Gene regulatory networks

The Special Feature on gene regulatory networks in this issue of PNAS highlights an emerging field in the biosciences: gene regulatory networks that control animal development. The complex control systems underlying development have probably been evolving for more than a billion years. They regulate the expression of thousands of genes in any given developmental process. They are essentially hardwired genomic regulatory codes, the role of which is to specify the sets of genes that must be expressed in specific spatial and temporal patterns. In physical terms, these control systems consist of many thousands of modular DNA sequences. Each such module receives and integrates multiple inputs, in the form of regulatory proteins (activators and repressors) that recognize specific sequences within them. The end result is the precise transcriptional control of the associated genes. Some regulatory modules control the activities of the genes encoding regulatory proteins. Functional linkages between these particular genes, and their associated regulatory modules, define the core networks underlying development.

Gene regulatory networks explicitly represent the causality of developmental processes. They explain exactly how genomic sequence encodes the regulation of expression of the sets of genes that progressively generate developmental patterns and execute the construction of multiple states of differentiation.

As this new field takes shape, the following are among the key emergent concepts:

- The regulatory genome as a logic processing system: Every regulatory module contained in the genome receives multiple disparate inputs and processes them in ways that can be mathematically represented as combinations of logic functions (e.g., "and" functions, "switch" functions, "or" functions). At the system level, a gene regulatory network consists of assemblages of these information-processing units; thus, it is essentially a network of analogue computational devices, the functions of which are conditional on their inputs.
- Causality in the regulatory genome: The reasons why genes are expressed when and where they are in the spatial domains of the developing organism are revealed in network "architecture," that is, in the total aggregate pattern of regulatory linkages. Definitive regulatory functions emerge only from the architecture of intergenic linkages, and these functions are not visible at the level of any individual genes. Examples are the multigenic circuits that act to produce positive or negative feedback loops. It is most important to determine regulatory network architecture, and this can be done by experimental perturbation followed by measurement of the effects on function of many individual genes. But gene regulatory network architecture can be authenticated only by experimental molecular biology in which the functional meaning of given regulatory sequences is directly determined.
- Network substructure: Gene regulatory networks are inhomogeneous compositions of different kinds of subcircuits, each performing a specific kind of function. This concept is important, because it holds the key to network design principles. Some subcircuits are used in many diverse biological contexts, for example, most signal transduction subcircuits, just as particular electronic subcircuit devices are used in diverse kinds of processors. Others are more complex and are dedicated to similar biological functions wherever they appear. Subcircuits of these latter kinds are "wired" in such a way that they are not easily reorganized. However, evolutionary comparison shows that other types of network linkages are far more flexible and malleable. Even subtle modifications in these latter linkages can create morphological diversity among related animal groups.
- Reengineering genomic control systems: To redesign these most potent of all biological control systems, to both intellectual and practical ends, it is necessary to understand thoroughly the flow of causality in a genomically encoded gene regulatory network. Such understanding requires a uniquely interdisciplinary mix of theory and experiment, computation and molecular biology, high-end technology, and sophisticated command of the biological system. Once appreciated and experimentally controlled, the inbuilt richness of genomic control is certain to provide insights into processes that we can only begin to define.

Eric Davidson,
Member, PNAS Editorial Board

Michael Levine,
Special Feature Editor