I

OVERVIEW

1

ADVANCES IN COMPUTATIONAL SYSTEMS BIOLOGY

Huma M. Lodhi

Department of Computing, Imperial College London, London, SW7 2AZ, UK

1.1 INTRODUCTION

Computational systems biology, a rapidly evolving field, is at the interface of computer science, mathematics, physics, and biology. It endeavors to study, analyze, and understand complex biological systems by taking a coordinated integrated systems view using computational methodologies. From the middle of the twentieth century till present, we have been witnessing breakthrough discoveries in biology that range from molecular structure of deoxyribonucleic acid (DNA) to the generation of the sequence of the euchromatic portion of the human genome. There have also been recent advances in sophisticated computational methodologies, high-throughput biotechnologies, and computational power. The stunning developments in diverse disciplines such as biology and computer science are playing a key role in the fast progression of the emerging field. Computational systems biology provides a point of convergence for genomics, proteomics, metabolomics, and computational modeling. It is characterized by its focus on experimental data, computational techniques, and hypotheses testing [1–3].

Open and unsolved problems in biology range from understanding structure and dynamics of biological systems to prediction and inference in the complex systems.

Elements of Computational Systems Biology Edited by Huma M. Lodhi and Stephen H. Muggleton

In the postgenomic era, systems-based approaches may provide a solution to such unsolved problems. It is believed that some answer to the question "what is life" may be obtained by taking a broader, integrated view of biology [4]. However, applications of systems-based techniques to biology are not new. Such methods and frameworks have been applied to analyze biological processes since early twentieth century [5, 6]. Norbert Wiener's groundbreaking work [7] is a well-known example of these applications.

The purpose and objective of this chapter is to review cutting-edge and long-ranging research in the field of computational systems biology in the recent years. However, the review is not meant to be exhaustive. We briefly describe novel methodologies to build multiscale biological models in Section 1.2. In Section 1.3, we present an overview of the applications of proteomics techniques to study biological processes. We then summarize computational systems biology methods to examine and understand aging in Section 1.4. Section 1.5 describes systems-based techniques for drug design, where such methods are revolutionizing the process of drug discovery. Efficient software tools and infrastructure are crucial to solving complex biological problems. In Section 1.6, we review tools for systems biology.

1.2 MULTISCALE COMPUTATIONAL MODELING

In the postgenomic era, researchers seek to focus their attention to studying and analyzing biological networks and pathways by the use of multiscale computational modeling techniques. A model can be viewed as a representation of a biological system, where the representation can comprise a set of differential equations [8], a set of first-order logic clauses [9], and so on. Biological models that incorporate multiple scales such as time and space or multiple timescales may be viewed as multiscale models [10]. Chapter 2 gives an in-depth account of mathematical and computational models in systems biology.

Development of efficient and effective computational methodologies to perform modeling, simulation, and analysis of complex biological processes is a challenging task. Traditionally, mathematical and computational models have been developed by considering a single scale. However, it is now feasible to incorporate multiple scales in the process of model building due to recent advances in computational power and technology. Generally, multiscale models are constructed by using sophisticated techniques including numerical methods and integration approaches. Multiscale model of the heart [11, 12] is a well-known example of an application of these modeling techniques.

Multiscale computational modeling and simulation methods are showing promising results in the field of oncology. The development of three-dimensional multiscale brain tumor model by Zhang et al. [13] is an attempt in this direction. The dynamics of tumor growth were simulated by using an agent-based multiscale model where microscopic scale, macroscopic scale, and molecular scale were incorporated in the *in silico* model. In micro-macroscopic environment, a virtual brain tissue block was represented by points in three-dimensional lattice. The lattice was

divided into four cubes that illustrated the behavior of chemotactically acting tumor cells. The chemotaxis distribution of transforming growth factor alpha (TGFα), glucose, and oxygen tension were illustrated in a set of mathematical equations. It was observed that the amount of TGFα and glucose was chemoattractant, and diffusion of glucose occurred at a constant rate. In order to incorporate molecular scale, epidermal growth factor receptor (EGFR) gene–protein interaction network model [14] was used in conjunction with cell cycle module. The authors used a simplified EGFR network that comprised of EGFR and TGFα genes. The mathematical model of EGFR gene–protein network was represented as a set of differential equations. The authors utilized the cell cycle model presented in Tyson and Novak [15] and Alacron et al. [16]. The implementation of the software systems was carried out by combining in-house code with an agent-based software tool, namely, MASON (`http://cs.gmu.edu/ eclab/projects/mason/`). In order to study and analyze tumor growth and spread, 10 simulations were performed. The results demonstrated an increase in tumor volume with respect to time, where the relationship between tumor volume and time was not linear. There was a sharp increase in volume growth at later time intervals. The study found that migrating and proliferating cells exhibited a dynamic behavior with respect to time. Furthermore, the cells caused spatiotemporal tumor growth. The results showed that the number of migrating cells was greater than the number of proliferating cells over time, where the high concentration of phospholipase C gamma (PLCγ) might be the key factor behind the phenomenon. In summary, the study demonstrated a successful construction of multiscale computational model of the complex multifaceted biological process. However, the approach is not free from shortcomings as described below:

- A simple EGFR network was used.
- Clonal heterogeneity within tumor was not examined.

