Cell

Leading Edge
Essay

Network News: Innovations in 21st Century Systems Biology

Adam P. Arkin[1,4,*] and David V. Schaffer[1,2,3,4]
[1]Department of Bioengineering
[2]Department of Chemical and Biomolecular Engineering
[3]The Helen Wills Neuroscience Institute
University of California, Berkeley, Berkeley, CA 94720, USA
[4]Physical Biosciences Division, E.O. Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
*Correspondence: aparkin@lbl.gov
DOI 10.1016/j.cell.2011.03.008

A decade ago, seminal perspectives and papers set a strong vision for the field of systems biology, and a number of these themes have flourished. Here, we describe key technologies and insights that have elucidated the evolution, architecture, and function of cellular networks, ultimately leading to the first predictive genome-scale regulatory and metabolic models of organisms. Can systems approaches bridge the gap between correlative analysis and mechanistic insights?

System biology aims to understand how individual elements of the cell interact to generate behaviors that allow survival in changeable environments and collective cellular organization into structured communities. Ultimately, these cellular networks assemble into larger population networks to form large-scale ecologies and thinking machines, such as humans. Given this central focus on codifying the organizational principles and algorithms of life, we argue that systems biology is not a newly emerging field, but rather a mature synthesis of thought about the implications of biological structure and its dynamic organization, ideas that have been brewing for more than a century.

To many scientists, the beginning of the last decade marked the definition and rise of the field of systems biology. However, systems biology's conceptual origins date back almost 100 years. In 1917, D'Arcy Thompson formalized the first link between development, evolution, and physics in his treatise *On Growth and Form*, when he observed that shapes and function of biological systems were fundamentally determined by physical requirements and mechanical laws. In 1939, Walter Canon, then chairman of the Department of Physiology at Harvard Medical School, coined the term "homeostasis" when he noted that organisms hold essential physiological variables at constant values despite a fluctuating environment (Canon, 1939). In 1943, the American mathematician Norbert Weiner, along with his coauthors, proposed that negative feedback loops would be central to maintaining this stability in biological systems (Rosenbleuth et al., 1943), thus linking concepts of control and optimality with biological dynamics. Ten years later, the British developmental biologist Conrad Waddington laid some of the modern foundation for systems biology when he presciently conceptualized networks of cellular components (i.e., genes, cells, and tissues) as evolutionarily dynamical systems expressible as solutions to a series of simultaneous differential equations. Over his long career, Waddington argued for a truly dynamic systems theory of cellular decision making driven by gene expression and epigenetics (Waddington, 1954, 1977). When Jacques Lucien Jacob and François Monod unveiled the molecular mechanisms of gene regulation in 1962, they noted, "it is obvious from the analysis of these [bacterial genetic regulatory] mechanisms that their known elements could be connected into a wide variety of 'circuits' endowed with any desired degree of stability" (Jacob and Monod, 1962).

During the ensuing decade, scientists across a wide array of disciplines started exploring the nonlinear dynamics in biochemical networks. Although experimental data to support their theoretical hypotheses were still largely missing, this period was quite productive, as numerous fundamental principles came to light. These included the possible mechanisms and advantages of different biochemical switches and oscillators with and without biochemical noise (Goodwin, 1963); new models of metabolic control and engineering (Heinrich and Rapoport, 1974; Kacser and Burns, 1973); the reverse engineering of cellular networks (Bekey and Beneken, 1978); and abstracted models of these networks to understand the evolution and optimization of specific network "designs" (Kauffman, 1969). Indeed, these latter principles of how networks can be structured to achieve particular functions have been used more recently to explicitly predict natural network behavior.

Thus, by the early 1970s, the concepts and components were all in place for what encompasses most of what we call "systems biology"—the integrated molecular analysis of cellular networks. However, one roadblock remained: experimental data to support the models and hypotheses. This is where the last two decades have revolutionized the field of cellular network inference and analysis.

