

Science Symposium Abstracts
As of June 7, 2006

TRACK I

GENOMICS / PROTEOMICS

Opportunities and Obstacles in Future Genomics

Charles Cantor, Ph.D. Chief Scientific Officer, SEQUENOM

Abstract:

The cost of both the analysis of genomes and the synthesis of new genome elements is decreasing exponentially. This has profound implications for our future. On the positive side, we will be able to reduce the cost of health care by using genomic information to develop non-invasive methods of early disease discovery, progression, and therapeutic response. These same methods promise to make clinical trials much more cost effective. Examples will be shown for prenatal diagnosis and cancer. We will also increasingly gain the ability to accurately predict the behavior of cells and organisms, based on their genomes, and this too has profound implications for future healthcare. Examples will be shown for simple biological circuits in living cells.

However, these same advances are going to facilitate the ability to modify the properties of all of the organisms in our ecosystem including humans themselves. It is inevitable that this will happen, and it is essential that the focus on constructive applications and minimize the chances of destructive ones. Unfortunately, there is unlikely to be universal agreement on where to draw this line. Examples will be described in livestock management.

Analyze This and That: Genomes and Proteomes

Michael Snyder, Ph. D. Director, Yale Center for Genomics & Proteomics; Lewis B. Cullman Professor of Molecular & Cellular Biology; Professor of Molecular Biophysics & Biochemistry, Yale University

Abstract: We have used microarray technology to identify functional elements in the human genome. DNA tiling micorarray have been used to identify transcribed regions, transcription factor binding sites and copy number variation sin humans. We have also used microarray methods to map transcription regulatory and phosphorylation networks in yeast. Extensive analysis of transcription factor binding sites in yeast has revealed a regulatory cascade controlling pseudohyphal growth and identified key regulators within the cascade. Protein microarrays have been used to globally map phosphorylation networks in yeast. Integration of transcription factor binding, protein-protein and phosphorylation networks has been performed to form a meta-network. Searches within this network reveal common regulatory signatures used by eucaryotes. Overall, these studies provide significant information concerning basic principles underlying the organization of eucaryotic genomes and eucaryotic regulatory networks.

Widespread Operons in the C. Elegans Genome

Thomas Blumenthal, Ph.D. Professor & Chairman, Department of Biochemistry & Molecular Genetics, University of Colorado Health Sciences Center

Abstract: Several animal phyla, including nematodes, flatworms, cnidaria, and some chordates, have evolved ways to co-transcribe neighboring genes, producing multi-gene pre-mRNAs that get processed to make individual mRNAs. This is best understood in the nematode, *C. elegans*, where >15 % of genes are cotranscribed in operons, ranging from 2 – 8 genes long. These operons tend to contain genes with certain kinds of functions, but to exclude other types almost entirely. They also often result in cotranscription of genes of related function. The selection pressure for a small genome and for coregulation of genes appears to be sufficient for evolution of operons, as long as the mechanisms to process the resulting multigene pre-mRNAs exist. Both conventional 3' end cleavage and trans-splicing occur between closely spaced genes in operons. In studying the mechanism of polycistronic pre-mRNA processing we have gained insight into the evolutionary history of the molecules involved. We have studied the functions of novel proteins as well as novel RNAs responsible for 3' end cleavage and trans-splicing. There appears to be an intricate cycle of RNA/protein interactions required for catalysis of the reactions and perhaps for re-use of the proteins allowing subsequent rounds of processing.

Genomics-Based Insights into Pathogen Biology and Evolution

Claire M. Fraser-Liggett, Ph.D. President and Director, The Institute for Genomic Research

Abstract: No other field of research has embraced and applied genomic technology more than the field of microbiology. Microbial genomics has a broad range of applications, from understanding basic biological processes, host-pathogen interactions and protein-protein interactions to discover DNA variations that can be used in genotyping or forensic analyses to the design of novel anti-microbial compounds and vaccines. Comparative analysis of nearly 300 microbial species has demonstrated that the microbial genome is a dynamic entity shaped by multiple forces. Comparisons between closely related species and strains of pathogens have also begun to allow for correlations between genotype and phenotype. Metagenomics approaches are allowing us to begin to probe complex microbial communities for the first time and hold great promise in helping to unravel the relationships between microbial species. Lastly, the taxonomic potential of genomics is emerging, as studies of multiple strains allow us to revise and refine the bacterial species concept as well as the idea of a static genome.

