

Systems biology 101—what you need to know

Trey Ideker

Systems biology has spurred interest in thousands of researchers, some just starting their careers, others well established but interested in learning more about it. What is the best plan for scientists and students interested in a career in systems biology?

Why the excitement?

The use of systematic genomic, proteomic and metabolomic technologies to construct models of complex biological systems and diseases is becoming increasingly commonplace. These endeavors, collectively known as systems biology^{1,2}, establish an approach by which to interrogate and iteratively refine our knowledge of the cell. In so doing, systems biology integrates knowledge from diverse biological components and data into models of the system as a whole.

Although the notion of systems science has existed for some time³, these approaches have recently become far more powerful because of a host of new experimental technologies that are high-throughput, quantitative and large-scale⁴. As evidence of the impact 'systems' thinking has had on biology, consider the explosive growth of new research institutes, companies, conferences and academic departments that have the words 'systems biology' in the title or mission statement. Several journals are now either entirely devoted to reporting systems biology research or are sponsoring regular sections devoted to current issues in systems or computational biology, such as this inaugural section in *Nature Biotechnology*. And under the leadership of Elias Zerhouni, the National Institutes of Health (Bethesda, MD, USA) has released a new 'roadmap' that includes interdisciplinary science and integrative systems biology as core focus areas⁵; the UK's Biotechnology and Biological Sciences Research Council has also targeted predictive and integrated biology as a strategic aim over the next five years⁶.

*Trey Ideker is at the Department of Bioengineering, University of California, San Diego, La Jolla, CA 92093-0412, USA.
e-mail: trey@bioeng.ucsd.edu*

Where to start

Because of the need to couple computational analysis techniques with systematic biological experimentation, more and more universities are offering PhD programs that integrate both computational and biological subject matter (Table 1). Several of these programs, such as those recently initiated by the Massachusetts Institute of Technology (MIT, Cambridge, MA, USA) and Harvard University (Cambridge, MA, USA), include 'systems biology' directly in the name. Others offer courses of study from within physics, engineering or biology departments (e.g., the systems biology syllabus within the bioengineering department at the University of California, San Diego, CA, USA).

Apart from PhD programs with course offerings in systems biology, a number of institutions offer intensive short courses (Table 2). These include the Institute for Systems Biology (Seattle, WA, USA), Oxford University (Oxford, UK) and Biocentrum Amsterdam (Holland). There are also several other emerging initiatives and educational programs around the globe (Table 3).

Given the pace of the field, it is probably too early to endorse one particular syllabus as the correct or best option. However, clearly all programs must provide a rigorous understanding of both biology and quantitative modeling. Thus, many require that all students, regardless of background, perform hands-on research in both computer programming and in the wet laboratory. Required course work in biology typically includes genetics, biochemistry, molecular and cell biology, with laboratory work associated with each of these. Course work in quantitative modeling might include probability, statistics, information theory, numerical optimization, artificial intelligence and machine learning, graph and network theory, and nonlinear dynamics. Of the biological course work, genetics is particularly important, because the logic of genetics is, to a large degree, the logic of systems biology. Of the course work in quantitative modeling, graph theory and machine learning techniques are of particular interest, because systems approaches often reduce cellular function to a search on a network of

biological components and interactions^{7,8}. A course of study integrating life and quantitative sciences helps students to appreciate the practical constraints imposed by experimental biology and to effectively tailor research to the needs of the laboratory biologist. At the same time, knowledge of the major algorithmic techniques for analysis of biological systems will be crucial for making sense of the data.

Other paths

An alternative to pursuing a cross-disciplinary program is to tackle one field initially and then learn another in graduate school. Examples would include choosing an undergraduate major in engineering and then obtaining a PhD in molecular biology, or starting within biochemistry then pursuing course work in computer programming. This leads to a common question: when contemplating a transition, is it better to switch from quantitative sciences to biology or vice versa?