Since the early 1990s, a vast array of technologies has dramatically improved the efficiency of manipulating cells genetically, the measurement of cellular components at high precision and completeness, and the dissemination of materials and information at unprecedented speeds (due to the other network revolution, which

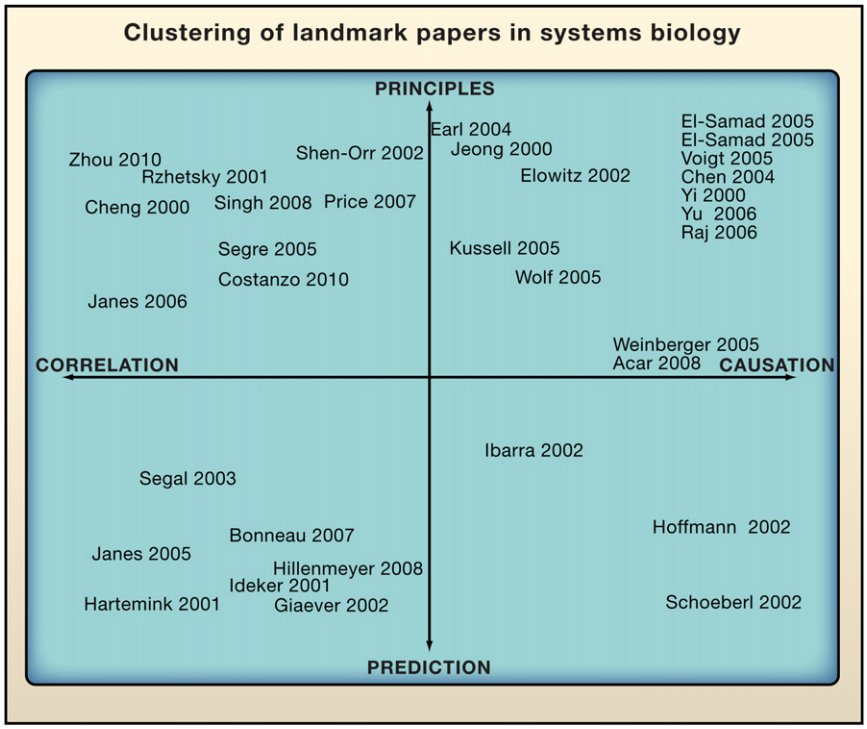

Figure 1. A Simplified Scheme for Organizing Results in the Field of Systems Biology
References are placed (subjectively) into this space according to whether their respective study focused more on mechanistic insight or on large-scale correlation analysis (the x axis) and whether the results were primarily principles about cellular networks or predictions of their behavior (the y axis). (Because of space constraints, only the last name of the first author is given).

has also left a conceptual mark on systems biology). Many of these biological technologies are scaling by a Moore's Law-type (Moore, 1965) dynamic in which every few years, the amount of DNA that can be sequenced or synthesized doubles in size for half the cost (as has the number of transistors on a microchip) (Carlson, 2003). Clearly, this ability to read and write genomic information has profoundly accelerated systems biology.

Principles versus Prediction and Correlation versus Causation

This brief historical perspective suggests that discoveries in systems biology may be organized within a conceptual space (Figure 1). The y axis distinguishes between two relatively distinct objectives: deducing principles of network organization necessary for behaviors versus reverse engineering networks to predict their behavior. Strikingly, with the advent of scaling biological data, two general approaches have evolved to meet these objectives. On one hand, correlative studies, which are usually on the genomic scale, infer relationships among genes and modules of function. These studies can also annotate genes and their products by a "guilt-by-association" approach in which detailed biochemical information available about one gene or system is transferred to others with correlated behaviors. This strategy contrasts with a "casual" approach in which direct interactions among molecules are tracked to glean mechanistic insights. Interestingly, as genetic and biochemical technologies climb the scaling curves, correlative and causal studies have become more intermingled. In other words, as it becomes possible to rapidly alter any gene (Paddison et al., 2004), modulate any gene's expression level, and perhaps even reorganize large regions of the genome (Gibson et al., 2010; Wang et al., 2009; Warner et al., 2010), mechanistic studies will become available at a genome scale.

