Interfacing Optoelectronics with Biosystems – Reading, Writing & Interpreting Genomes

George Church
Department of Genetics, Harvard Medical School

As the semiconductor roadmap approaches the molecular scale (sub-10nm), bionanotechnology is reaching back from the other side with new materials, 3D-CAD, and yield improvements. Another interesting opportunity lies in harnessing evolution to test billions of complex physical designs. Finally, there are contributions of exponential technologies for I/O of optoelectronic bits to I/O of biological bits. The use of CCD/CMOS detectors and Digital Micromirror Devices (DMD/DLP) and ink-jet printers has brought the cost down for reading and writing DNA and other biological systems by 5 logs in 5 years. The impact of this on applications like personal genomes, synthetic life and bioenergy will be discussed. Bioinformatics plays key roles in harvesting parts for synthetic biology and generation hypotheses from large human Genome, Environment and Trait (GET) datasets.

George Church is Professor of Genetics at Harvard Medical School and Director of the Center for Computational Genetics. With degrees from Duke University in Chemistry and Zoology, he co-authored research on 3D-software & RNA structure with Sung-Hou Kim. His PhD from Harvard in Biochemistry & Molecular Biology with Wally Gilbert included the first direct genomic sequencing method in 1984; initiating the Human Genome Project then as a Research Scientist at newly-formed Biogen Inc. and a Monsanto Life Sciences Research Fellow at UCSF with Gail Martin. He invented the broadly-applied concepts of molecular multiplexing and tags, homologous recombination methods, and array DNA synthesizers. Technology transfer of automated sequencing & software to Genome Therapeutics Corp. resulted in the first commercial genome sequence (the human pathogen, H. pylori, 1994). This multiplex solid-phase sequencing evolved into polonies (1999), ABI-SOLiD (2005) & open-source Polonator.org (2007) and Personal Genomes.org. He has served in advisory roles for 12 journals (including Nature Molecular Systems Biology), 5 granting agencies and 24 biotech companies (e.g. 23andme & recently founding Codon Devices, Knome and LS9). Current research focuses on integrating biosystems-modeling with Personal Genomics & synthetic biology.
The Origin and Early Evolution of Life: Can Bioinformatics Bridge the Gap Between Prebiotic Chemistry and Deep Phylogenies?

Antonio Lazcano
School of Sciences, National Autonomous University of Mexico

Analysis of carbon-rich meteorites and the laboratory simulations of the primitive Earth suggest that prior to the emergence of the first living systems the prebiotic environment was endowed with a large suite of organic compounds of biochemical significance, many organic and inorganic catalysts, purines and pyrimidines, i.e., the potential for template-dependent polymerization reactions; and membrane-forming compounds. The remarkable coincidence between the monomeric constituents of living organisms and those synthesized in Miller-type experiments appears to be too striking to be fortuitous. However, how the ubiquitous nucleic acid-based genetic system of extant life may have originated from such a mixture is one of the major unsolved problems in contemporary biology. The discovery of catalytically active RNA molecules provided considerable credibility to prior suggestions that the first living entities were largely based on ribozymes, in an early stage called the RNA world. There are many indications of the robustness of the RNA world hypothesis, but at the time being the hiatus between the primitive soup and the RNA world is discouragingly enormous, and the problem of how RNA came into being is still an open one.

Bioinformatics and comparative genomics provide important insights into some very early stages of biological evolution, but it is difficult to see how their applicability can be extended beyond a threshold that corresponds to a period in which protein biosynthesis was already in operation, i.e., the RNA/protein world. They are also not applicable to the origin of life itself, since all possible intermediates that may have once existed have long since vanished. Given the huge gap in the current descriptions of the evolutionary transition between the prebiotic synthesis of biochemical compounds and the last common ancestor of all extant living beings, it may be naive to attempt to describe the origin of life and the nature of the first living systems based on molecular cladistics.