Although some feel that it is easier to move from engineering into biology, the honest answer is that either trajectory can work. Some practical advice is that if coming from biology, start by becoming familiar with Unix, Perl and Java before diving into more complex computational methodologies. If coming from the quantitative sciences, jump into a wet laboratory as soon as possible—when shaky hands become steady, you're well on your way.

The job market

What jobs are new systems biologists likely to find? With the formation of myriad new academic departments and centers, the academic job market is booming. On the other hand, biotech firms and 'big pharma' have been more cautious about getting involved⁹. However, most agree that in the long-term, systems approaches promise to influence drug development in several areas: first, target identification, in which drugs are developed to target a specific molecule or molecular interaction within a pathway; second, prediction of drug mechanism-of-action (MOA), in which a compound has known therapeutic effects but the molecular

Table 1 Selected programs offering systems biology courses

Institution	URL
Asia	
A*Star Bioinformatics Institute, Singapore	http://www.bii.a-star.edu.sg/
University of Tokyo, Japan, Graduate School of Information Science and Technology	http://www.i.u-tokyo.ac.jp/index-e.htm
Europe	
Flanders and Ghent University, Belgium, Department of Plant Systems Biology	http://www.psb.ugent.be/
Max Planck Institute of Molecular Genetics, Berlin, Germany	http://lectures.molgen.mpg.de/
Max Planck Institute of Dynamics of Complex Systems, Madgeburg, Germany	http://www.mpi-magdeburg.mpg.de/
University of Rostock, Germany, Systems Biology & Bioinformatics	http://www.sbi.uni-rostock.de/
University of Stuttgart, Germany, Systems Biology Group	http://www.sysbio.de/
North America	
Cornell, Memorial Sloan-Kettering Hospital, and Rockefeller Universities, New York, Physiology, Biophysics & Systems Biology, Program in Computational Biology and Medicine	http://www.cs.cornell.edu/grad/cbm/
Harvard University, Cambridge, MA, Department of Systems Biology	http://sysbio.med.harvard.edu/
Massachusetts Institute of Technology, Cambridge, MA, Computational and Systems Biology Initiative	http://csbi.mit.edu/
Princeton University, Princeton, NJ, Lewis-Sigler Institute for Integrative Genomics	http://www.genomics.princeton.edu
Stanford University, Stanford, CA, Medical Informatics (SMI) and BioX	http://smi-web.stanford.edu/
University of California, Berkeley, CA, Graduate Group in Computational & Genomic Biology	http://cb.berkeley.edu/
University of California, San Diego, CA, Department of Bioengineering	http://www-bioeng.ucsd.edu/
University of Toronto, Canada, Program in Proteomics and Bioinformatics	http://www.utoronto.ca/medicalgenetics/
University of Washington, Seattle, WA, Department of Genome Sciences	http://www.gs.washington.edu/
Virginia Polytechnic Institute and State University, Blacksburg, VA, Program in Genetics, Bioinformatics and Computational Biology	http://www.grads.vt.edu/gbcb/phd_gbcb.htm
Washington University, St. Louis, MO, Computational Biology Program	http://www.ccb.wustl.edu/

mechanisms by which it achieves these effects are unclear; and third, prediction of drug toxicity and properties related to absorption, distribution, metabolism and excretion. In all of these cases, the key contribution of systems biology would be a comprehensive understanding blueprint of cellular pathways used for identifying proteins at key pathway control points, or proteins for which the predicted perturbation phenotypes most closely resemble those observed experimentally with a pharmacologic or toxic agent.

It is revealing that outside of biology and biotech, many sectors of manufacturing already depend heavily on computer simulation and modeling for product development. Using computer-aided design tools, digital circuit manufacturers explore the wiring of transistors and other components on the silicon wafer, just as automotive engineers estimate how many miles per gallon to expect from the next-generation sedan long before it is built on the assembly line. Biology will undoubtedly also benefit from these 'classical' engineering approaches. Given that

more than six out of every seven drugs that undergo human testing ultimately fail due to unanticipated side effects, software simulations may act as a much needed filter between high-throughput screening for drug candidates and the time-consuming and costly follow-up of human trials.