Intimate Encounters: Frontiers at the Human-Microbial Interface

David A. Relman, M.D. Director, Proteomics/Genomics Core at the Digestive Disease Center, Stanford University School of Medicine (Stanford, CA); Chief, Infectious Diseases Section, Veteran's Affairs Palo Alto Health Care System (Palo Alto, CA)

Abstract: Humans are home to multiple diverse microbial communities. These communities play integral roles in guiding human development, maturation of immune defenses, protection against potential microbial invaders, maintenance of mucosal barrier functions, and regulation of human nutrition and fat metabolism. At the same time, disturbances in these microbial communities contribute to important human disease and suffering. Basic features of the human microbial ecosystem remain poorly described. Genomic tools and approaches have enabled a more detailed description of host-microbe encounters, the composition and function of these indigenous microbial communities, and shed light on fundamentally important processes, including the cellular responses associated with infection. Genome-wide transcript-abundance

profiles, like other comprehensive molecular readouts of host physiological state, provide a detailed blueprint of the host-pathogen dialogue during microbial disease, and can reveal functional gene-based modules associated with mechanisms of virulence and host defense. Yet, the highly dynamic and compartmentalized aspects of the host response to pathogens complicates efforts to identify predictive signatures for infectious diseases. Early explorations in some of these areas indicate the potential feasibility of this approach but also point to important unmet challenges.

TRACK II NEUROSCIENCE

Signal Transduction Pathways Used by Therapeutic Agents and Drugs of Abuse

Paul Greengard, Ph.D. Director, Laboratory of Molecular and Cellular Science, Vincent Astor Professor, Rockefeller University Nobel Laureate, 2000 Nobel Prize in Physiology or Medicine

Abstract: The concept that nerve cells communicate with each other through two distinct mechanisms, referred to as fast and slow synaptic transmission, is now well-established. A number of components of the two signal transduction pathways have been identified. Fast synaptic transmission occurs via activation by a neurotransmitter of a ligand-gated ion channel. In contrast slow synaptic transmission occurs via a signal transduction cascade that can be remarkably complex and that usually involves second messengers and/or protein phosphorylation/dephosphorylation reactions. A growing body of knowledge concerning slow signal transduction pathways has been utilized to elucidate the mechanism of action of therapeutic agents used for the treatment of schizophrenia, Parkinsonism, and depression, as well as of drugs of abuse, such as caffeine, cannabis, amphetamine, PCP, and LSD.

Addressing the Molecular and Cellular Complexities of the Mammalian Brain

Nathaniel Heintz, Ph.D. Director, Laboratory of Molecular Biology, James & Marilyn Simons Professor and Investigator, Howard Hughes Medical Institute

Abstract: In many instances, our understanding of mammalian physiology and disease is hampered by the molecular and histological diversities of tissues and organs. Perhaps the most extreme example of this type of complexity is the mammalian brain, which contains an unknown number of distinct cell types, expresses fully two thirds of the genome, and is characterized by a precise wiring diagram of interconnected neurons that is required for brain function. During the last decade, the Heintz laboratory has developed a variety of methods to gain genetic access to specific CNS cell types and circuits, to determine their anatomic and molecular properties, and to dissect their roles in mammalian CNS circuitry. This presentation will focus on the large-scale application of these approaches to advance our understanding of mammalian brain development, function and disease.

\Dissecting the Complex Genetic Etiology of Autism

Margaret Pericak-Vance, Ph.D. Director of the Center for Humans Genetics at Duke University Medical Center and James B. Duke Professor of Medicine and Chief of the Section of Medical Genetics at Duke University Medical Center