Antonio Lazcano, a professor at the Universidad Nacional Autónoma de México (UNAM) in Mexico City, has studied the origin and early evolution of life for over 30 years. He was trained both as an undergraduate and graduate student at the UNAM, where he rapidly focused on the study of prebiotic evolution and the emergence of life. An academic deeply committed to public education, he has also devoted considerable efforts to scientific journalism and teaching. He is the author of several books published in Spanish, including The Origin of Life, which has become a bestseller with over 650,000 copies sold. He is considered the foremost promoter of evolutionary biology and the study of the origins of life in Latin America, and has been Professor-in-Residence or Visiting Scientist in France, Spain, Italy, Cuba, Switzerland, Russia and the USA. He has organized a number of international symposia and scientific meetings, and has been member of number of editorial boards of major journals. He has served on many international advisory and review boards for NASA and other international organizations. He was twice President of the International Society for the Study of the Origin of Life, the first Latin American scientist to occupy this position.
Amyotrophic Lateral Sclerosis and Frontotemporal Lobar Degeneration: Connecting the Dots through TDP-43

Virginia Lee
Department of Pathology and Laboratory Medicine,
University of Pennsylvania School of Medicine

The disease protein in frontotemporal lobar degeneration with ubiquitin-positive inclusions (FTLD-U) and amyotrophic lateral sclerosis (ALS) was identified recently as the TAR DNA-binding protein (TDP-43) thereby providing a molecular link between these two disorders. Moreover, the discovery of mutations in the TDP-43 gene (TARDBP) in familial ALS and FTLD-U with similar TDP-43 pathology suggests that it is a primary cause of these disorders. Genome-wide association (GWA) studies on patients with clinical FTLD-U and autopsy confirmed TDP-43 inclusions showed association to multiple single nucleotide polymorphisms (SNP) within a single linkage disequilibrium (LD) block on 7p21 that contains the TMEM106B gene. Expression data suggest the TMEM106B variants may confer disease risk by increasing TMEM106B expression via a cis-acting mechanism thereby implicating TMEM106B as a strong risk factor for FTLD-U and suggest a pathogenic mechanism for risk-associated genotype. Other studies also showed that TDP-43 aggregates are present to variable extent in other neurodegenerative disorders including Alzheimer’s and Parkinson’s. Furthermore, TDP-43 accumulations are the first non-amyloidogenic aggregates found in any neurodegenerative diseases. This presentation will summarize the rapidly advancing field of TDP-43 proteinopathies.

Dr. Virginia M.-Y. Lee is the John H. Ware 3rd Professor in Alzheimer’s Research in the Department of Pathology and Laboratory Medicine. She is the Director of the Center for Neurodegenerative Disease Research and Co-director of the Marian S. Ware Alzheimer Drug Discovery Program at the University of Pennsylvania, School of Medicine. Dr. Lee is the recipient of the Metropolitan Life Foundation Award for Medical Research in Alzheimer’s Disease (1991, 1996), the Potamkin Prize for Medical Research in Alzheimer’s Disease (1998) and the Bristol-Myers Squibb Biomedical Research Grant in Neuroscience Research (2003). Dr. Lee was a member of the National Advisory Council on Aging (NIH) and elected to membership in the Institute of Medicine of the National Academies. Dr. Lee’s research focuses on disease proteins that become misfolded and accumulate as pathological inclusions in hereditary and sporadic Alzheimer’s disease (AD), Parkinson’s disease (PD), frontotemporal dementia (FTD), amyotrophic lateral sclerosis (ALS) and related neurodegenerative disorders of aging. Her work was instrumental in identifying and demonstrating that tau, α-synuclein and TDP-43 proteins form unique brain aggregates in numerous neurodegenerative diseases and provided critical evidence supporting her hypothesis that aggregation of brain proteins into potentially toxic lesions is a common mechanistic theme in diverse neurodegenerative diseases including AD, PD, FTD, ALS and related disorders. Significantly, studies of human postmortem brains and model systems of these diseases, including hereditary forms thereof, have enabled Lee to implicate the abnormal aggregation of tau, α-synuclein and TDP-43 in mechanisms leading to the formation of these inclusions, to show that these inclusions compromise neuronal viability, and, most importantly, to identify targets in drug discovery to develop better treatments for these disorders.
Statistical Models for Predicting HIV Phenotypes and Effectiveness of Antiretroviral Therapies

Thomas Lengauer
Max-Planck-Institut für Informatik

The prime problem in applying antiviral drug therapy to AIDS patients is the variability of the AIDS virus HIV. Presented with a given drug therapy the virus follows evolutionary escape paths to drug resistance, which renders the drug therapy ineffective. Statistical computer models can help estimate the level of resistance of a given viral variant against a given antiviral drug. Furthermore, such models can estimate the evolutionary path to resistance followed by the virus.