Obviously, prediction is not truly antipodal to principles, nor is correlation distantly removed from causation; indeed, the quadrants are connected. However, when we asked a group of colleagues which systems biology papers over the last decade have been most important to the field, the resulting set of landmark studies naturally clustered into different regions of this systems biology "plane" (Figure 1).

Correlative Approaches

Genome-scale data have fundamentally changed the types of questions that we ask about cellular systems. We can now observe how genomes dynamically change expression in response to environmental conditions and then correlate these results to other phenotypes, such as growth, fate choices, and biosynthetic productivity. Such experiments have inspired several classes of analysis that can vastly improve the data-driven annotation of genomes, more strongly link genotype to phenotype through inferred networks of interaction, and predict behaviors of cellular systems (Figure 1, lower-left quadrant). They have also led to a wide array of conceptual interpretations about the organization and evolution of cellular networks into evolvable modules, the decomposition of these networks into recurrent regulatory "motifs" with useful dynamical function, and the robustness of these architectures to mutation (Figure 1, upper-left quadrant).

Correlative Approaches to Predicting Function

One type of analysis infers properties of biomolecules from correlated changes of genome-scale RNA, protein, DNA copy number, or metabolite abundance as it varies in time and across conditions. Most often, genes sharing common expression dynamics are inferred to share regulators and possibly functional roles, as least at some level (Brown and Botstein, 1999). The challenge in this area has been isolating the set of correlated genes from the background of measurement noise and from those genes with merely coincident coexpression. Although clustering techniques have been used for decades to derive relationships in complex correlative data sets such as those found in gene expression compendia, in 2000, Cheng and Church introduced an algorithm called "biclustering" that explicitly discovers "modules" from such data. This method identifies groups of genes, or "modules,"

with similar patterns of expression over a specific subset of conditions (Cheng and Church, 2000). Individual genes may belong to multiple modules, thereby allowing inference of their numerous functions and combinatorial regulation. This important work inspired an increasing number of algorithms concerned with identifying related sets of biomolecules from complex data and inferring their "modular" function. These algorithms thereby opened the door to discovering an apparent hierarchical modular architecture to cellular regulation, which complements the more informal "pathway" organization with which biologists were familiar. The modules of coherent function also greatly simplify construction and interpretation of predictive models, as they enabled prediction of how different modules, rather than the individual constituent genes, are dynamically deployed—a system formulation that has far fewer variables and thus requires far less data.

Gene expression can be an indirect measurement of a component's contribution to a particular cellular process, and thus, genetic perturbations and activity assays may be required. In seminal work, Giaever et al. (2002) constructed a bar-coded deletion library for the entire genome of *Saccharomyces cerevisiae*. This library enabled single-pot assays of the relative growth or fitness of each strain when exposed to a specific condition (Giaever et al., 2002). In a subsequent study, a growth phenotype for nearly every gene in yeast was identified using 1000 chemical perturbations (Hillenmeyer et al., 2008). These types of studies can rapidly dissect the cellular targets of drugs and even directly identify specific transporters involved. In addition, these studies have shown that genes displaying changes in expression under a given condition are not always the genes necessary for responding functionally to that condition (Giaever et al., 2002). Although the implications of this result are not fully understood, one obvious conclusion is that different types of experiments are required to deduce or even predict function of genes.

Correlative Prediction of Organization

Another type of analysis seeks to infer relationships among gene modules; in other words, the strategy used to infer function of a single gene is now extended to infer the underlying biochemical network (Arkin et al., 1997). In 2001, Ideker et al. combined genetic, macromolecular interactions and expression data (both protein and gene) to infer how the galactose utilization network in yeast is regulated (Ideker et al., 2001). They then used the resulting "influence network" to predict how the system responds to genetic perturbations. Some of these predictions were validated by experiments, yet others were proven incorrect, suggesting that properties of this well-characterized regulatory network still await discovery.

