A unifying view of 21st century systems biology

Marc Vidal

Center for Cancer Systems Biology (CCSB) and Department of Cancer Biology, Dana-Farber Cancer Institute, and Department of Genetics, Harvard Medical School, 44 Binney Street, Boston, MA 02115, USA

ARTICLE INFO

Article history:
Received 6 November 2009
Revised 10 November 2009
Accepted 10 November 2009
Available online 12 November 2009

Edited by Stefan Hohmann

Keywords:
Systems biology
Network biology
Interactome network

ABSTRACT

The idea that multi-scale dynamic complex systems formed by interacting macromolecules and metabolites, cells, organs and organisms underlie some of the most fundamental aspects of life was proposed by a few visionaries half a century ago. We are witnessing a powerful resurgence of this idea made possible by the availability of nearly complete genome sequences, ever improving gene annotations and interactome network maps, the development of sophisticated informatic and imaging tools, and importantly, the use of engineering and physics concepts such as control and graph theory. Alongside four other fundamental “great ideas” as suggested by Sir Paul Nurse, namely, the gene, the cell, the role of chemistry in biological processes, and evolution by natural selection, systems-level understanding of “What is Life” may materialize as one of the major ideas of biology.

© 2009 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

1. Introduction

As argued by Sir Paul Nurse [1], four widely accepted “great ideas” are inherent to our current view of “What is Life?” [2]. Modern biologists take for granted that: (i) the gene is the basis for heredity, (ii) the cell is the fundamental unit of organisms, (iii) biology is based on chemistry, and (iv) species evolve by natural selection. These four ideas are an integral part of how we think, teach, address biological problems, organize experiments, communicate our findings to the public, and design modern preventive medicine and therapeutic strategies. However, as powerful as reductionist models limited to individual genes, cells, and chemical reactions might have been throughout the 19th and 20th centuries, it is increasingly clear that they are insufficient to fully “explain” or “describe” life as we enter the 21st century.

The four ideas mentioned above each represent a different prism through which life can be understood. They each touch on a fundamental aspect of life. It is rather obvious that no life form can be imagined without genes, cells (viruses are not considered cells per se but require host cells for their parasitic life cycles), enzymes, or evolution by natural selection as the driving force of their emergence. In addition to, or as a consequence of, representing fundamental aspects of life, these four concepts gradually developed into distinct, specialized and widely accepted fields of biology: genetics and molecular biology, cell biology, biochemistry and evolutionary biology, respectively.

Systems biology represents a fifth prism through which life can be understood touching on a fifth fundamental aspect of life. Although certainly not obvious to everyone at this point and definitely requiring further formal proof, the idea of systems biology presupposes that no life form can be imagined without complex systems formed by interacting genes and macromolecules, or cells at a higher scale, and in the context of which natural selection operates. Gene products do not act alone, individual cells separated from their neighbors lose many of their functional and structural attributes, macromolecules and metabolites are intimately linked to each other. Importantly, evolution rarely acts on separate biochemical reactions, individual cells or distinct species, but rather, impinges upon complex multi-scale systems in which these components are intricately interconnected. At a recent Nobel Symposium organized near Stockholm, Sweden, and entitled “systems biology”, the scientists gathered in the beautiful setting of the Sånga-Säby Conference Center seemed to have converged on that very point. We will not completely understand biology until we fully embrace a “fifth great idea” that can be summarized as follows: multi-scale dynamic complex systems formed by interacting macromolecules and metabolites, cells, organs, and organisms underlie most biological processes.

This review, primarily targeted to non-specialists, attempts to summarize briefly how the concept of systems understanding of biology, or “systems biology”, proposed by a few visionaries half a century ago has re-emerged during this ending decade. Systems biology can be viewed as a unifying framework in which talks presented at this 2009 Nobel Symposium can easily be incorporated.
2. Systems biology

It is important to realize that the development of today’s systems biology concepts lay on theoretical and empirical grounds that go back half a century. Despite the often over-simplified *a posteriori* perception whereby the creation of a field relates to a single person or a single discovery, the acceptance of new fundamental ideas is often slow and depends upon the contributions of many scientists over many years. As examples, consider Mendel for the gene as the basis of heredity or Darwin for evolution by natural selection [3].

