Cell, Vol. 121, 503–504, May 20, 2005, Copyright ©2005 by Elsevier Inc. DOI 10.1016/j.cell.2005.05.005

The Meaning of Systems Biology

Commentary

Marc W. Kirschner*
Department of Systems Biology
Harvard Medical School
Boston, Massachusetts 02115

With the new excitement about systems biology, there is understandable interest in a definition. This has proven somewhat difficult. Scientific fields, like species, arise by descent with modification, so in their earliest forms even the founders of great dynasties are only marginally different than their sister fields and species. It is only in retrospect that we can recognize the significant founding events. Before embarking on a definition of systems biology, it may be worth remembering that confusion and controversy surrounded the introduction of the term "molecular biology," with claims that it hardly differed from biochemistry. Yet in retrospect molecular biology was new and different. It introduced both new subject matter and new technological approaches, in addition to a new style.

As a point of departure for systems biology, consider the quintessential experiment in the founding of molecular biology, the one gene one enzyme hypothesis of Beadle and Tatum. This experiment first connected the genotype directly to the phenotype on a molecular level, although efforts in that direction can certainly be found in the work of Archibald Garrod, Sewell Wright, and others. Here a protein (in this case an enzyme) is seen to be a product of a single gene, and a single function; the completion of a specific step in amino acid biosynthesis is the direct result. It took the next 30 years to fill in the gaps in this process. Yet the one gene one enzyme hypothesis looks very different to us today. What is the function of tubulin, of PI-3 kinase or of rac? Could we accurately predict the phenotype of a nonlethal mutation in these genes in a multicellular organism? Although we can connect structure to the gene, we can no longer infer its larger purpose in the cell or in the organism. There are too many purposes; what the protein does is defined by context. The context also includes a history, either developmental or physiological. Thus the behavior of the Wnt signaling pathway depends on the previous lineage, the "where and when" questions of embryonic development. Similarly the behavior of the immune system depends on previous experience in a variable environment. All of these features stress how inadequate an explanation for function we can achieve solely by trying to identify genes (by annotating them!) and characterizing their transcriptional control circuits.

That we are at a crossroads in how to explore biology is not at all clear to many. Biology is hardly in its dotage; the process of discovery seems to have been perfected, accelerated, and made universally applicable to all fields of biology. With the completion of the human genome and the genomes of other species, we have a glimpse of many more genes than we ever had before to study. We are like naturalists discovering a new continent, enthralled with the diversity itself. But we have also at the same time glimpsed the finiteness of this list of genes, a disturbingly small list. We have seen that the diversity of genes cannot approximate the diversity of functions within an organism. In response, we have argued that combinatorial use of small numbers of components can generate all the diversity that is needed. This has had its recent incarnation in the simplistic view that the rules of *cis*-regulatory control on DNA can directly lead to an understanding of organisms and their evolution. Yet this assumes that the gene products can be linked together in arbitrary combinations, something that is not assured in chemistry. It also downplays the significant regulatory features that involve interactions between gene products, their localization, binding, posttranslational modification, degradation, etc. The big question to understand in biology is not regulatory linkage but the *nature of biological systems that allows them to be linked together* in many nonlethal and even useful combinations. More and more we come to realize that understanding the conserved genes and their conserved circuits will require an understanding of their special properties that allow them to function together to generate different phenotypes in different tissues of metazoan organisms. These circuits may have certain robustness, but more important they have adaptability and versatility. The ease of putting conserved processes under regulatory control is an inherent design feature of the processes themselves. Among other things it loads the deck in evolutionary variation and makes it more feasible to generate useful phenotypes upon which selection can act.

Systems biology offers an opportunity to study how the phenotype is generated from the genotype and with it a glimpse of how evolution has crafted the phenotype. One aspect of systems biology is the development of techniques to examine broadly the level of protein, RNA, and DNA on a gene by gene basis and even the posttranslational modification and localization of proteins. In a very short time we have witnessed the development of high-throughput biology, forcing us to consider cellular processes in toto. Even though much of the data is noisy and today partially inconsistent and incomplete, this has been a radical shift in the way we tear apart problems one interaction at a time. When coupled with gene deletions by RNAi and classical methods, and with the use of chemical tools tailored to proteins and protein domains, these high-throughput techniques become still more powerful.