Variants of this approach that applied additional, more sophisticated algorithms from multivariate statistics and machine learning quickly began to have a strong impact on the field. In particular, Hartemink et al. (2001) offered perhaps the first Bayesian approach for rating different network structural hypotheses (i.e., different patterns of molecular interaction) against data. Using a collection of 52 conditions, they demonstrated that it was possible to infer the regulatory interactions in the galactose pathway (Hartemink et al., 2001). Two years later, Segal et al. (2003) increased the power of these algorithms to infer the sets of genes (i.e., modules) regulated by particular transcription factors under specific conditions. This algorithm also correctly predicted new regulatory roles for less-characterized proteins (Segal et al., 2003). In particular, the model predicted that one putative transcription factor (Ypl230w) and two signaling molecules (Kin82 and Ppt1) were important for cellular response to three different conditions: heat shock, hypo-osmotic shift, and entry into stationary phase, respectively. Disrupting the genes elicited no expression phenotype in rich, unstressed conditions but strong changes in expression relative to wild-type in the condition predicted to be relevant for a given gene.

Applying a different statistical approach called "Partial Least Squares Regression," Janes and colleagues undertook herculean efforts to measure and correlate mammalian cell survival, apoptosis, intracellular protein phosphorylation states, and kinase activities (thereby generating a data set with 7980 intracellular measurements) in response to combinations of extracellular growth factor and cytokine inputs (Janes et al., 2005, 2006). The resulting model successfully predicted the level of apoptosis as a function of cytokine inputs and led to the new mechanistic insight that cascades of autocrine signaling were involved in mediating downstream cell responses to the extracellular cues.

Shortly thereafter, in another landmark paper, Bonneau et al. (2007) demonstrated how the output of a new gene expression biclustering algorithm provided input to a clever regression algorithm that deciphers the transcriptional regulatory network of an Archaea (*Halobacterium salinarum* NRC-1) and predicts expression responses to > 100 conditions (Bonneau et al., 2007). Recently, such correlative systems analyses are scaling up to link biomolecular networks to ecological networks. These pioneering studies are uncovering new scales of biological organization that should lead to entirely new principles of ecosystem function (Zhou et al., 2010).

Nevertheless, it is not yet clear how to optimally design perturbation repertoires to achieve maximum accuracy in annotating gene function and regulation and in predictive model inference with minimal expense. Also, it has yet to be proven that the models obtained in these types of studies are sufficiently accurate or inexpensive to have an impact in a medical or industrial setting. Nonetheless, the ability to collect such compendia of data, even from diverse types of experiments, is rapidly becoming a feasible task for even a single laboratory to accomplish. We predict that the increased accessibility to these large-scale data sets will enable the detailed characterization of organisms after their genomes are sequenced and may, ultimately, change what it means to "complete" the genome of an organism.

Uncovering Principles of Network Organization

The fact that clear functional modules of gene expression can be inferred from correlative data sets implies the existence of underlying organizational principles for these networks. Similar hierarchies of modules have been found in large-scale protein interaction data and metabolic networks. Certain "scale-free" topologies of molecular interaction networks have received considerable attention in biology

and other fields. Such topologies, which seem to arise often in both natural and human designed systems, are characterized by a pattern of interconnectedness among the nodes (e.g., proteins) in which the number of interactions per node follows a power law. Influential papers have suggested that these topologies lead to robustness to perturbation (Jeong et al., 2000) and in the case of proteins, naturally arise due the evolutionary process of duplication and divergence (Rzhetsky and Gomez, 2001). Likewise, in developmental biology, it has been argued for decades that for integrated cellular processes to evolve, they must be dissociable into hierarchical, modular units that can adapt their behavior with little interference from other such units. Thus, interaction and expression modules may allow rapid, effective rewiring and tuning of internal dynamics (Price et al., 2007; Singh et al., 2008), such that this ability to evolve may even be a selectable trait (Earl and Deem, 2004). However, caution must be taken in assigning evolutionary meaning to apparent modularity (Lynch, 2007).

On slightly smaller size scales, certain topological motifs—that is, stereotypical small networks of regulatory interactions and chemical reactions—may have important control functions for cellular networks (Rao and Arkin, 2001). The availability of large-scale data has, in the last decade, enabled the discovery that certain motifs appear more than expected by random chance (Shen-Orr et al., 2002), including feed-forward and feedback loops (for more on feed-forward loops, see Review by Yosef and Regev on page 886 of this issue). These motifs have potential functional importance, such as noise rejection, and appear physiologically robust but also evolutionarily flexible with tunable function (Voigt et al., 2005). Milo et al. (2002) hypothesized that these motifs might form a sort of basis set of dynamic functions from which complex optimized networks could be assembled in numerous contexts within and outside of biology (Milo et al., 2002).