First and foremost, the idea of systems biology starts from basic intuition. When considered as separate entities, ~25 000 genes, ~10¹³ cells, or a few thousand enzymes fail to completely capture the mystery of life as wonderfully revealed by the example of a developing baby. Having closely observed our daughter Ava since she was born seven months ago, I can intuitively report that the four widely accepted great ideas of biology are not sufficient to explain such beauty. A gene list apparently not much longer than that of a round worm simply doesn’t account for how she grows, develops, learns, behaves, and how she is at once so resilient, yet so fragile [4].

Historically, a systems-based intuition of biology dates back two centuries when the word “organism”, originally used for “organic structure, organization,” was picked to refer to “living animals or plants”. Further capturing that intuition of organization or inter-connectedness, potentially at the heart of all biological phenomena, is the notion expressed by Kant at the end of the 18th century that “organisms are organized natural products in which every part is reciprocally both end and means”.

The idea of systems biology is of course also firmly grounded in a theoretical framework supported by increasingly sophisticated sets of empirical observations. Interestingly, one of the earliest and most inspirational statements one can find about cellular complexity was provided in a somewhat prophetetic abstract opening the famous Beadle and Tatum 1941 paper that first described the “one-gene/one-enzyme/one-function” hypothesis [5]. In that abstract, Beadle and Tatum express the need to understand “how an organism consists essentially of an integrated system of chemical reactions controlled in some manner by genes”. They go on to state that: “since the components of such system are likely to be interrelated in complex ways, it would appear that there must exist orders of directness of gene control ranging from simple one-to-one relations to relations of great complexity”. Clearly Beadle and Tatum, while launching a long tradition of reductionist one-gene-at-a-time molecular biology studies, were fully aware of the long-term implications of their work upon the discovery of most if not all genes and the functional interactions they mediate with each other.

A first theoretical example of how complexity can emerge in relatively small biological systems composed of a few macromolecules was proposed in the late 1940s by Delbrück [6] as an attempt to explain the phenomenon of differentiation. How can two cells harboring exactly the same genotype behave as differently as skin and retina cells? In this rather short but extremely powerful piece, Delbrück states that two molecular entities (e.g. retina cells? In this rather short but extremely powerful piece, Delbrück states that two molecular entities (e.g. retina cells? In this rather short but extremely powerful piece, Delbrück states that two molecular entities (e.g. retina cells?) in the basis of heredity or Darwin for evolution by natural selection [3].

The first empirical demonstration of this idea emerged from studies by Novick and Weiner, and later by Cohn and Horibata. The 1957 Novick and Weiner paper [7] represents a milestone of enormous consequence for today’s field of systems biology. Entitled “Enzyme induction as an all-or-none phenomenon” the paper shows how induction of lacZ-encoded β-galactosidase (β-Gal) triggered by a lactose analog is extremely rapid if one observes the response at the cellular level rather than for an entire cell population. Amazingly, it was also shown that, upon extreme reduction of the inducer concentration, down to a level far below what is required for activation at the first place, high levels of β-Gal are maintained for over 150 generations. A systems-level explanation accounting for these two observations was provided by the authors as follows: a lactose permease, which facilitates transport of lactose across the cell membrane, is induced along with β-Gal induction (the permease-encoding gene lacY is co-expressed concomitantly with lacZ as part of the lacZ operon), once the permease is induced the system enters a positive feedback circuit-type mode leading to the on state, which can last as long as traces of the analog are available in the culture.

Monod and Jacob summarized in another landmark paper entitled “General conclusions: teleonomic mechanisms in cellular mechanisms, growth and differentiation” [8] how, while positive feedback circuits are expected to be required for bistability, negative feedback circuits would be expected to be required for homeostasis and oscillatory phenomena. A few years prior, Umbarger, Pardee and others had observed early versions of negative feedback circuits with the phenomenon of enzymatic feedback inhibition [9,10]. The Monod and Jacob teleonomic arguments were subsequently formalized by Thomas and others in terms of theoretical requirements for positive and negative feedback circuits using Boolean modeling [11].