Abstract: Autism is a neurodevelopmental disorder characterized by impairments in reciprocal social interaction and communication accompanied by restricted and repetitive patterns of interest or behavior. With the improved surveillance and a broadening of the diagnostic criteria, the most recent prevalence study suggests that autism affects as many as 1 in 300 children in the US. Treatments are inadequate and provide little impact on the profound morbidity. Little is known about the etiology of autism, but there is a strong genetic component. Studies over the past decade have clearly shown that the underlying genetics are complex with several genes likely acting independently as well as interactively to significantly raise the risk for autism. With this realization the field of autism genetics arrives at a critical juncture. To move forward we must embrace new and creative paradigms to successfully dissect the genetic etiology of this disease. We need to expand and build on previous results embracing the paradigm that the wedding of new genomic technology with novel statistical methodology and unique study designs will bring about success. We are applying this approach to our large data set of autism patients and their families. Specifically, we have undertaken a SNP genome wide linkage screen in a unique series of large extended multigenerational autism families and have mapped a locus to chromosome 12. We have also identified a separate set of families that map to a region on chromosome 19 and have found significant association in the candidate region. Finally, we have hypothesized that autism is in part due to potential geneXgene interactions and have examined variations in the GABA receptor (GABAR) subunit genes for association and interactions to attempt to explain the spectrum of autism risk. The integration and results of this research will be used to address an important problem in childhood disease, the genetics of autism spectrum disorders.

Linking Stem Cells and Molecular Motors to Signaling and Neurodegenerative Disease

Lawrence S.B. Goldstein, Ph.D. Professor of Cellular and Molecular Medicine at University of California San Diego and Investigator, Howard Hughes Medical Institute

Abstract: Neurons are large highly polarized cells that depend upon long-distance microtubule-based transport between synapses and cell bodies to support signaling and neurotransmission. We have been using fruit flies, mice, and humans to study the nature of molecules that link molecular motor proteins to vesicular cargoes in axons and have discovered close relationships between these molecules, signaling, and neurodegenerative diseases. In my talk, I will discuss our recent research on basic mechanisms of movement and long-range signaling in neurons and how this information is leading to new and testable hypotheses about neurodegenerative diseases such as Lou Gehrig's Disease (ALS cmm-1433), Huntington's Disease, and Alzheimer's Disease. In addition, I will discuss how we are trying to use human embryonic stem cells to develop new systems for understanding, and eventually treating these disorders by generating new approaches for drug discovery or therapy. For example, we are currently working to generate human neurons from human embryonic stem cells that carry familial Alzheimer's Disease mutants to try and test several hypotheses of disease causation. In the case of ALS we have been working to identify which cells are at fault in this disease, and

have begun to develop methods that might allow stem cells to someday be used for therapeutic intervention.

Molecular Neurobiology of Alzheimer's Disease

Sangram S. Sisodia, Ph.D. Director, Center for Molecular Neurobiology, and Thomas Reynolds, Sr. Family Professor of Neurosciences, University of Chicago

Abstract: Alzheimer's disease (AD) a prevalent, adult-onset, neurodegenerative disease, is clinically characterized by progressive impairments in cognition and memory. These clinical features are accompanied by characteristic histological changes in the brain, including neuronal loss, extracellular deposition of fibrillogenic A β peptides in senile plaques and intracellular neurofibrillary tangles. The principal risk factors for AD are age and inheritance. Mutations in genes encoding amyloid precursor protein (APP), or presenilins (PS1 and PS2) cause autosomal dominant, familial Alzheimer's disease (FAD). While the function of APP in the nervous system is presently unclear, this membrane protein is subject to rapid anterograde transport and subject to proteolytic processing at, or near, terminal fields in the CNS. Genetic and biochemical studies have revealed that PS interacts with nicastrin, APH1 and PEN-2 in high molecular weight complexes. I will discuss the temporal assembly of these membrane proteins, the nature of subunit interactions and the enzymatic mechanism(s) by which the complex promotes intramembranous, "ε-secretase" processing of APP, Notch and several type 1 membrane proteins. In parallel, I will discuss gain- and loss-of function effects of mutant PS on axonal trafficking, neuronal vulnerability, neurogenesis, gene expression and APP/A β metabolism in transgenic mice. In summary, this lecture will provide a synopsis of current strategies to understand the function of APP and PS1 in the nervous system and the dysfunction of mutant variants in disease models.