A beautiful theoretical paper by Segrè et al. (2005) determined another organizational principle of cellular networks. They not only showed that functional modules could be inferred from growth phenotypes of double knockout mutants, but also that the epistatic interactions between pairs of genes in these modules always fell into one of two classes of interactions: buffering, in which epitasis diminishes the individual phenotypic effects of the two mutations, or aggravating, in which the deleterious, individual effects of two mutations are worsened by their combination (Segrè et al., 2005). Modules were thus ''monochromatic'' and never contained mixed type genes, a principle that was recently verified experimentally (Costanzo et al., 2010).

These architectural principles uncovered from large sets of correlative data are evocative and well supported, but the challenge remains to find incontrovertible evidence for evolutionary selection of these architectures and to fully characterize their functional consequences.

Mechanistic Approaches to Study Causal Relationships

Although large-scale genomic data sets lend themselves to statistical analysis of correlation, causal analysis necessitates more detailed biochemical data on the networks' effectors, such as proteins, second messengers, and metabolites. Unfortunately, the experimental analyses of these components have not enjoyed the same growth in scale as those of nucleic acids. That is, whereas volumes of data on one-dimensional genomes are readily available, causal analysis also requires multidimensional data on biomolecules' interactions, reactions and their rates, localization, and transport. Mass spectrometry, imaging, genetic sensors, chemical probes, and other technologies are increasingly providing such data, but not yet at the same magnitude as genomic information. As a result, causal analyses of cellular networks initially focused on elucidating functional principles but are becoming increasingly empowered with data to enable prediction.

Uncovering Principles of Function

Large-scale models of biological networks face the challenges that molecular mechanisms are often complex and nonlinear (e.g., cooperative protein interactions and epigenetic regulation) and many of their inherent parameters are unknown (e.g., affinities and rate constants). However, in some model systems, the biochemistry is sufficiently well characterized to enable the construction of elegant, large-scale models.

As a prime example, Tyson and colleagues (Chen et al., 2004) modeled the cell-cycle control system of *Saccharomyces cerevisiae* using a set of 35 ordinary differentiation equations (ODE) representing molecular mechanisms and mass action (Chen et al., 2004) (for more on modeling the cell cycle, see Primer by Ferrell et al. on page 874 of this issue). The goal of the model was not to account for the full complexity of the system but instead to provide a reasonable approximation of network behavior and to uncover dynamical principles of the architecture. Indeed, their model succeeded in accounting for a majority of mutant phenotypes simulated.

Using a similar framework, El-Samad et al. (2005a, 2005b) modeled the heat shock response in *Escherichia coli.* Despite the simplicity of the response—deploying chaperones to keep proteins folded at higher temperature—this model uncovered complexity in the modular control structure of the system. It also demonstrated how the many feedback loops in this system confer the ability to respond quickly and robustly while also trying to minimize the energetic cost of heat shock protein expression (El-Samad et al., 2005a, 2005b). In another important study, Yi et al. used dynamical systems control theory to analyze bacterial chemotaxis (Yi et al., 2000), another system with well-characterized biochemistry. Building on the principle that negative feedback is often central to biological stability (Rosenbleuth et al., 1943), the study found that integral feedback control underlies the robustness of network adaptation to significant perturbations in both the amounts and kinetic parameters of its component proteins. Interestingly, control engineers ''reinvented'' this strategy and proved that it is required, in certain conditions, to build robustness into electrical circuits and other systems.

Deterministic representations of networks are compromised when their constituents are present at low concentrations or undergo slow reactions. Moreover, early studies suggested that noise can significantly influence network function (Arkin et al., 1998). Elowitz et al. (2002) explored the principle that fluctuations in the quantities and reaction rates of gene expression machinery can cause noise in gene expression at both a global

level in a cell (extrinsic), as well as for an individual gene (intrinsic) (Elowitz et al., 2002). Indeed, subsequent single-molecule imaging studies directly confirmed that both translation (Yu et al., 2006) and transcription (Raj et al., 2006) can underlie such noisy protein expression.