We describe the science behind the geno2pheno Server www.geno2pheno.org which offers such models freely via the internet: Given the relevant portion of an HIV genome, geno2pheno predicts the resistance of HIV to any of a number of antiviral drugs that are in clinical use. Furthermore the server ranks combination drug therapies with respect to their expected effectiveness against the given HIV variant.

Entry inhibitors aim at preventing HIV from entering the host immune cell. Here, we are confronted with a different computational challenge: When entering the human host cell, HIV uses one of two coreceptor molecules on the cell surface. We present a statistical model offered via the geno2pheno server that predicts which of the two coreceptors the virus uses. Determining this “viral tropism” is essential for disease prognosis and in the context of applying the new entry inhibitors.

Both model classes are trained using various linear and nonlinear statistical learning procedures. The training data are carefully assembled databases comprising relevant genotypic, phenotypic and clinical parameters. While the resistance models only incorporate sequence features, one version of the tropism models also involves information on the structure of the relevant portion of the viral gp120 protein that docks to the human cell.

The geno2pheno server has been developed in the context of the Arevir consortium, a German National research consortium targeted at the bioinformatical analysis of HIV resistance data, and is currently in prototypical use for research purposes. Members of the consortium and their associated practices treat about two thirds of the AIDS patients in Germany. In recent years, data collection and method development has been lifted to a European level in the EuResist project (www.euresist.org).

Prof. Dr. Dr. Thomas Lengauer, (born 1952) is Director at the Max Planck Institute for Informatics and a professor of Informatics at Saarland University in Saarbrucken, Germany. His background is in Math (Dr. rer. nat. Berlin, Germany 1976) and Computer Science (Ph.D., Stanford 1979). In the seventies he performed research in theoretical computer science, in the 80s on design methods for integrated circuits. He has been engaged in research in computational biology since the beginning of the 90s. His major focuses of research are protein bioinformatics, computational drug screening and design and bioinformatics for understanding and curing diseases. Previously, he held the positions of a full professor at the University of Paderborn, Germany (1984-1992) and of a Director of the Institute for Algorithms and Scientific Computing at the German National Research Center for Computer Science in Sankt Augustin, near Bonn Germany (1992-2001). Dr. Lengauer is a founding member of the International Society for Computational Biology (ISCB), a member of the steering board of the international conference series RECOMB, and he headed the steering board of the European bioinformatics conference series ECCB since its foundation in 2002 until 2005. In 2001 he co-founded the BiosolveIT GmbH, Sankt Augustin, Germany, which develops and distributes Cheminformatics software. In 2003 he received the Konrad Zuse Medal of the German Informatics Society and the Karl Heinz Beckurts Award. He is a member and a senator of the German Academy of Sciences Leopoldina as well as a member of acatech – German Academy of Science and Engineering.
A Novel Modeling Strategy for Dynamic Modeling of Metabolic Systems

James Liao
Department of Chemical and Biomolecular Engineering, University of California at Los Angeles

Complete modeling of metabolic networks is desirable but difficult for the lack of kinetics. As a step towards this goal, we develop an approach to build an ensemble of dynamic models which reach the same steady state. The models in the ensemble are based on the same mechanistic framework, including known regulations, and span the space of all kinetics allowable by thermodynamics. This ensemble allows for the examination of possible phenotypes of the network upon perturbations, such as changes in enzyme expression levels. The size of the ensemble is reduced by acquiring data for such perturbation phenotypes. If the mechanistic framework is approximately accurate, the ensemble converges to a smaller set of models and becomes more predictive. This approach bypasses the need for detailed characterization of kinetic parameters and arrives at a set of models that describes relevant phenotypes upon enzyme perturbations.

