Cell, Vol. 121, 511–513, May 20, 2005, Copyright ©2005 by Elsevier Inc. DOI 10.1016/j.cell.2005.04.020

Systems Biology: Its Practice and Challenges

Commentary

Alan Aderem*
The Institute for Systems Biology
1441 North 34th Street
Seattle, Washington 98103

Systems biology is a comprehensive quantitative analysis of the manner in which all the components of a biological system interact functionally over time. Such an analysis is executed by an interdisciplinary team of investigators that is also capable of developing required technologies and computational tools. In this model, biology dictates what new technology and computational tools should be developed, and, once developed, these tools open new frontiers in biology for exploration. Thus, biology drives technology and computation, and, in turn, technology and computation revolutionize biology.

Systems approaches have long been taken, particularly in immunology, physiology, development, and neurobiology. However, technology development during the 1980s permitted the concepts generated by many years of reductionist inquiry to be analyzed in the context of the entire system. Automated DNA sequencers enabled the sequencing of genomes and the definition of polymorphisms among individuals; microarray analysis permitted global transcriptional profiling, and advances in mass spectrometry led to large-scale proteomic and metabolomic analysis. The mountains of data generated by these high-throughput platforms led to the rapid growth of computational biology and bioinformatics. Thus, knowledge of the complete sequences of genomes, together with technology that permit the monitoring of information flow leading to specific cellular functions, set the stage for systems biology.

There are three basic concepts that are crucial to understanding complex biological systems: they are emergence, robustness, and modularity.

Emergence. Complex systems display properties, often called "emergent properties," that are not demonstrated by their individual parts and cannot be predicted even with full understanding of the parts alone. For example, understanding the properties of hydrogen and oxygen does not allow us to predict the properties of water. Life is an example of an emergent property. It is not inherent in DNA, RNA, proteins, carbohydrates, or lipids but is a consequence of their actions and interactions. A comprehensive understanding of such emergent properties requires systems-level perspectives and cannot be gleaned from simple reductionist approaches.

Robustness. Biological systems maintain phenotypic stability in the face of diverse perturbations imposed by the environment, stochastic events, and genetic variation. Robustness often arises through positive and negative feedback loops and other forms of control that constrain a gene's output. This feedback insulates the system from fluctuations imposed on it by the environment. Positive feedback, in general, enhances sensitivity, whereas negative feedback can dampen noise and reject perturbations. Robustness is an inherent property of all biological systems and is strongly favored by evolution.

Modularity. A further characteristic of complex systems is their modularity. Multiple useful definitions of a module exist. To an engineer, a module is a functional unit, a collection of parts that interact together to perform a distinct function. Such a module would have distinct inputs, things it is sensitive to, and outputs, things it controls. To a biologist, a module in a network is a set of nodes that have strong interactions and a common function. Modularity can contribute to both robustness of the entire system, by confining damage to separable parts, and to evolution, by simply rewiring modules. Furthermore, modularity decreases the risk of failure of the system by preventing the spread of damage in one part of the network throughout the entire network.

Network Modeling by Iterative Refinement

To practice systems biology, one must capture and integrate global sets of biological data from as many hierarchical levels of information as possible (see Figure 1). These could include DNA sequences, RNA and protein measurements, protein-protein and protein-DNA interactions, biomodules, signaling and gene regulatory networks, cells, organs, individuals, populations, and ecologies. The data is then transferred to a comprehensive database, where it is warehoused and annotated. Human minds are incapable of inferring the emergent properties of a system from thousands of data points, but we have evolved to intelligently interpret an enormous amount of visual information. The data is therefore transferred to visualization programs. This is the initiation point for the formulation of detailed graphical or mathematical models, which are then refined by hypothesis-driven, iterative systems perturbations and data integration. In this manner, the phenotypic features of the system are tied directly to the behavior of the protein and gene regulatory networks. Cycles of iteration will result in a more accurate model; ultimately, these models will explain the systems or emergent properties of the biological system of interest. Once the model is sufficiently accurate and detailed, it will allow biologists to accomplish two tasks never before possible: (1) predict the behavior of the system given any perturbation and (2) redesign or perturb the gene regulatory networks to create completely new emergent systems properties. This latter possibility lies at the heart of preventative medicine. Thus, systems biology is hypothesis driven, global, quantitative, iterative, integrative, and dynamic.

Some Technical Challenges for Systems Biology

- Data quality and standardization. Systems approaches rely heavily on information in the public databases. The datasets are often incomplete, not standardized, or properly annotated; worse yet, the

---

*Correspondence: aderem@systemsbiology.org

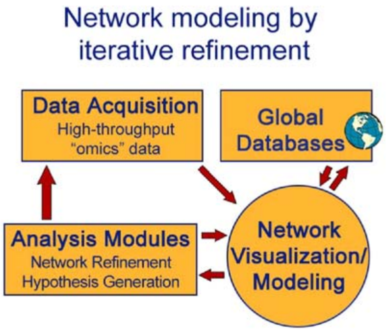

Figure 1. Network Modeling by Iterative Refinement

The main features of this model are described in the text.

quality of the data is often uncertain. It is essential that metrics be developed for the validation of large data sets.