The principle that noise is inherent in biological networks raised the question of whether its effects on biological fitness are neutral, positive, or negative. Although the value of noise depends on the system, in certain cases, noise appears to make positive contributions to fitness. Organisms have a need to adapt to changing environments, and two adaptation strategies are sensing and responding to change or stochastically switching phenotype.

Two theoretical studies arrived at the principle that, under some conditions, such as when transitions in selective environments are slow or cannot be sensed, stochastic fluctuations in an organism's phenotype can increase its fitness (Kussell and Leibler, 2005; Wolf et al., 2005). In a study that combined experimental approaches with simulations, Weinberger et al. (2005) investigated this principle by analyzing stochastic effects in HIV infection (Weinberger et al., 2005). Low initial numbers of viral molecules, slow gene expression, and amplification by a positive feedback loop lead to very noisy gene expression, which for some infections yielded long delays in gene expression. This delayed expression contributed to the formation of latent HIV, which is clinically recognized as the most formidable barrier to the elimination of virus from a patient.

In an elegant study, Acar et al. (2008) engineered *Saccharomyces cerevisiae* strains that stochastically switched phenotypes at different rates. Interestingly, they found that the fast-switching strain outgrew the slow-switching strain in environments undergoing rapid fluctuations, whereas the slow-switching strains were more fit in environments that fluctuated slowly (Acar et al., 2008).

Predictive Analysis of Network and Cell Function

The complexity of molecular mechanisms and scarcity of biochemical parameters often makes the development of predictive models challenging. Ibarra et al. (2002) created a constraints-based whole-cell metabolic model for *E. coli*, in which stoichiometric, thermodynamic, and other constraints mathematically yielded a solution space of allowed metabolic network states (Ibarra et al., 2002). This model, which requires fewer parameters than full dynamical models, can make predictions of network function that optimize growth under different environmental conditions. Indeed, when Ibarra et al. grew *E. coli* on a new carbon substrate, the cells evolved to the metabolic state predicted by the model.

In some systems, substantive comparison to data can yield deterministic models increasingly capable of prediction. Hoffmann and colleagues (2002) analyzed the mammalian NF-κB system (Hoffmann et al., 2002), in which activation of this transcription factor upregulates expression of IκBα, a negative regulator of NF-κB. Integrating experimental data with a deterministic model enabled prediction of the oscillatory behavior of this module upon stimulation and perturbation. Finally, Schoeberl et al. (2002) developed a model with 94 ODEs to simulate epidermal growth factor signaling through MAP kinase, including receptor trafficking dynamics and intracellular phosphorylation cascades (Schoeberl et al., 2002). This is the first dynamic model of a large cellular signaling network that was carefully parameterized by prior experimental measurements and that yielded prediction on signal transduction dynamics, which were subsequently validated experimentally.

The Next Decade

As systems biology matures, the number of studies linking correlation with causation and principles with prediction continues to grow (Figure 1). Advances in measurement technologies that enable large-scale experiments across an array of parameters and conditions will increasingly meld these correlative and causal approaches, including correlative analyses leading to mechanistic hypothesis testing as well as causal models empowered with sufficient data to make predictions. In addition, the increasing number of organisms sequenced and the increasing ease of measurement and genetic manipulation will enable deep comparison of systems across phylogenetic trees, thereby enhancing our understanding of mechanistic features that are necessary for function and evolution. The increasing integration of experimental and computational technologies will thus corroborate, deepen, and diversify the theories that the earliest systems biologists used logic to infer, thereby inching us ever closer to that central question: "What is Life"?

ACKNOWLEDGMENTS

We would like to thank our colleagues for suggesting a number of the papers that we reference in this work, and we apologize for not being able to include all of them. The authors would like to acknowledge the National Institutes of Health (R01 GM073010-01), and work conducted by ENIGMA was supported by the Office of Science, Office of Biological and Environmental Research of the US Department of Energy under contract number DE-AC02-05CH11231.

