

Question & Answer

Q&A: Systems biology

James E Ferrell Jr

What is systems biology?

Systems biology is the study of complex gene networks, protein networks, metabolic networks and so on. The goal is to understand the design principles of living systems.

How complex are the systems that systems biologists study?

That depends. Some people focus on networks at the 'omics'-scale: whole genomes, proteomes, or metabolomes. These systems can be represented by graphs with thousands of nodes and edges (see Figure 1). Others focus on small subcircuits of the network; say a circuit composed of a few proteins that functions as an amplifier, a switch or a logic gate. Typically, the graphs of these systems possess fewer than a dozen (or so) nodes. Both the large-scale and small-scale approaches have been fruitful.

Why is systems biology important?

Stas Shvartsman at Princeton tells a story that provides a good answer to this question. He likens biology's current status to that of planetary astronomy in the pre-Keplerian era. For millennia people had watched planets wander through the night-time sky. They named them, gave them symbols, and charted their complicated comings and goings. This era of descriptive planetary astronomy culminated in Tycho Brahe's careful quantitative studies of planetary motion at the end of the 16th century. At this point planetary motion had been described but not understood.

Then came Johannes Kepler, who came up with simple theories (ellipti-

cal heliocentric orbits; equal areas in equal times) that empirically accounted for Brahe's data. Fifty years later, Newton's law of universal gravitation provided a further abstraction and simplification, with Kepler's laws following as simple consequences. At that point one could argue that the motions of the planets were understood.

Systems biology begins with complex biological phenomena and aims to provide a simpler and more abstract framework that explains why these events occur the way they do. Systems biology can be carried out in a 'Keplerian' fashion - look for correlations and empirical relationships that account for data - but the ultimate hope is to arrive at a 'Newtonian' understanding of the simple principles that give rise to the complicated behaviors of complex biological systems.

Note that Kepler postulated other less-enduring mathematical models of planetary dynamics. His *Mysterium Cosmographicum* showed that if you nest spheres and Platonic polyhedra in the right order (sphere-octahedron-sphere-icosahedron-sphere-dodecahedron-sphere-tetrahedron-sphere-cube-sphere), the sizes of the spheres correspond to the relative sizes of the first six planets' orbits. This simple, abstract way of accounting for empirical data was probably just a happy

James E Ferrell Jr, Department of Chemical and Systems Biology, Stanford University School of Medicine, Stanford, CA 94305-5174, USA. Email: james.ferrell@stanford.edu

coincidence. Happy coincidences are a potential danger in systems biology as well.

Is systems biology the antithesis of reductionism?

In a limited sense, yes. Some 'emerging properties', as discussed below, disappear when you reduce a system to its individual components.

However, systems biology stands to gain a lot from reductionism, and in this sense systems biology is anything but the antithesis of reductionism. Just as you can build up to an understanding of complex digital circuits by studying individual electronic components, then modular logic gates, and then higher-order combinations of gates, one may well be able to achieve an understanding of complex biological systems by studying proteins and genes, then motifs (see below), and then higher-order combinations of motifs.

What are emergent properties?

Systems of two proteins or genes can do things that individual proteins/genes cannot. Systems of ten proteins or genes can do things that systems of two proteins/genes cannot. Those things that become possible once a system reaches some level of complexity are termed emergent properties.

Can you give a concrete example of an emergent property?

Three proteins connected in a simple negative-feedback loop ($A \rightarrow B \rightarrow C \rightarrow A$) can function as an oscillator; two proteins ($A \rightarrow B \rightarrow A$) cannot. Two

proteins connected in a simple negative-feedback loop can convert constant inputs into pulsatile outputs; a one-protein loop ($A \dashv A$) cannot. So pulse generation emerges at the level of a two-protein system and oscillations emerge at the level of a three-protein system.

In systems biology there is a lot of talk about nodes and edges. What is a node? An edge?

Biological networks are often depicted graphically: for example, you could draw a circle for protein A, a circle for protein B, and a line between them if A regulates B or vice versa. The circles are the nodes in the graph of the A/B system. Nodes can represent genes, proteins, protein complexes, individual states of a protein, and so on.

A line connecting two nodes is an edge. The edge can be directed: for example, if A regulates B, we write an arrow - a directed edge - from A to B, whereas if B regulates A we write an arrow from B to A. Or the edge can be undirected; for example, it represents a physical interaction between A and B.

