirov received her Ph.D. in Mathematics from Brandeis University. Mesirov came to the Whitehead in 1992 from IBM where she was Manager of Computational Biology and Bioinformatics in the Healthcare/Pharmaceutical Solutions Organization. Before joining IBM, she was Director of Research at Thinking Machines Corporation for ten years. She has also held positions as lecturer at the University of California at Berkeley, research mathematician at IDA's Center for Communications Research in Princeton, and Associate Executive Director of the American Mathematical Society.

Matthew Pincus, M.D., Ph.D.

Molecular Modeling of Anti-Cancer Peptides that Block Cancer but not Normal Cell Growth

Oncogenic ras-p21 protein induces cell transformation and is a major causative factor of about one-third of all common human cancers. It differs from its wild-type protein by containing arbitrary single amino acid substitutions at critical positions in the amino acid sequence such as at Gly 12 or Gln 61. These substitutions induce changes in the three-dimensional structure of the ras-p21 protein. To identify the domains of p21 that undergo these changes, we have used molecular dynamics based on AMBER potentials and the electrostatically-driven Monte Carlo (EDMC) method based on ECEPP to compute the average structures of oncogenic and wild-type ras-p21. Superposition of these structures reveals domains that differ in their three-dimensional structures, i.e., residues 35-47, 55-71, 81-93, 96-110, 115-126, 122-138. We find that these changes occur independently of the method employed in phase space sampling. We have synthesized peptides corresponding to these sequences and have tested them in *Xenopus laevis* (frog) oocytes. Oncogenic ras-p21 induces oocytes to undergo maturation as does insulin, which activates endogenous wild-type ras-p21. We find three peptides, residues 35-47 (PNC-7), 96-110 (PNC-2) and 115-126 (PNC-1) that block only oocyte maturation induced by oncogenic ras-p21. This suggests that these peptides selectively block mitogenesis induced by oncogenic ras-p21. We have therefore proceeded to test these peptides on ras-transformed cancer cell lines. In these experiments PNC-2 and PNC-7 were linked to a penetratin peptide sequence that allows the peptide to be transferred across the cell membrane. We find that these peptides induce *phenotypic reversion* of TUC-3 pancreatic cancer cells but have no effect on the corresponding untransformed pancreatic acinar BMRPA1 cell line. These peptides likewise induce phenotypic reversion of HT1080 fibrosarcoma and SW480 colon cancer cells. They induce necrosis of MIAPaCa-2 pancreatic cancer and U251 astrocytoma cell lines. These peptides are therefore promising as possible agents that can block cancer growth in humans without affecting normal cell growth.

Using similar modeling techniques we have identified and synthesized peptides from the *MDM-2-binding domain of the p53 protein*, i.e., residues 12-26, 17-26 and 12-20), each attached to a penetratin sequence: p53 12-26-penetratin, p53 17-26-penetratin and p53 12-20-penetratin are called PNC-27, 28 and 21, respectively. These peptides induce tumor cell necrosis, not apoptosis, of a wide variety of human cancer cells, including pancreatic, colon, breast and cervical. Remarkably, they have no effect on the growth of normal cells such as BMRPA1 pancreatic acinar cells, normal breast epithelial cells (MCF-10-2A) and human stem cells from cord blood. We have tested PNC-28 in nude mice and found that it, but not control peptides, induces tumor cell necrosis but has no toxic effects on the mice over a 45-day observation period. These peptides appear to be excellent candidates, therefore, for clinical trials.

Dr. Pincus received his M.D. from Downstate Medical School in his beloved borough of Brooklyn, and his Ph.D. from Cornell University in Chemistry where he worked with the renowned biophysical chemist Harold Scheraga who also grew up in NYC. Dr. Pincus has been a pioneer in the application of computational structural biology to problems of clinical importance. From 1979 -1983 he was senior staff Medical Fellow at NCI where he collaborated with Drs. Blumenthal, Weinstein, DeLisi and Klausner. He is currently Professor, Department of Pathology, State University of New York Health Sciences Center at Brooklyn; Chairman, Department of Pathology and Laboratory Medicine, New York Harbor Veterans Affairs Health Care System, Brooklyn, NY. As one of the Nation's most distinguished Pathologists, he is co-editor of the standard reference book in the field, Henry's Clinical Diagnosis and Management by Laboratory Methods, 21st edition.

Chris Sander

Biomolecular networks: representation, perturbation and function

The analysis of biomolecular networks and computational predictions are cornerstones of systems biology. I will cover the database representation of pathway models, experimental-theoretical perturbation analysis of biomolecular networks and a potential application of functional network theory to cancer therapy.