Dr. James C. Liao, Chancellor’s professor, Chemical and Biomolecular Engineering, UCLA, is a pioneer in Metabolic Engineering, Synthetic Biology, and Systems Biology. He received his BS degree from National Taiwan University and PhD from University of Wisconsin-Madison. After working as a research scientist at Eastman Kodak Company, Rochester, NY, he started his academic career at Texas A&M University in 1990 and moved to UCLA in 1997. He was elected Fellow of American Institute for Medical and Biological Engineering, 2002, and received numerous awards, including the NSF Young Investigator Award (1992), the Merck Award for Metabolic Engineering (2006), the Food, Pharmaceutical, and Bioengineering Division award of American Institute of Chemical Engineers (AIChE) (2006), the Charles Thom Award of the Society for Industrial Microbiology (2007), the Marvin Johnson Award of American Chemical Society (2009), the Alpha Chi Sigma Award of AIChE (2009), and the James E. Bailey Award of Society for Biological Engineering (2009).
Breakthroughs in Imaging Using Photoactivatable Fluorescent Proteins

Jennifer Lippincott-Schwartz
Cell Biology and Metabolism Branch, National Institutes of Health

Photoactivatable fluorescent proteins (PA-FPs) are molecules that switch to a new fluorescent state in response to activation to generate a high level of contrast. Several types of PA-FPs have been developed, including PA-FPs that fluoresce green or red, or convert from green to red in response to activating light. The optical “highlighting” capability of PA-FPs has led to the rise of novel imaging techniques providing important new biological insights. These range from in cellulo pulse-chase labeling for tracking subpopulations of cells, organelles or proteins under physiological settings, to super-resolution imaging of single molecules for determining intracellular protein distributions at nanometer precision. The use of PA-FPs in super-resolution imaging of single molecules is a rapidly emerging field of microscopy that improves the spatial resolution of light microscopy by over an order of magnitude (10-20 nm resolution). It is based on the controlled activation and sampling of sparse subsets of photoconvertible fluorescent molecules whose illumination centroids are fitted and then summed into a final super resolution image, revealing the complex distribution of dense populations of molecules within subcellular structures with nanometer precision. The full potential of PA-FPs in conventional, diffraction-limited and super-resolution imaging is only beginning to be realized. Here, I discuss the diverse array of PA-FPs available to researchers and the new imaging techniques they make possible for unraveling long-standing biological questions.

Jennifer Lippincott-Schwartz obtained her Ph.D from Johns Hopkins University in Baltimore, MD, received post-doctoral training with Dr. Richard Klausner at the National Institute of Child Health and Human Development (NICHD), NIH, Bethesda, MD, and is currently Chief of the Section on Organelle Biology in the Cell Biology and Metabolism Branch of the NICHD. Lippincott-Schwartz’s research uses live cell imaging approaches to analyze the spatio-temporal behavior and dynamic interactions of molecules in cells. These approaches have helped to change the conventional ‘static’ view of protein distribution and function in cells to a more dynamic view that integrates information on protein localization, concentration, diffusion and interactions that are indiscernible from protein sequences and in vitro biochemical experiments alone. The projects in Lippincott-Schwartz’s lab cover a vast range of cell biological topics, including protein transport and the cytoskeleton, organelle assembly and disassembly, and the generation of cell polarity. Analysis of the dynamics of fluorescently labeled proteins expressed in cells is performed using numerous live cell imaging approaches, including FRAP, FCS and photoactivation. Most recently, her research employs photoactivation localization microscopy, called PALM, that enables visualization of molecule distributions at high density at the nano-scale. Dr. Lippincott-Schwartz serves as Editor for Current Protocols in Cell Biology and The Journal of Cell Science and she is on the Editorial Boards of Cell and Molecular Biology of the Cell. She is an active member of the scientific community, serving as a member of the Council for the American Society of Cell Biology and on the Executive Board of the Biophysical Society. She was elected to the National Academy of Sciences in 2008.