Staying with graphs, what's a motif?

As defined by Uri Alon, a motif is a statistically over-represented sub-graph of a graphical representation of a network. Motifs include things like negative feedback loops, positive feedback loops, and feed-forward systems.

Isn't positive feedback the same thing as feed-forward regulation?

No. They are completely different. In a positive-feedback system, A activates B and B turns around to activate A. A transitory stimulus that activates A could lock the system into a self-perpetuating state where both A and B are active. In this way, the positive-feedback loop can act like a toggle switch or a flip-flop. A positive-feedback loop

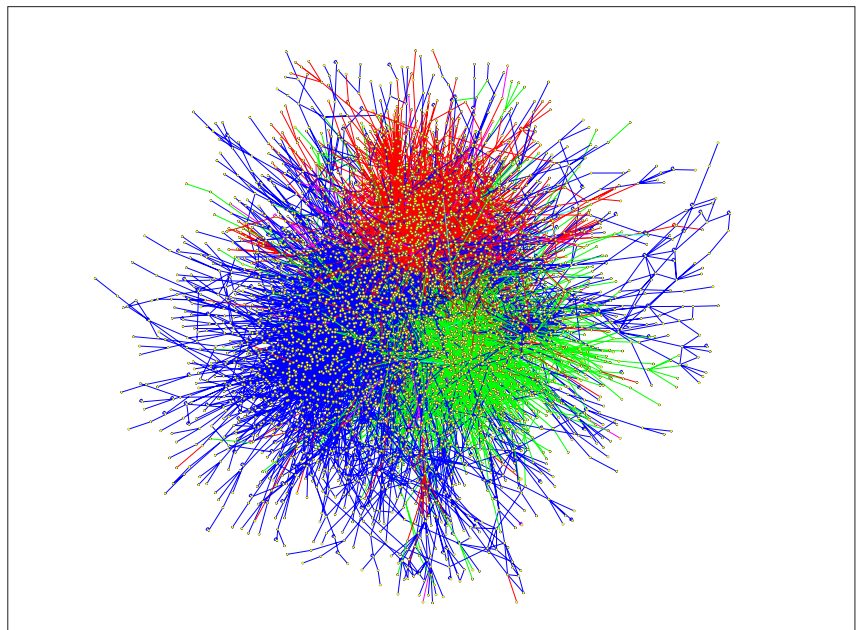


Figure 1

Human protein-protein interaction network. Proteins are shown as yellow nodes. Interactions from CCSB-HII (Rual *et al.*, *Nature* 2005, **437**:1173-1178) and from (Stelzl *et al.*, *Cell* 2005, **122**:957-968) are shown as red and green edges, respectively. Literature-Curated Interactions (LCI) extracted from databases (BIND, DIP, HPRD, INTACT and MINT) that are supported by at least 2 publications are shown as blue edges. Interactions common to two of those 3 datasets are represented with the corresponding mixed color (yellow for (Rual *et al.*, 2005) and (Stelzl *et al.*, 2005), magenta for Rual and LCI, cyan for (Stelzl *et al.*, 2005) and LCI). Interactions common to all 3 datasets are shown as black edges. (Figure kindly provided by Nicolas Simonis and Marc Vidal.)

behaves much like a double-negative feedback loop, where A and B mutually inhibit each other. That system can act like a toggle switch too, except that it toggles between A on/B off and A off/B on states, rather than between A off/B off and A on/B on states. Good examples of this type of system include the famous lambda phage lysis/lysogeny toggle switch, and the CDK1/Cdc25/Wee1 mitotic trigger.

In a feed-forward system, A impinges upon C directly, but A also regulates B, which regulates C. A feed-forward system can be either 'coherent' or 'incoherent', depending upon whether the route through B does the same thing to C as the direct route does. There is no feedback - A affects C, but C does not affect A - and the system

cannot function as a toggle switch. A good example of feed-forward regulation is the activation of the protein kinase Akt by the lipid second messenger PIP3 (PIP3 binds Akt, which promotes Akt activation, and PIP3 also stimulates the kinase PDK1, which phosphorylates Akt and further contributes to Akt activation). Since both routes contribute to Akt activation, this is an example of coherent feed-forward regulation. Uri Alon's classic analysis of motifs in *Escherichia coli* gene regulation identified numerous coherent feed-forward circuits in that system.