Dr. Sander is Head of the Computational Biology Center at Memorial Sloan Kettering Cancer Center and tri-institutional professor at Rockefeller and Cornell Universities, and is internationally acknowledged as a founder of computational biology. His principal research interests are in computational and systems biology, including predictive simulations of biological processes, integrated molecular profiling of disease states, gene regulation by small RNAs and structural genomics. He is a leader in community efforts to create an open-source information resource for biological pathways. Previously, Dr. Sander served as Chief Information Science Officer with Millennium Pharmaceuticals, as Senior Scientist at the European Bioinformatics Institute in Cambridge, England, as founding chair of the Department of Biocomputing at the European Molecular Biology Laboratory and researcher at the Max-Planck Institute for Medical Research in Heidelberg. He is editor of Bioinformatics, a leading journal in computational biology, and currently an advisor to the Protein Structure Initiative of the National Institutes of Health and the IBM Deep Computing Initiative. Dr. Sander was trained at the universities of Berlin, Berkeley, Copenhagen and Stony Brook and his Ph.D. is in theoretical physics. Drs. Sander, Kanehisa and DeLisi have been colleagues for more than 2 decades and co-led one of the first Human Frontier Projects in the late 80s.

Boris Shakhnovich, Ph.D.

Selective Constraints in Evolution of Gene Families and TBP-Dependent Promoters

This talk is divided into two related parts. First, I explore the origins and impacts of constraints in evolution of gene families. Surprisingly, I find that selection is homogeneous for members of the same paralogous gene family. Thus, I partition gene families in several genomes into two classes: those that include at least one essential gene (E-families) and those without essential genes (N-families). I find that weaker purifying selection causes N-families to evolve in a more dynamic regime with higher rates both of duplicate fixation and pseudogenization. Because genes in E-families are subject to significantly stronger purifying selection than those in N-families, they survive longer and exhibit much greater sequence diversity. Longer average survival time also allows for divergence of upstream regulatory regions, resulting in larger changes of transcriptional control mechanisms among paralogs in E-families. Second, I explore the determinants of promoter architecture that differentiate promoters with a TATA-box that are regulated by the TATA Binding Protein (TBP) from promoters that have the TATA sequence but are not regulated by TBP. I show that some global properties of the extended promoter such as length and G+C content correlate with activity of TBP. I then discuss a physical model of TBP function in terms of efficiency in localization of the Pre-initiation complex. Finally, I find that there is differential selection on promoter sequences in light of the physical constraints imposed by the TBP. If time permits, I will discuss a putative relationship between selection in the evolution of gene families and TATA promoters.

Dr. Shakhnovich received his degree in Bioinformatics from Boston University, working with Charles DeLisi. He is currently a Postdoctoral Associate in the Department of Molecular and Cellular Biology at Harvard University. He has contributed to several journal publications in addition to multiple talks and poster presentations he has given. He has also organized the first student invitational series of lectures at the Boston University Bioinformatics Department.

Dr. Shakhnovich's research focuses on molecular evolution, in particular the co-evolution of structure and function. The research also includes the molecular evolution of genes. He has begun preliminary studies in the area of identification and elucidation of cellular control mechanisms.

Avrum Spira, M.D.**Developing Airway Epithelial Gene Expression Biomarkers for Early Lung Cancer Detection**

Lung cancer is the leading cause of cancer death in the US and the world. While cigarette smoking is the major cause of lung cancer, only 10-20% of heavy smokers develop the disease. Lung cancer is most often diagnosed at a late stage (with a resulting five year mortality of 80-85%) because of the inability to identify current or former smokers who are at risk for developing the disease and the lack of sensitive biomarkers that can be used for early diagnosis. Based on the premise that cigarette smoke creates a field of injury in all airway epithelial cells, we have developed profiles of gene expression in histologically normal bronchial airway epithelial cells that can serve as a sensitive and specific diagnostic for lung cancer in smokers. This gene expression biomarker exceeds the diagnostic yield of standard bronchoscopy tests and is effective in diagnosing lung cancer at an early and potentially curable stage. We are currently extending our studies to epithelial cells that can be obtained in a non-invasive fashion from the upper respiratory tract and potentially serve as an effective screening tool for lung cancer. These tools are destined to impact patient care and are the direct consequence of Dr. Charles DeLisi's vision for developing and applying post-genomic technologies to clinical research and the practice of medicine.