- Network biology is in its infancy. Future needs range from the development of new theoretical methods to characterize network topology, to insights into the dynamics of motif clusters and biological function.
- Sensitive tools for identifying and quantifying the concentrations, fluxes, and interactions of various types of molecules at high resolution both in space and time are required. These dynamic measurements must be made in the appropriate context of specific networks, cells, and organisms.
- Miniaturized and automated microfluidics/nanotechnology platforms capable of parallel multiparameter analysis that integrates operations such as cell sorting and single-cell gene and protein profiling are necessary. In addition, nanomechanical and nanoelectronic devices will also permit the quantification of the forces and kinetics associated with protein/protein, protein/DNA, and protein/drug interactions.
- Imaging will need to be extended to dynamic, spatial, multiparameter measurements within single cells. Furthermore, hypotheses must ultimately be tested in whole animals. Such testing requires advances in molecular imaging ranging from bioluminescence and fluorescence to positron emission tomography (PET) and magnetic resonance imaging (MRI).

Novel Challenges for Universities in the Practice of Systems Biology

Teamwork, Fair Credit, and Data Ownership. Systems biology is an interdisciplinary endeavor that necessitates investigators to work in teams. There are a number of reasons that interdisciplinary teamwork is difficult in a typical university context. First, the departmental structures found within universities work against interdisciplinary interactions; biologists, engineers, and computer scientists rarely interact. Second, academics are reluctant to work in teams because their advancement is dependent upon their individual contribution; large, multiauthored papers are not particularly favored by university promotions committees. Clearly, a new set of standards for academic advancement will have to be adopted. University promotions committees will need to find a way of evaluating investigators who have not taken traditional academic routes.

Access to Technology. Systems approaches require access to the technologies necessary to capture the various levels of biological information and to integrate these information types with hypothesis-driven science. Clearly, a way must be found to democratize the technologies and permit access to a wide range of investigators. One approach has been the establishment of core facilities. There are a number of drawbacks to such facilities. It is extremely difficult to maintain state-of-the-art technologies, and, most importantly, core facilities usually lack the ability to integrate technology and biology. Thus, biological experiments are shoehorned into preset technology platforms. Despite these shortcomings, core facilities remain a very important resource.

The Problem of Funding. Traditionally, NIH study sections are relatively conservative in funding what they consider to be risky projects and new approaches. Skilled grant writers know that a grant examining questions to which the answer is essentially known will be funded with the highest priority. Labels such as "fishing expeditions" or "stamp collecting," the derogatory terms for the discovery component of a systems biology proposal, are usually the death knell of any grant. However, a comprehensive collection of information that defines an entire system is a critical starting point to the integration of discovery science and hypothesis-driven science, and the education of study sections in this regard is a challenging problem, indeed.

Contrasting Academia and Industry with Respect to the Practice of Systems Biology

It is interesting to contrast the strengths and weaknesses of academia and industry when considering the practice of systems biology. Academic laboratories are flexible and can take long-term perspectives. Unfortunately, as discussed above, they do not have easy access to large platform technologies, and they have difficulty working in interdisciplinary teams. In contrast, industrial laboratories have access to large-scale technologies, and the scientists are organized in interdisciplinary teams. But industrial laboratories are answerable to stock holders and must deliver products, two constraints that prevent them taking long-term perspectives in science. I believe that a new academic model is possible, situated between the classic academic and industrial positions, with the advantages of both and the disadvantages of neither. The Institute for Systems Biology is designed as a prototype of such an academic institution.

The Institute for Systems Biology

The Institute for Systems Biology (ISB) was founded in 2000 to respond to the opportunities in biological research emerging from the Human Genome Project. Its mission to develop and apply the tools necessary for systems biology has required the development of large-scale facilities for DNA sequencing; genotyping; DNA arrays; proteomics; high-speed, multiparameter cell sorting; nanotechnology and microfluidics; imaging; and a strong computational infrastructure. The staff comprises crossdisciplinary scientists from the computer sciences, engineering, mathematics, physics, and, of course, biology. The crossdisciplinary scientists are focused on the development of new global technologies and improvements in the existing technologies, in developing powerful new computational technologies, and in the creation of the tools necessary for the graphical display and mathematical modeling of biological systems.

The ISB also seeks academic and industrial partnerships. Collaborations with academic laboratories take a number of forms, including project-based collaborations, courses, and sabbaticals. These partnerships are an important component of efforts to integrate large- and small-scale science and to democratize the accessibility of high-throughput technologies.

Indeed, if one had to summarize the essence of the ISB, it would be integration: integration of the various technologies, integration of the various hierarchical levels of biological information, integration of technology and biology, integration of crossdisciplinary scientists, integration of industry and academia, and, finally, integration of discovery and hypothesis-driven science.

The Promise of Systems Biology

Systems biology will bring about a revolution in the practice of medicine. Knowledge of specific genetic traits, coupled with multiparameter diagnostics, will generate routine procedures for assessing health and disease status. Knowledge of disease-perturbed networks will facilitate drug discovery, and pharmacological intervention will focus on preventing disease-mediated transitions. This predictive and preventative medicine will lead naturally to a personalized medicine, in which therapeutic strategies will be tailored to individual needs. In this sense, systems biology will fundamentally transform society.