In high school I hated physics and math, but I loved biology. Should I go into systems biology?

No.

What kind of physics and math is most useful for understanding biological systems?

Some level of comfort in doing simple algebra and calculus is a must. Beyond that, probably the most useful math is nonlinear dynamics. The Strogatz textbook mentioned below is a great introduction to nonlinear dynamics.

Do I need to understand differential equations?

Systems biologists often model biological processes with ordinary differential equations (ODEs), but the fact is that almost none of them can be solved exactly. (The one that can be solved exactly describes exponential approach to a steady state, and it's something every biologist should work out at some point in his or her training.) Most often, systems biologists solve their ODEs numerically, often with canned software packages like Matlab or Mathematica.

Ideally, a model should not only reproduce known biology and predict unknown biology, it should also be 'robust' in important respects.

What is robustness, and why is it important to systems biologists?

Robustness is the imperviousness of some performance characteristic of a system in the face of some sort of insult - such as stochastic fluctuations, environmental insults, or deletion of nodes from the system. For example, the period of the circadian oscillator is robust with respect to changes in the temperature of the environment. Robustness can be quantitatively defined as the inverse of sensitivity, which itself can be defined a few ways - often sensitivity is taken to be:

$$\frac{d\ln Response}{d\ln Perturbation}$$

so that robustness becomes

$$\frac{d\ln Perturbation}{d\ln Response}$$

Robustness is important to systems biologists because of the attractiveness of the idea that a biological system must function reliably in the face of myriad uncertainties. Maybe robustness, more than efficiency or speed, is what evolution must optimize to create successful biological systems.

Modeling can provide some insight into the robustness of particular networks and circuits. Just as a biological system must be robust with respect to insults the system is likely to encounter, a successful model should also be robust with respect to parameter choice. If a model 'works', but only for a precisely chosen set of parameters, the system it depicts may be too finicky to be biologically useful, or to have been 'found' in evolution.

What other types of models are useful in systems biology?

ODE models assume that each dynamical species in the model - each protein, protein complex, RNA, or whatever - is present in large numbers. This is sometimes true in biological systems. For example, regulatory proteins are often present at concentrations of 10 to 1,000 nM. For a four picoliter eukaryotic cell, this corresponds to 24,000 to 2,400,000 molecules per cell. This is probably large enough to warrant ODE modeling. However, genes and some mRNAs are present at concentrations of one or two molecules per cell. At such low numbers, each individual transcriptional event or mRNA degradation event becomes a big deal, and the appropriate type of modeling is stochastic modeling.

Sometimes systems are too complicated, or have too many unknown parameters to warrant ODE modeling. In these cases, Boolean models and

probabilistic Bayesian models can be particularly useful.

Sometimes it is important to see how dynamical behaviors propagate through space, in which case either partial differential equation (PDE) models or stochastic reaction/diffusion models may be just the ticket.

Where can I go for more information?

Review articles

- Hartwell LH, Hopfield JJ, Leibler S, Murray AW: **From molecular to modular biology.** *Nature* 2005, **402(Suppl):C47-C52**.
 Kirschner M: **The meaning of systems biology.** *Cell* 2005, **121**:503-504.
 Kitano H: **Systems biology: a brief overview.** *Science* 2002, **295**:1662-1664.

Textbooks

- Alon U: *An Introduction to Systems Biology: Design Principles of Biological Circuits.* Boca Raton, FL: Chapman & Hall/CRC; 2006.
 Heinrich R, Schuster S: *The Regulation of Cellular Systems.* Berlin: Springer; 1996.
 Klipp E, Herwig R, Kowald A, Wierling C, Lehrach H: *Systems Biology in Practice: Concepts, Implementation and Application.* Weinheim, Germany: Wiley-VCH; 2005.
 Palsson B: *Systems Biology: Properties of Reconstructed Networks.* Cambridge University Press; 2006.
 Strogatz SH: *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry and Engineering.* Boulder, CO: Westview Press; 2001.

Published: 26 January 2009

Journal of Biology 2009, **8**:2
 (doi:10.1186/jbiol107)

The electronic version of this article is the complete one and can be found online at <http://jbiol.com/content/8/1/2>

© 2009 BioMed Central Ltd