Extending from such “local” systems properties to more “global” complex networks views of larger numbers of interacting genes and gene products at the scale of whole cells, Waddington [12] proposed the notion of epigenetic landscape as a metaphor for the “trajectory” that a complex biological system might be traversing in response to genetic, developmental and/or environmental cues. This powerful concept, together with fascinating theoretical modeling of “randomly constructed genetic nets” by Kauffman [13], helps us to conceive how, at any given moment, a cellular system can be described in terms of “states” resulting from particular combinations of genes, gene products, or metabolites all considered either active or inactive at any given time. Complex wiring diagram-like sets of functional and logical interconnectivities between macromolecules acting upon each other were imagined to account for how systems “travel” from state to state over time throughout a “state space” determined by combinations of genotype and environment.

In summary, it was realized relatively early on and concomitantly with the development of the field of molecular biology that complex interconnections between macromolecules, both at local and global levels, might be able to generate systems properties or behaviors that would ultimately be recognized and understood as fundamental to life. It seemed that by the 1970s the available framework for a systems understanding of biology was in place. However, it would take several decades before such a notion could mature into a solid field of scientific research. Amazingly, some of these early concepts were proposed before we even knew the molecular nature of the gene.

Before systems biology could develop fully, molecular biology had first to come to maturity in its own right. Before gene–gene, gene–protein, or protein–protein based systems could be studied, one first needed to learn how to isolate, sequence, manipulate, and perturb genes, exogenously express proteins, characterize their biochemical activities and determine their structural
properties. Before the need for a systems understanding would be more widely accepted, nearly complete gene lists were needed for organisms of interest. It took a few decades for systems biology concepts to reach critical mass and for a robust field of research to emerge. At the dawn of the 21st century, most of the basic ideas of systems biology remained to be proven and/or applied to relevant problems.

3. The “emerging” field of systems biology

Standing on the shoulders of the giants mentioned above and many others, we are now witnessing a “renaissance” of local and global systems biology understanding, powered by the introduction of modern molecular and imaging technologies as well as mathematics, physics and engineering concepts and modeling strategies. Importantly, systems properties have been discovered in the context of the cell cycle, signaling, apoptosis, circadian clocks and are likely to become relevant for stem cells, wide association studies and the microRNA world. One could almost argue that this resurgence throughout the first decade of the 21st century is, in some ways, analogous to the evolution of the concept of the gene as the basis for heredity in the 1900s. Approximately four decades after the 1866 Mendel paper, now famous but completely unnoticed at the time, seemingly unrelated observations published by de Vries, von Tschermak, Correns, Cuénot, Johannsen, Bateson, Garrod and others converged toward a single unifying theme over the course of the very first decade of the 20th century: the laws of heredity, apparently universally applicable, involve hereditary units interchangeably referred to over time as “gemmules”, “factors”, “elements”, “determinants”, “pangenes”, “mnémons” and finally, as “genes”, a “simplifying, to-the-point, and appealingly short word”, as proposed by Johannsen.

As new fields develop, new fundamental concepts need to be formulated, usually leading to lists of neologisms that are sometimes perceived as annoying at first by scientists not directly involved, but that eventually gain acceptance if determined useful. Consider that while the words “gene” and “genetics” were coined in the 1900s, “genomics” only appeared as recently as in the turn of this century. Approximately four decades after the 1866 Mendel paper, now famous but completely unnoticed at the time, seemingly unrelated observations published by de Vries, von Tschermak, Correns, Cuénot, Johannsen, Bateson, Garrod and others converged toward a single unifying theme over the course of the very first decade of the 20th century: the laws of heredity, apparently universally applicable, involve hereditary units interchangeably referred to over time as “gemmules”, “factors”, “elements”, “determinants”, “pangenes”, “mnémons” and finally, as “genes”, a “simplifying, to-the-point, and appealingly short word”, as proposed by Johannsen.

As new fields develop, new fundamental concepts need to be formulated, usually leading to lists of neologisms that are sometimes perceived as annoying at first by scientists not directly involved, but that eventually gain acceptance if determined useful. Consider that while the words “gene” and “genetics” were coined in the 1900s, “genomics” only appeared as recently as in the 1980s. In many instances new metaphors need to be invented to reflect new fundamental findings or novel hypotheses, as in “genetic code”, “transcription” and “translation” for molecular biology. Accordingly, biologists along with mathematicians, physicists and engineers are in the process of introducing their own vocabulary with old and newer notions such as positive and negative “feedback circuits”, “feedforward loops” [14], “interactome” networks [15], “scale-free” [16] and “small-world” [17] networks, “hubs” [18], “date and party hubs” [19]. Although it remains unclear at this stage which of these concepts and terms will sustain the test of time, they illustrate nicely the creative forces of our burgeoning field.