REFERENCES

Acar, M., Mettetal, J.T., and van Oudenaarden, A. (2008). Nat. Genet. *40*, 471–475.

Arkin, A., Ross, J., and McAdams, H.H. (1998). Genetics *149*, 1633–1648.

Arkin, A.P., Shen, P.-D., and Ross, J. (1997). Science *277*, 1275.

Bekey, G.A., and Beneken, J.E.W. (1978). Automatica *14*, 41–47.

Bonneau, R., Facciotti, M.T., Reiss, D.J., Schmid, A.K., Pan, M., Kaur, A., Thorsson, V., Shannon, P., Johnson, M.H., Bare, J.C., et al. (2007). Cell *131*, 1354–1365.

Brown, P.O., and Botstein, D. (1999). Nat. Genet. *21*(1, Suppl), 33–37.

Canon, W. (1939). The Wisdom of the Body (London: Norton).

Carlson, R. (2003). Biosecurity and Bioterrorism: Biodefense Strategy, Practice, and Science *1*, 203–214.

Chen, K.C., Calzone, L., Csikasz-Nagy, A., Cross, F.R., Novak, B., and Tyson, J.J. (2004). Mol. Biol. Cell *15*, 3841–3862.

Cheng, Y., and Church, G.M. (2000). Proc. Int. Conf. Intell. Syst. Mol. Biol. *8*, 93–103.

Costanzo, M., Baryshnikova, A., Bellay, J., Kim, Y., Spear, E.D., Sevier, C.S., Ding, H., Koh, J.L., Toufighi, K., Mostafavi, S., et al. (2010). Science *327*, 425–431.

Earl, D.J., and Deem, M.W. (2004). Proc. Natl. Acad. Sci. USA *101*, 11531–11536.

El-Samad, H., Khammash, M., Homescu, C., and Petzold, L. (2005a). Proceedings 16th IFAC World Congress. http://engineering.ucsb.edu/~cse/Files/IFACC_HS_OPT04.pdf.

El-Samad, H., Kurata, H., Doyle, J.C., Gross, C.A., and Khammash, M. (2005b). Proc. Natl. Acad. Sci. USA *102*, 2736–2741.

Elowitz, M.B., Levine, A.J., Siggia, E.D., and Swain, P.S. (2002). Science *297*, 1183–1186.

Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Véronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., André, B., et al. (2002). Nature *418*, 387–391.

Gibson, D.G., Glass, J.I., Lartigue, C., Noskov, V.N., Chuang, R.-Y., Algire, M.A., Benders, G.A., Montague, M.G., Ma, L., Moodie, M.M., et al. (2010). Science *329*, 52–56.

Goodwin, B.C. (1963). (London: Academic Press).

Hartemink, A.J., Gifford, D.K., Jaakkola, T.S., and Young, R.A. (2001). In Pacific Symposium on Biocomputing 2001 (PSB01), R. Altman, A.K. Dunker, L. Hunter, K. Lauderdale, and T. Klein, eds. (New Jersey:: World Scientific), pp. 422–433.

Heinrich, R., and Rapoport, T.A. (1974). Eur. J. Biochem. *42*, 89–95.

Hillenmeyer, M.E., Fung, E., Wildenhain, J., Pierce, S.E., Hoon, S., Lee, W., Proctor, M., St Onge, R.P., Tyers, M., Koller, D., et al. (2008). Science *320*, 362–365.

Hoffmann, A., Levchenko, A., Scott, M.L., and Baltimore, D. (2002). Science *298*, 1241–1245.

Ibarra, R.U., Edwards, J.S., and Palsson, B.O. (2002). Nature *420*, 186–189.

Ideker, T., Thorsson, V., Ranish, J.A., Christmas, R., Buhler, J., Eng, J.K., Bumgarner, R., Goodlett, D.R., Aebersold, R., and Hood, L. (2001). Science *292*, 929–934.

Jacob, F., and Monod, J. (1962). Cold Spring Harb. Symp. Quant. Biol. *26*, 193–211.