Thus, the key question may not be whether systems biology will be important

for the biotech industry, but how quickly it will take hold⁹. To effect real changes, what systems biology needs most is a major 'success story' within industry—a well-publicized example of how systems biology principles, experimentation and/or integrative modeling have led directly to selection of a lead compound or in pushing a drug to market.

Table 2 Selected short courses in systems biology

Institute/Program	URL
Berlin Graduate Program, Dynamics & Evolution of Cellular and Macromolecular Processes	http://www.biologie.hu-berlin.de/
Biocentrum Amsterdam, Molecular Systems Biology Course	http://www.science.uva.nl/biocentrum/
Cold Spring Harbor Laboratory, Course in Computational Genomics	http://meetings.cshl.org/
Institute of Systems Biology, Introduction to Systems Biology	http://www.systemsbiology.org/
University of Oxford, UK, Biology & Proteomics Informatics courses; Genomics, Proteomics & Beyond	http://www.conted.ox.ac.uk/cpd/biosciences/courses/short_courses/Genome_Analysis.asp

Already, numerous biotechnology start-ups such as Genstruct, Genomatica, Beyond Genomics, Ingenuity, Gene Network Sciences, Entelos, Bioseek and GeneGO are hard at work on developing just such 'proof-of-principles.' As for big pharma, systems concepts are influencing many companies to evolve their drug development process, but it is unclear whether what evolves will at all resemble what is currently called 'systems biology'.

A cautionary note

The field of systems biology still includes many challenges and holds much promise, especially for new cross-disciplinary scientists entering the field. Of course, although any number of educational options can place researchers in a potentially powerful and sought-after position in the field, students should beware of the universal curse of systems biology: as you quickly attain a breadth of knowledge in biology and mathematics, you risk losing, or fail to attain,

Table 3 Some emerging initiatives and educational programs in systems biology

German Systems Biology Research Program	http://www.systembiologie.de/
Manchester Interdisciplinary Biocenter, Manchester, UK	http://www.mib.umist.ac.uk/
University of Texas Southwestern, Dallas, TX, Program in Molecular, Computational & Systems Biology, Integrative Biology Graduate Program	http://www.utsouthwestern.edu/utsw/home/education/integrativebiology/

depth in either. Jack of all trades, or master of one? The choice is yours.

ACKNOWLEDGMENTS

The author would like to thank Eric Neumann, Adam Feist, Benno Schwikowski, Bernhard Palsson and Kristyn Ideker for valuable discussions and material for this commentary. Funding was generously provided by National Institutes of Health grant 1-P41RR018627-01.

- Ideker, T., Galitski, T. & Hood, L. A new approach to decoding life: Systems Biology. *Annu. Rev. Genomics Hum. Genet.* **2**, 343–372 (2001).
- Kitano, H. Computational systems biology. *Nature* **420**, 206–210 (2002).
- Bertalanffy, L.v. General systems theory: foundations, development, applications. (George Braziller, New York, 1969).
- Zhu, H. & Snyder, M. "Omic" approaches for unraveling signaling networks. *Curr. Opin. Cell Biol.* **14**, 173–179 (2002).
- Zerhouni, E. Medicine. The NIH Roadmap. *Science* **302**, 63–72 (2003).
- UK Biotechnology and Biological Sciences Research Council. *World Class Science. Strategic Plan 2003–2008*. (BBSRC, Swindon, UK, 2003).
- Ideker, T. & Lauffenburger, D.A. Building with a scaffold: emerging strategies for high- to low-level cellular modeling. *Trends in Biotechnol.* **21**, 255–262 (2003).
- Friedman, N. Inferring cellular networks using probabilistic graphical models. *Science* **303**, 799–805 (2004).
- Cohen, H. Systems biology: a pale beacon for biotech. *The Scientist* **17**, 52 (2003).