With increasing challenges and novel questions emerging, new technologies and methodologies need to be invented. Conversely, new technologies often trigger new ways of addressing biological problems. Such technologies are often confused with the fundamentals of the concepts they are designed to address. For example, molecular biology is more often then not understood as a set of technologies such as DNA restriction, ligation, cloning and sequencing, while, fundamentally, it is a field that addresses life through the prism of the properties of macromolecules. Likewise, genetics is sometimes understood as a set of techniques such as dissecting yeast tetrads, pushing flies or worms, etc., while it really represents the study of life from the angle of what heredity can teach us. A similar confusion applies to systems biology. The fundamental aspect of systems biology is often blurred together with experimental strategies. There is a common confusion with the term “systematic” biology, which usually refers to studying a molecular biological problem by systematically testing all genes or proteins of an organism as a way to identify all macromolecules involved in that particular process. In addition, systems biology is too often thought of as a restricted set of specific modeling approaches, such as the supposedly required use of differential equations or any other specific mathematics-based methodology, while what really matters are the fundamental biological aspects that our newly emergent field represents.

4. Twenty first century systems biology

By all accounts, systems biology is thus becoming a mature field of biology, the goal of which is to understand how complex systems underlie life. As often happens in science, multiple seemingly unrelated events took place during a relatively short time at the turn of this century.

Just in the opening month of the decade that followed the scare about the Y2 K bug, a series of papers appeared with some of the major elements of today’s systems approach to biology. Two publications demonstrated fundamental properties of positive and negative feedback circuits by virtue of synthetic molecular reconstruction of de novo designed and modeled wiring diagrams [20,21], while a third similar story appeared shortly thereafter [22]. A differential equation-based quantitative model of the budding yeast cell cycle control was released [23]. And, on the other side of the systems biology spectrum, two papers published that same month confirmed the high level of macromolecular interconnectivity in proteome-wide protein–protein interaction or “interactome” networks, and strongly suggested that obtaining high-quality interactome maps was possible and would be extremely useful for cellular network modeling [24,25]. Such a serendipitous convergence of events in a single month is mentioned here to illustrate the wide diversity of systems biology approaches.

Graph theory papers describing global properties of non-biological networks appeared [26,27] at about the same time, which would turn out to be critical to understand overall organization features of interactome networks [18,28]. Modeling regulatory circuits [29–31] and larger scale networks [32,33], reconstructing metabolic networks [34], investigating network robustness [35,36] and stochastic gene expression in biological systems [37], were among the necessary conditions for a full reemergence of systems biology.

This issue of *FEBS Letters* nicely illustrates some more recent contributions to a long list of accomplishments, namely: (i) feedback mechanisms are of the highest importance in cell cycle
regulation, signaling, circadian oscillators and other regulatory networks [38–43]; (ii) evolutionary mechanisms can be better understood in light of complex molecular systems [44,45]; (iii) interactome networks together with genetic, transcriptome and metabolic networks contribute to our global understanding of cellular systems [46–48]; (iv) characterization of the components of biological systems, their regulations and molecular functions remains a great challenge [49,50]; and finally (v) increasingly sophisticated modeling concepts remain to be developed before the promise of systems biology can be fully realized [51–53].

Most of the participants of this 2009 Systems Biology Nobel Symposium would agree at this stage that there are indeed many facets to a systems understanding of life: mapping and modeling macro- and cellular networks, understanding cellular organization, information flow and logical relationships, mathematical modeling of local molecular relationships, and importantly de novo engineering of molecular circuits with predefined systems properties as exemplified by the growing field of “synthetic biology”. What appears to have emerged from the Symposium is how seemingly different approaches appear to be bound by a single common thread: further developing the 5th great idea of biology.

Acknowledgements

The work performed at the DFCI Center for Cancer Systems Biology (CCSB) is funded by grants from the Keck and Ellison Foundations, NHGRI, NIGMS, NCI, NSF, and DOE. Special thanks to M. Cusick, D. Hill, A.-R. Carvunis, B. Charloteaux, E. Rietman and R. Vidal for their help with the manuscript.

References