High-throughput biology has opened up another important area of systems biology: it has brought us out into the field again or at least made us aware that there is a world outside our laboratories. Our model systems have been chosen intentionally to be of limited genetic diversity and examined in a highly controlled and reproducible environment. The real world of ecology, evolution, and human disease is a very different place. When genetics separated from the rest of biology in the early

*Correspondence: marc@hms.harvard.edu

part of the 20th century, most geneticists sought to understand heredity and chose to study traits in the organism that could be easily scored and could be used to reveal genetic mechanisms. This was later extended to powerful effect to use genetics to study cell biological and developmental mechanisms. Some geneticists, including a large school in Russia in the early 20th century, continued to study the genetics of natural populations, focusing on traits important for survival. That branch of genetics is coming back strongly with the power of phenotypic assays on the RNA and protein level. As human beings we are most concerned not with using our genetic misfortunes to unravel biology's complexity (important as that is) but with the role of our genetics in our individual survival. The context for understanding this is still not available, even though the data are now coming in torrents, for many of the genes that will contribute to our survival will have small quantitative effects, partially masked or accentuated by other genetic and environmental conditions. To understand the genetic basis of disease will require not just mapping these genes but an understanding of how the phenotype is created in the first place and the messy interactions between genetic variation and environmental variation.

I find myself personally drawn to another part of systems biology. It is a smaller scale view, totally compatible with and partially dependent on the global analysis of high-throughput biology. This view spans in vitro biochemistry to what is now called synthetic biology and it has as its goal the reconstruction and description of partial but complex systems. I sometimes wonder if I have witnessed almost the entire molecular biology revolution only to emerge unscathed as an unrepentant and unrehabilitated biochemist. One particular experience may have foreshadowed this interest, the work with Tim Mitchison on microtubule assembly. The unusual dynamics of this polymer was a lesson in the importance of variation and selection in the spatial realm morphogenesis. Here was a simple adaptive system able to support structure throughout biology and throughout the evolution of eukaryotic cells. The phenotype of microtubules could not be understood by studying tubulin on its own, except in a very limited way. Yet although the biochemistry of tubulin is simple and conserved, it provided insight into the special unconstrained character of biology. Later John Gerhart and I tried to think about the connections among cell biology, biochemistry, development, and evolution. We looked at biology in terms of conserved processes and circuits and asked what features were selected and what changes occurred in evolution. We summarized our findings in a book that took almost a decade to write. By that time I had unconsciously become a systems biologist, awaiting, I assume, merely social acceptance of the term.

The biochemical and synthetic aspects of systems biology can occur on many levels and use many techniques. One level is the reconstitution of complex processes from purified components and the study of the dynamical nature of those processes. On another level will be the development and study of extract systems that recapitulate fairly faithfully cellular processes or explant systems that recapitulate developmental events. Extracts and explants are relatively accessible to synthetic manipulation. Next there is the explicit reconstruction of circuits within cells or the deliberate modification of those circuits. This has occurred for a while in biology, but the difference is that now we wish to construct or intervene with the explicit purpose of describing the dynamical features of these synthetic or partially synthetic systems. There are more and more tools to intervene and more and more tools to measure. Although these fall short of total descriptions of cells and organisms, the detailed information will give us a sense of the special life-like processes of circuits, proteins, cells in tissues, and whole organisms in their environment. This meso-scale systems biology will help establish the correspondence between molecules and large-scale physiology.

You are probably running out of patience for some definition of systems biology. In any case, I do not think the explicit definition of systems biology should come from me but should await the words of the first great modern systems biologist. She or he is probably among us now. However, if forced to provide some kind of label for systems biology, I would simply say that systems biology is the study of the behavior of complex biological organization and processes in terms of the molecular constituents. It is built on molecular biology in its special concern for information transfer, on physiology for its special concern with adaptive states of the cell and organism, on developmental biology for the importance of defining a succession of physiological states in that process, and on evolutionary biology and ecology for the appreciation that all aspects of the organism are products of selection, a selection we rarely understand on a molecular level. Systems biology attempts all of this through quantitative measurement, modeling, reconstruction, and theory. Systems biology is not a branch of physics but differs from physics in that the primary task is to understand how biology generates variation. No such imperative to create variation exists in the physical world. It is a new principle that Darwin understood and upon which all of life hinges. That sounds different enough for me to justify a new field and a new name. Furthermore, the success of systems biology is essential if we are to understand life; its success is far from assured—a good field for those seeking risk and adventure.