Dr. Spira received his M.D. at McGill University and his M.Sc. in Bioinformatics at Boston University where he collaborated with Professor DeLisi. Dr. Spira is Assistant Professor in the Departments of Medicine, and Pathology and Laboratory Medicine, and Assistant Professor of Bioinformatics. He attends in the Medical Intensive Care Unit and on the Pulmonary/Critical Care Interventional service at Boston Medical Center. Dr. Spira's laboratory research interests focus on applying high-throughput genomic and bioinformatic tools to the translational study of lung cancer and emphysema, and COPD. He is funded by the Doris Duke Charitable Foundation and is the only recipient of that highly select award in BUMC history.

John N. Weinstein, M.D., Ph.D.**Integromic Molecular Profiling in Cancer Pharmacology and Therapeutics**

When it comes to understanding cancer cells and their pharmacology, molecular profiles – generated, for example, using microarrays -- are having a profound impact. Transcript expression, protein expression, DNA copy number, DNA methylation, chromosomal aberrations, microRNA expression, and other molecular characteristics, analyzed individually, all play their parts. But, to cite a well-worn metaphor, examining the elephant from many different perspectives can lend new insight into the beast. In that spirit, according to the 'integromic hypothesis' (1), synergies (in the non-mathematical sense of the term) are to be had by integrative analysis of cancer cells from many perspectives at once. To pursue the integrative program, we've focused on the panel of 60 human cancer cell lines (the NCI-60) used by the NCI Developmental Therapeutics Program to screen >100,000 compounds for anticancer activity since 1990. We and our many collaborators are using >25 different microarray platforms and other high-throughput technologies for integromic characterization of the NCI-60 and/or derivatives of them at the DNA, RNA, protein, functional, pharmacological, and chromosomal levels. In tandem with that experimental enterprise, we've been developing a set of web-based computational tools, the 'Miner Suite', for analysis and integration of the data. Included, among others, are programs for navigating the forest of gene and protein identifiers (MatchMiner), leveraging the Gene Ontology (GoMiner and High-Throughput GoMiner), incorporating splice variant information (SpliceMiner), organizing and querying the molecular profile databases and metadatabases (CellMiner), and navigating Kohn Molecular Interaction Maps (MIMminer), and creating clustered heat maps (CIMminer) -- which we introduced in the mid-1990's (2). The Miner suite tools and databases are freely available at <http://discover.nci.nih.gov>.

Our proofs of principle to date for 'integromic' analysis in the context of cancer pharmacology include: (i) identification of candidate biomarkers for differential diagnosis of ovarian and colon metastases (3); (ii) analyses that led to clinical development of oxaliplatin, now standard-of-care for colon cancer; (iii) discovery of "MDR1-inverse" compounds, which, paradoxically, are more potent in cells that express MDR1; (iv) development of the "Permissive Apoptosis-Resistance" (PAR) model for acquired resistance to cancer drugs; (v) identification of asparagine synthetase (ASNS) as a potential 'causal' biomarker for treatment of ovarian cancers with L-asparaginase (L-ASP) (4). In addition to those specific biomedical outcomes, we think of our multi-faced profiling of the NCI-60 as providing grist for the mill of integrative systems biology (5). With thanks to our many other collaborators in the overall enterprise.

References: (1) Weinstein & Pommier. *Comptes Rendus Biology* 2003; 326:909. (2) Weinstein, et al. *Science* 1997; 275:343. (3) Ludwig and Weinstein, *Nature Rev. Cancer* 2005; 5: 845. (4) Lorenzi, et al., *ibid* 2006; 5: 2613. (5) Weinstein, *Mol. Cancer Ther.* 2006; 5: 2601.

Dr. Weinstein has a B.A. in Biology at Harvard College, then an M.D. and a Ph.D. in Biophysics at Harvard University. His 230 publications include 10 as first author in Science. He presents 30 to 40 lectures a year nationally and internationally and was recently nominated for the Medal of Technology as a pioneer of the "postgenomic era" in biomedical science. After an internship and residency in Medicine at Stanford he joined the NIH, and currently heads the Genomics & Bioinformatics Group in the NCI's Laboratory of Molecular Pharmacology. He is a Captain (retired), U.S. Public Health Service. Dr. Weinstein founded and heads the NCI Bioinformatics, Biostatistics, and Computational Biology Faculty.

His research initially focused on the development of novel approaches to therapy of cancer and AIDS using liposomes, monoclonal antibodies, cytokines, and other biologicals. Since 1992, he has been applying a mix of genomic, proteomic, bioinformatic, and computational chemistry tools to the pursuit of new biomarkers and therapeutic strategies for cancer. He is also an adjunct Professor in BU's Bioinformatics Program, and serves as a research sponsor and mentor of the Programs doctoral students. He and Dr. DeLisi are former racquetball partners, and friends for nearly three decades.