Supported by NIA, Ruth K. Broad Biomedical Research Foundation, Ellison Medical Foundation, Alzheimer's Association, American Health Assistance Foundation and Edward Levi Fund

TRACK III

BIOMEDICAL INFORMATICS

Genome-Wide Identification of Modulators of Transcriptional Regulation in Human B Lymphocytes: The Kinome Meets the Transcriptome

Andrea Califano, Ph.D. Professor, Department of Biomedical Informatics, Columbia University; Co-Director, Center for Computational Biochemistry and Biosystems; Chief of Bioinformatics Division, Department of Biomedical Informatics; Director, Genome Center Bioinformatics

Abstract: Most transcriptional interactions in the cell are modulated by a variety of genetic and epigenetic mechanisms, resulting in transcriptional networks that are both dynamic and context-dependent. However, current efforts to reverse-engineer these circuits have provided representations that are mostly static and limited to lower eukaryotes. In this talk we discuss MINDY (Modulator Inference by Network Dynamics), a new general method for the genome-wide discovery of post-translational modulators of transcriptional regulation processes in mammalian cells from large scale microarray expression profile data. We first show that the method is effective in identifying signaling genes in the BCR pathway of human B cells, which

is known to modulate the transcriptional profile of the MYC proto-oncogene. We then apply the method to the systematic identification of novel post-translational modulators of all transcription factors in human B cells that are active in normal, tumor related, and experimentally manipulated human B cells. Inferred modulators are highly enriched in kinases and transcription factors, recapitulate established regulation pathways, and identify novel post-translational regulators of transcriptional activity. Finally, we provide a genome-wide, all-against-all view of the kinome mediated modulation of the human transcriptome. The method is of broad applicability for the identification of cellular network dynamics in mammalian cells and for their comparison in normal and pathologic phenotypes.

Informatics Tools to Support Translational Cancer Research

Michael Becich, M.D., Ph.D. Director of the Benedum Oncology Informatics Center Director of the Center for Pathology Informatics Vice Chairman and Professor of Pathology, Information Sciences and Telecommunications Department of Pathology, University of Pittsburgh Medical School

Ashok Patel MD, Rebecca Crowley, MSIS MD, Mike Davis, Tim Fennell, Doug Fridsma MD, William Gross, Rajnish Gupta, James Lyons-Weiler PhD, John Milnes, Kevin Mitchell MS, Sambit Mohanty, MD PhD, Yimin Nie MD, Anil Parwani, MD PhD, Andrew Pople, Linda Schmandt MS, Susan Urda, Sharon Winters MS, and Michael J. Becich, MD PhD

Abstract: Transforming translational research is the focus of a new NIH Roadmap Initiative and biomedical informatics is viewed as a critical to accomplishing this. This poses significant challenges as well as opportunities for biomedical informatics. At the University of Pittsburgh we are using our experience in Pathology and Oncology to scale a variety of solutions for clinical trials, tissue banking, honest brokering and bioinformatics in service of a general model for supporting translational research in medicine. We describe our work in the area of clinical trials by profiling our Clinical Trials Management System (CTMA) and our work with the Biomedical Research Integrated Domain Group (BRIDG) and the Cancer Biomedical Informatics Grid (caBIG) program. We also talk about our work with caTISSUE a suite of tools that manage and inventory biospecimens in tissue banks. caTISSUE also automates the annotation of biospecimens with pathology, tumor marker, staging, grading and clinical outcome data.. Earlier versions of this tool is already being shared with 14 cancer centers and research institutes in the Pennsylvania Cancer Alliance Bioinformatics Consortium (http://www.pcabc.upmc.edu/main.cfm?dis=db_ov) and the Cooperative Prostate Cancer Tissue Resource (<http://cpctr.cancer.gov/todatabase.html>). We will also describe our honest broker tracking system so critical to our data sharing and tissue/data request fulfillment for the hundreds of researchers using our tissue bank. Finally, we describe our bioinformatics development efforts, particularly our Gene Expression Data Analysis resource (caGEDA, <http://bioinformatics.upmc.edu/cageda>). Our goal is to integrate these systems into one architecture supported by a common set of ontology (vocabulary) services as we transform our research environment and enable whole sale sharing of our software, biospecimens and data for the caBIG program.

NOTE: This work was supported by the following grants:

Cooperative Prostate Cancer Tissue Resource – (see URL above) – NCI - UO1 CA86735-07

Shared Pathology Informatics Network (SPIN) – (<http://spin.nci.nih.gov>) – NCI - UO1 CA91338-05

Univ of Pittsburgh Cancer Institute Cancer Center Support Grant – Cancer Informatics Service – NCI - P30 CA47904
Pennsylvania Cancer Alliance Bioinformatics Consortium – (see URL above) PA DOH - ME 01-740
Cancer Biomedical Informatics Grid – (<http://cabig.nci.nih.gov>) – NCI - 79207CBS10

The Practical Impact of Ontologies on Biomedical Informatics

James Cimino, M.D. Professor of Biomedical Informatics, Columbia University College of Physicians and Surgeons; Director, American Medical Informatics Association

Abstract:

Objectives: To examine recent research work in the development and evaluation of controlled biomedical terminologies - especially, the representation of structured, controlled definitional knowledge about the terms themselves; such terminologies are often referred to as "ontologies".