Janes, K.A., Albeck, J.G., Gaudet, S., Sorger, P.K., Lauffenburger, D.A., and Yaffe, M.B. (2005). Science *310*, 1646–1653.

Janes, K.A., Gaudet, S., Albeck, J.G., Nielsen, U.B., Lauffenburger, D.A., and Sorger, P.K. (2006). Cell *124*, 1225–1239.

Jeong, H., Tombor, B., Albert, R., Oltvai, Z.N., and Barabási, A.L. (2000). Nature *407*, 651–654.

Kacser, H., and Burns, J.A. (1973). Symp. Soc. Exp. Biol. *27*, 65–104.

Kauffman, S.A. (1969). J. Theor. Biol. *22*, 437–467.

Kussell, E., and Leibler, S. (2005). Science *309*, 2075–2078.

Lynch, M. (2007). Proc. Natl. Acad. Sci. USA *104* (*Suppl 1*), 8597–8604.

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., and Alon, U. (2002). Science *298*, 824–827.

Moore, G.E. (1965). Electronics *38*, 114–117.

Paddison, P.J., Silva, J.M., Conklin, D.S., Schlabach, M., Li, M., Aruleba, S., Balija, V., O'Shaughnessy, A., Gnoj, L., Scobie, K., et al. (2004). Nature *428*, 427–431.

Price, M.N., Dehal, P.S., and Arkin, A.P. (2007). PLoS Comput. Biol. *3*, 1739–1750.

Raj, A., Peskin, C.S., Tranchina, D., Vargas, D.Y., and Tyagi, S. (2006). PLoS Biol. *4*, e309.

Rao, C.V., and Arkin, A.P. (2001). Annu. Rev. Biomed. Eng. *3*, 391–419.

Rosenbleuth, A., Wiener, N., and Bigelow, J. (1943). Philos. Sci. *10*, 18–43.

Rzhetsky, A., and Gomez, S.M. (2001). Bioinformatics *17*, 988–996.

Schoeberl, B., Eichler-Jonsson, C., Gilles, E.D., and Müller, G. (2002). Nat. Biotechnol. *20*, 370–375.

Segal, E., Shapira, M., Regev, A., Pe'er, D., Botstein, D., Koller, D., and Friedman, N. (2003). Nat. Genet. *34*, 166–176.

Segrè, D., Deluna, A., Church, G.M., and Kishony, R. (2005). Nat. Genet. *37*, 77–83.

Shen-Orr, S.S., Milo, R., Mangan, S., and Alon, U. (2002). Nat. Genet. *31*, 64–68.

Singh, A.H., Wolf, D.M., Wang, P., and Arkin, A.P. (2008). Proc. Natl. Acad. Sci. USA *105*, 7500–7505.

Voigt, C.A., Wolf, D.M., and Arkin, A.P. (2005). Genetics *169*, 1187–1202.

Waddington, C.H. (1954). Proceedings of the 9th International Congress of Genetics *9*, 232–245.

Waddington, C.H. (1977). Tools for thought (New York: Basic Books).

Wang, H.H., Isaacs, F.J., Carr, P.A., Sun, Z.Z., Xu, G., Forest, C.R., and Church, G.M. (2009). Nature *460*, 894–898.

Warner, J.R., Reeder, P.J., Karimpour-Fard, A., Woodruff, L.B., and Gill, R.T. (2010). Nat. Biotechnol. *28*, 856–862.

Weinberger, L.S., Burnett, J.C., Toettcher, J.E., Arkin, A.P., and Schaffer, D.V. (2005). Cell *122*, 169–182.

Wolf, D.M., Vazirani, V.V., and Arkin, A.P. (2005). J. Theor. Biol. *234*, 227–253.

Yi, T.-M., Huang, Y., Simon, M.I., and Doyle, J. (2000). Proc. Natl. Acad. Sci. USA *97*, 4649–4653.

Yu, J., Xiao, J., Ren, X., Lao, K., and Xie, X.S. (2006). Science *311*, 1600–1603.

Zhou, J., Deng, Y., Luo, F., He, Z., Tu, Q., and Zhi, X. (2010). MBio. *1*, e00169-10.