Fifty Years after Jacob and Monod: What Are the Unanswered Questions in Molecular Biology?

Tomorrow's Molecular Biology

Marc W. Kirschner
Harvard University

Today, the term "molecular biology" has a rather prosaic sound. After all, there is little in contemporary biology that isn't molecular. It's hard to remember that 50 years ago, molecular biology was distinctly countercultural. Whereas biochemistry proudly traced its roots to the sober field of chemistry, molecular biology stood for pure bravado and opportunism. Its immodest goal was to understand the program of life from gene to phenotype. Has it fulfilled its promise? The operon theory of Jacob and Monod stands as one of its signal achievements, but a modern view of how the circuit actually works illustrates its limitations as well as its accomplishments. To fully understand the lac operon requires precisely the kind of biochemistry and physical chemistry that early molecular biologists eschewed. Today, the path from genotype to phenotype looks nothing like a precise Boolean circuit. It is ridden with interactions of low specificity, futile cycles, and redundancy. The eukaryotic chromosome is messy and complex, perhaps ultimately understandable but hardly predictable. Whatever new field emerges will have to transcend space and time, connecting molecules not just to reactions of high specificity but to byzantine pathways concerned with cellular, tissue, and organism homeostasis. It must move beyond binding and catalysis to encompass cellular morphogenesis, development, and physiology, each with its own metalogic. It is obvious that molecular biology cannot do this with the same intellectual tools it possessed at its birth. Tomorrow's molecular biologists might wish to look to their predecessors, take risks, and defy convention.

Regulatory Circuitry

Lucy Shapiro and Harley McAdams
Stanford University

The Jacob/Monod explanation of *E. coli*'s *lac* operon provided the initial concepts of bacterial genetic circuitry that we now know to be an integrated system involving the entire cell. This system includes not only transcriptional networks, but also regulation by dynamically localized phospho-signaling proteins and proteases that together control the orderly activation of subsystems and, in some species, asymmetric cell division. Key regulatory proteins in the *Caulobacter* cell-cycle control network are widely conserved among the α-proteobacteria, but connectivity to other subsystems has been rewired to reflect the environmental niches and fitness strategies of individual species. Thus, much like the "kernels" of gene regulatory networks proposed by Eric Davidson to regulate metazoan body plan development, the core connectivity is conserved, while the periphery connectivity is highly plastic under evolutionary pressure. Moving forward, a concerted and focused effort to develop a whole-cell- or system-level understanding of the integrated regulatory circuitry in a few well-chosen, single-cell eukaryotes will have high payoff. Multicellular organisms will introduce another level of complexity because the essential properties of an integrated regulatory system are emergent—they cannot be recognized or studied in the absence of detailed knowledge of the constituent subsystems or without system-level analysis needed to understand the functioning of the whole. Clearly, these challenges and their resolution will yield valuable and far-reaching insights when applied to understanding cells and tissues whose regulatory systems have been co-opted by viruses or oncogene function.

Oper-ating in Chromatin

Genevieve Almouzni
Institut Curie

The operon model by Jacob and Monod has for 50 years been a major conceptual framework to explain how gene activity is regulated. This vision, which can be applied to "*E. coli* all the way up to the elephant" (Jacques Monod), provides a basis for understanding the rich complexity of gene regulation. One challenge that molecular biology faces today is to study the genome as it exists in the cell: packaged in chromatin and organized within the nucleus, as a template that can respond to external signals. Indeed, it will be important to move beyond the simple linear mode of regulation using individual elements and to integrate the concepts of space, time, and cellular context into our understanding of gene regulation. To accomplish this, the ability to map and image, with high resolution, both genome organization and function as an integrated network will be key. Then, to further determine the plasticity of the system and how, in the context of a whole organism, each cell lineage emerges with its own characteristics during development, will have to come into the picture. Finally, an intriguing question that will keep many people busy is how the memory of cell identity is preserved throughout cell division and what, beyond mere DNA, is actually transmitted. The future promises a lot of excitement for the next generations of biologists, as imagining multiscale gene regulatory mechanisms is just as fascinating as orchestrating an intricate opera.

Integrated Regulation

Phillip A. Sharp
Massachusetts Institute of Technology

The lac operon model for gene regulation introduced a fundamental paradigm that has persisted over the past half century. In modern terms, the model indicates that DNA binding factors such as transcription factors are the "master" regulators in controlling cell fate and responses to environmental signals external to the cell. This paradigm neatly explained Yamanaka's landmark experiment on induction of pluripotent stem cells from somatic cells by introduction of genes encoding the transcription factors Oct 4, Nanog, and Sox2 (Takahashi, K., and Yamanaka, S. [2006]. Cell *126*, 663–676). In fact, more recent experiments have shown that many cell types can be transformed to another cell type by expression of specific transcription factors. The discovery of cell-type-specific regulatory RNAs, however, challenges the concept that sequence-specific DNA binding factors are the "master" regulators of cell identity. Indeed, since most cells are stable units of functions through divisions, all regulatory factors must interact in feed-forward and feedback systems to maintain a homeostatic state. One major challenge of the coming decades is to integrate the system of interactions between various types and layers of regulatory factors to model this homeostatic state. Consistent with the model of Jacob and Monod, it would probably be wise to start to analyze these complex layers of regulatory factors beginning at recognition of genomic sequences.

Human Development and Disease

Richard A. Young
Whitehead Institute for Biomedical Research

The major challenge I see for molecular biology is to learn how human development and disease are programmed in the genome. This will require that we discover how cell state is controlled in hundreds of cell types. How transcription factors, chromatin regulators, noncoding RNAs, and signaling pathways contribute to the regulatory circuitry of cells is a fundamental and challenging problem, and knowledge of this circuitry is certain to give us new insights into disease processes. Reprogramming experiments tell us that a small number of master transcription factors can establish and maintain cell state, so for at least some cell types, it seems possible to tackle the core elements of this problem in the near future.

As more human genomes are sequenced, there is increased interest in the contribution of sequence variation to human health and disease. Although alterations in the levels and functions of transcription factors and other regulators are known to contribute to many diseases, little is known about the effect of sequence variation on the functions of these regulators. Thus, new approaches are needed to understand how variation contributes to altered regulatory circuitry, both within and between cells, and to learn when this variation contributes to disease.

Meaningful Networks

Uri Alon
Weizmann Institute of Science

My favorite mystery is how networks of billions of communicating cells—such as the immune system and central nervous system—generate meaning. How do they make sense of the world and generate proper responses to unforeseen contexts? It is not clear yet whether networks on this scale can be understood by human beings. Whereas the molecular networks in individual cells are beginning to be understood as integrated circuitry made of recurring circuit elements with defined functions, it is not clear if one can define analogous circuits on the level of communicating cells. Understanding immune and nervous systems, and even characterizing their hugely complex states, may require concepts and experimental methods beyond the current horizon.

As a profession, I believe that our main challenge is the increasing competition and isolation of researchers. I hope that this can be addressed by the rising movement to add education and discussion of the subjective and emotional aspects of doing science: how people can fruitfully mentor, communicate, choose problems, and give and receive feedback. I don't mean vacuously being nice to each other; I mean interacting to reach our full potential as scientists—following our curiosity to figure out how natural systems work.