Methods: A review of the published literature using PubMed, as well as full-text searches of recent Medinfo and American Medical Informatics Association (AMIA) Symposia proceedings, searching for the terms "ontology" and "ontologies" and for articles discussion specific, prominent ontological work.

Results: We summarize the ontologic aspects of twelve current terminology projects: Galen, the Unified Medical Language System (UMLS), the Medical Entities Dictionary (MED), SNOMED-CT, LOINC, the Foundational Model of Anatomy (FMA), the Gene Ontology (GO), ISO Reference Terminology Model for Nursing Diagnosis, NDF-RT, RxNorm, the NCI Thesaurus, and DOLCE+. We discuss the origins, domain, and ontologic representation of each of these and attempt to summarize the impact that each has had on terminologic work and biomedical applications. We also note the contributions of the Protg tool to many of these efforts.

Conclusion: Terminologic research and development have advanced significantly in the past 20 years, especially since the recent orientation toward controlled biomedical ontologies. This work has had significant impact on the development of terminologies themselves, their acceptance and dissemination as standards, and their use in supporting biomedical information systems.

From Bedside to Bench and Back: Healthcare and Life Sciences Semantic Interoperability -- Challenges, Expectations, Possibilities

Charles Mead, M.D., MSc - Senior Associate, Booz Allen Hamilton

Abstract: Fundamental advances in genetics and molecular biology can inform patient care at the bedside, enabling so-called "individualized" healthcare. Likewise, knowledge gained in the course of individual patient care can provide targeted insights at the research bench. However, in order to achieve a scientifically and clinically meaningful level of information sharing, phenotypic and genotypic data will need to be analyzed by – and exchanged between – a number of computer systems along the 'bedside to bench' continuum. Most importantly, the system-specific analysis and data filtering tasks will require that the various systems share data *semantics* (i.e. meaning) rather than simply data *syntax* (i.e. structure).

Over its multiple-hundred-year history, architectural engineering has developed processes and tools for managing and integrating the complexity inherent in designing and building large structures such as airports, freeways, and skyscrapers. Software engineering has recently begun to agree on a number of 'best practice and lessons learned' over the relatively brief 50-year history. Unfortunately, healthcare and life sciences IT has lagged behind other industries (e.g. aerospace) in adoption of these Best Practices. It is now critical that this gap be closed. In parallel, a number of life sciences and healthcare efforts have emerged both nationally and internationally that define various data and information standards. Unfortunately, these standards are often overlapping in scope, conflicting in semantics, and, consequently, difficult to utilize across an integrated continuum. It is now critical that the necessary effort be expended to critically analyze the various existing standards so that they can be effectively integrated or appropriately evolved.

The bad news is that problem of passing the meaning of data between healthcare and life science information systems -- *computable* semantic interoperability -- is *hard*. Emerging technologies such as those associated with the Semantic Web will undoubtedly make certain aspects of the problem more approachable. However, the solution to the problem is not likely to become *easy* in the foreseeable future. The good news is that the problem is *solvable* through the application of a number of existing standards and best practices and technologies in combination with targeted, focused, and informed *human effort*.

TRACK IV **BIOENGINEERING, MEDICAL DEVICES AND LIFE SCIENCES INDUSTRY -
HYATT REGENCY**

Innovation in Bioengineering Technologies

Stelios Andreadis, Ph.D., Associate Professor, Chemical & Biological Engineering, UB
Frank Bright, Ph.D., UB Distinguished Professor of Chemistry, Conger Goodyear Chair of
Chemistry & Associate Chairperson, Department of Chemistry, UB
Kenneth Hoffman, Ph.D., Associate Professor, Department of Neurosurgery, UB
Toshiba Stroke Research Center

Success Stories in Building a Life Sciences Company in WNY

Robert Carey, General Manager, Analytical Instruments, Reichert, Inc.
Dave Barthel, CEO, SmartPill Corporation
Anthony Johnson, Interim CEO, Empire Genomics