Zhiping Weng, Ph.D.

Transcription Factor Binding and Modified Histones in Human Bidirectional Promoters

Bidirectional promoters have received considerable attention because of their ability to regulate simultaneously two downstream genes (divergent genes). They are also highly abundant, directing the transcription of approximately 11% of genes in the human genome. We categorized the presence of DNA sequence motifs, binding of transcription factors, and modified histones as overrepresented, shared, or underrepresented in bidirectional promoters with respect to unidirectional promoters. We found that a small set of motifs, including GABP, cMyc, E2F1, E2F4, Nrf-1, CCAAT, YY1, and ACTACAnnTCC, are overrepresented in bidirectional promoters, while the majority (73%) of known vertebrate motifs are underrepresented. We performed chromatin-immunoprecipitation (ChIP) followed by quantitative PCR for GABP on 118 regions in the human genome and showed that it binds to bidirectional promoters more frequently than unidirectional promoters and its position specific scoring matrix is highly predictive of binding. Categorization based on computational sequence analysis are supported by categorization based on transcription factor binding in living cells, as detected by ChIP assays on the ENCODE regions. Signatures of active transcription, such as occupancy of RNA polymerase II and the modified histones H3K4me2, H3K4me3, and H3ac, are overrepresented in regions around bidirectional promoters, suggesting that a higher fraction of divergent genes are transcribed in a given cell than the fraction of other genes. Accordingly, analysis of wholegenome microarray data indicates that 68% of divergent genes are transcribed compared with 44% of all human genes. This analysis also reveals that divergent genes of different bidirectional promoters have more correlated expression patterns than genes of unidirectional promoters, and divergent gene pairs of the same bidirectional promoters have the most correlated expression patterns. Examination of the occupancy of the modified histones around all bidirectional promoters in the ENCODE regions indicates that divergent gene pairs tend to be transcribed concurrently. By combining the analysis of publicly available ENCODE data and a detailed study of GABP, we survey bidirectional promoters with breadth and depth, leading to biological insights concerning their motif composition and bidirectional regulatory mode.

Dr. Weng received her Ph.D. in Biomedical Engineering from Boston University (1997) working with Charles DeLisi. She is currently Associate Professor of Biomedical Engineering, Bioinformatics, and Pharmacology. In addition to having published over 75 journal articles, she is a reviewer for nine different journals, and is a panelist for both the NSF and the NIH. She has won several awards including the Professional Opportunities for Women Award and the CAREER Award, both from the NSF. Dr. Weng's diverse areas of research include DNA and protein sequence analysis, protein-protein interactions, protein structure analysis, optimization algorithms and their applications in molecular biology, and drug and vaccine design.

Itai Yanai, Ph.D.

Chance and necessity in the manifestation of genetic programs

A gene's pattern of expression is generally assumed to correlate to its function, yet recent work has called this fundamental assumption into question. If a fraction of gene expression corresponds to non-functional regulation, how can it be distinguished from the functional component? From an evolutionary perspective, if a gene's expression profile in a specific context is under selection it is expected to be conserved throughout evolution, whereas free of selection, it may change and adopt different expression profiles. Thus, comparisons of expression profiles in related species can identify which expression profiles are under selection and likely to be functional. This approach is similar to comparing homologous gene sequences, where conservation is generally interpreted as separating the functional from the non-functional domains. Nematodes constitute a fitting test-bed for this approach since the early embryonic development of *C. elegans* is indistinguishable from that of *C. briggsae*, yet their genomes are roughly as distant as the human and mouse genomes. We conducted an embryonic time-course in both organisms, using a custom whole-genome microarray for each. We detected >10,000 *C. elegans* genes with significant temporal profiles and found that only 65% of these are conserved in *C. briggsae*. Strikingly, genes whose profile evolved are dramatically less likely to produce lethality when disrupted than genes with conserved profiles, suggesting that their embryonic expression is non-functional. These results indicate that expression comparisons over multiple species and developmental stages will identify the functional component of expression.

Dr. Yanai received his Ph.D. in Bioinformatics from Boston University, working with Charles DeLisi, and was the first student to receive a Ph.D. from the Program. He subsequently became Daniel Koshland Fellow at the Weizmann Institute of Science and is now completing a post-doctoral position as Ruth Kerstein Fellow in Biological Sciences at Harvard. He has developed and instructed courses at both the Weizmann Institute of Science and Northeastern University. His research interests include developmental genetics, the evolution of gene expression, and the evolution of gene function.