87e Network Medicine: Systems Biology in Health and Disease

Joseph Loscalzo

The field of human biology has progressed over the last three centuries largely as a result of the reductionist approach to the scientific problems that challenge the discipline. Biologists study the experimental response of a variable of interest in a cell or organism while holding all other variables constant. In this way, it is possible to dissect the individual components of a biologic system and assume that a thorough understanding of a specific component (e.g., an enzyme or a transcription factor) will provide sufficient insight to explain the global behavior of that system (e.g., a metabolic pathway or a gene network, respectively). Biologic systems are, however, much more complex than this approach assumes and manifest behaviors that frequently (if not invariably) cannot be predicted from knowledge of their component parts characterized in isolation. Growing recognition of this shortcoming of conventional biologic research has led to the development of a new discipline, *systems biology, which is defined as the holistic study of living organisms or their cellular or molecular network components to predict their response to perturbations*. Concepts of systems biology can be applied readily to human disease and therapy and define the field of *systems pathobiology,* in which genetic or environmental perturbations produce disease and drug perturbations restore normal system behavior.

Systems biology evolved from the field of systems engineering in which a linked collection of component parts constitute a network whose output the engineer wishes to predict. The simple example of an electronic circuit can be used to illustrate some basic systems engineering concepts. All the individual elements of the circuit—resistors, capacitors, transistors—have well-defined properties that can be characterized precisely. However, they can be linked (wired or configured) in a variety of ways, each of which yields a circuit whose response to voltage applied across it is different from the response of every other configuration. To predict the circuit's (i.e., system's) behavior, the engineer must study its response to perturbation (e.g., voltage applied across it) holistically rather than its individual components' responses to that perturbation. Viewed another way, the resulting behavior of the system is greater than (or different from) the simple sum of its parts, and systems engineering utilizes rigorous mathematical approaches to predict these complex, often nonlinear, responses. By analogy to biologic systems, one can reason that detailed knowledge of a single enzyme in a metabolic pathway or of a single transcription factor in a gene network will not provide sufficient detail to predict the output of that metabolic pathway or transcriptional network, respectively. Only a systems-based approach will suffice.

It has taken biologists a long time to appreciate the importance of systems approaches to biomedical problems. Reductionism has reigned supreme for many decades, largely because it is experimentally and analytically simpler than holism, and because it has provided insights into biologic mechanisms and disease pathogenesis that have led to successful therapies. However, reductionism cannot solve all biomedical problems. For example, the so-called off-target effects of new drugs that frequently limit their adoption likely reflect the failure of a drug to be studied in holistic context, i.e., the failure to explore all possible actions aside from the principal target action for which it was developed. Other approaches to understanding biology therefore are clearly needed. With the growing body of genomic, proteomic, and metabolomic data sets in which dynamic changes in the expression of many genes and many metabolites are recorded after a perturbation and with the growth of rigorous mathematical approaches to analyzing those changes, the stage has been set for applying systems engineering principles to modern biology.

Physiologists historically have had more of a (bio)engineering perspective on the conduct of their studies and have been among the first systems biologists. Yet, with few exceptions, they, too, have focused on comparatively simple physiologic systems that are tractable using conventional reductionist approaches. Efforts at integrative modeling of human physiologic systems, as first attempted by Guyton for blood pressure regulation, represent one application of systems engineering to human biology. These dynamic physiologic models often focus on the acute response of a measurable physiologic parameter to a system perturbation, and do so from a classic analytic perspective in which all the conventional physiologic determinants of the output parameter are known and can be modeled quantitatively.

Until recently, molecular systems analysis has been limited owing to inadequate knowledge of the molecular determinants of a biologic system of interest. Although biochemists have approached metabolic pathways from a systems perspective for over 50 years, their efforts have been limited by the inadequacy of key information for each enzyme (K_M, k_{cat}, and concentration) and substrate (concentration) in the pathway. With increasingly rich molecular data sets available for systems-based analyses, including genomic, transcriptomic, proteomic, and metabolomic data, biochemists are now poised to use systems biology approaches to explore biologic and pathobiologic phenomena.

PROPERTIES OF COMPLEX BIOLOGIC SYSTEMS

To understand how best to apply the principles of systems biology to human biomedicine, it is necessary to review briefly the building blocks of any biologic system and the determinants of system complexity. All systems can be analyzed by defining their static topology (architecture) and their dynamic (i.e., time-dependent) response to perturbation. In the discussion that follows, system properties are described that derive from the consequences of topology (form) or dynamic response (function). Any system of interacting elements can be represented schematically as a *network* in which the individual elements are depicted as nodes and their connections are depicted as links. The nature of the links among nodes reflects the degree of complexity of the system. *Simple systems* are those in which the nodes are linearly linked with occasional feedback or feedforward loops modulating system throughput in highly predictable ways. By contrast, *complex systems* are nodes that are linked in more complicated, nonlinear networks; the behavior of these systems by definition is inherently more difficult to predict owing to the nature of the interacting links, the dependence of the system's behavior on its initial conditions, and the inability to measure the overall state of the system at any specific time with great precision. Complex systems can be depicted as a network of lower-complexity interacting components or modules, each of which can be reduced further to simpler analyzable canonical motifs (such as feedback and feedforward loops, or negative and positive autoregulation); however, a central property of complex systems is that simplifying their structures by identifying and characterizing the individual nodes and links or even simpler substructures does not necessarily yield a predictable understanding of a system's behavior. Thus, the functioning system is greater than (or different from) the sum of its individual, tractable parts.

Defined in this way, most biologic systems are complex systems that can be represented as networks whose behaviors are not readily predictable from simple reductionist principles. The nodes, for example, can be metabolites that are linked by the enzymes that cause their transformations, transcription factors that are linked by the genes whose expression they influence, or proteins in an interaction network that are linked by cofactors that facilitate interactions or by thermodynamic forces that facilitate their physical association. Biologic systems typically are organized as *scale-free*, rather than stochastic, networks of nodes. Scale-free networks are those in which a few nodes have many links to other nodes (highly linked nodes, or hubs) but most nodes have only a few links (weakly linked nodes). The term *scale-free* refers to the fact that the connectivity of nodes in the network is invariant with respect to the size of the network. This is quite different from two other common network architectures: random (Poisson) and exponential distributions. Scale-free networks can be mathematically described by a power law that defines the probability of the number of