It has been found that the distribution of tumor cells is not homogeneous, and the cells exhibit heterogeneous patterns. Techniques that account for clonal heterogeneity of tumor cell populations can be vital to analyze and study the development of cancerous diseases. Furthermore, clonal heterogeneity can strongly impact the design of effective therapeutic strategies. Therefore, many studies examined heterogeneity in tumors [17, 18]. Zhang et al. [19] extended their multiscale computational modeling technique [13] to investigate the clonal heterogeneity by incorporating genetic instability. The extended model included doubling time of cell and cell cycle. Other parameters such as cell–cell adhesion were also considered so that the strength of the chemoattractants' (TGFα, oxygen tension, and glucose) impact on cancer cells adhesion and rate of cell migration could be investigated. The authors used Shannon's entropy for the quantification of tumor heterogeneity. Shannon entropy in this context can be calculated as follows: Let c_i denote the occurrence of clone i in the tumor, the entropy is given by $\sum_i c_i \ln(c_i)$, where the higher values of Shannon's entropy represent more clonal heterogeneity.

The results of the study showed an increase in tumor total volume over time, where the tumor was categorized into three regions on the basis of the distance between it

and the nutrient source. It was observed that there was a general increase in the values of Shannon's entropy for all the three regions. However, there was highest clonal heterogeneity in the region closest to the nutrient source at early time stages where the region exhibited a homogeneous pattern at later stages. The study inferred that cancer could spread faster due to clonal heterogeneity as compared to homogeneous cell populations in tumor.

The complexity of the mechanisms of development and morphogenesis establishes a need to design effective and efficient computational techniques to investigate and analyze the biological process. In a recent study, Robertson et al. [20] presented a multiscale computational framework to investigate morphogenesis mechanisms in *Xenopus laevis*. Mammalian cells share similarities with *X. laevis* in terms of signaling network and cell behavior. A multiscale model was constructed by integrating an intercellular signaling pathway model with the multicellular model of mesendoderm migration. The authors implemented Wnt/β-catenin signaling pathway model that was presented by Lee et al. [21], whereas an agent-based approach was applied to build mesendoderm migration model. In order to simulate mesendoderm cells' migration, it was viewed that each cell comprised of nine sections, where each section was modeled as an agent. Mesendoderm migration was facilitated by the use of fibronectin extracellular matrix substrate. The study found that fibronectin gradient was a key factor behind the cellular movement. It was also observed that polarity signals [22] might be important for mesendoderm migration and morphogenesis. The simulations also demonstrated the importance to keep the cadherin binding strength in balance with the integrin binding strength. Although the study establishes the efficacy of multiscale computational methodologies to studying morphogenesis, the proposed approach may not be computationally attractive for large-scale simulations.

Physiome project [12] is well known for the development of multiscale modeling infrastructures. Given that standard modeling languages are useful for sharing biological data and models, three markup languages, namely, CellML (`http://www.cellml.org/`), FieldML, and ModelML, have been developed in the project. CellML [23] is characterized by its ability to capture three-dimensional information regarding cellular structures. It can also incorporate mathematical knowledge and metadata. FieldML, a related language, is known for its incorporation of spatial information. The third systems biology modeling language, namely, ModelML, is characterized by its ability to encode physical equations that illustrate complex biological processes. The efficacy of the languages was established by building multiscale heart models [12].

It has been found that same input, to constituent parts of a system, can produce different outputs. Such variations may be produced by factors including alterations in the concentration of system's components. It is desirable to design techniques and methods that can provide robustness to variations. Shinar et al. [24] presented a robust method by exploiting molecular details. The authors coined the term "input–output relation" for the association between input signal strength and output. The study investigated the input–output relation in bacterial signaling systems.

1.3 PROTEOMICS

Proteomics, the study of proteins, is viewed crucial to analyze and understand biological systems, as protein is the building block of life. Mass spectrometry (for details see Chapter 17) is a well-known proteomics technology that is showing a huge impact on the development of the field of computational systems biology. Several recent studies have identified the significant role of proteomics techniques in solving complex biological problems [25–27].

Proteomics methods and data can be useful for the reconstruction of biological networks. Recently, Rho et al. [28] presented a computational framework to reconstruct biological networks. The framework is based on the use of proteomics data and technologies to build and analyze computational models of biological networks. It is termed as integrative proteomic data analysis pipeline (IPDAP). IPDAP incorporates a number of network modeling and analysis tools. The component tools of IPDAP can be applied to reconstruct biological networks by fusing different types of proteomics data. The successful application of IPDAP to different cellular and tissue systems demonstrated the efficacy and functionality of the framework.

In another study, Zhao et al. [29] investigated signal transduction by applying techniques from optimization theory and exploiting proteomics and genomics data. They formulated the network identification problem as an integer linear programming problem. The proteomics (protein–protein interaction) data were represented as weighted undirected graph, where the nodes and the edges represented proteins and interaction between pair of proteins, respectively. The results of the study confirmed the efficacy of the approach in searching optimal signal transduction networks from the data.

Cell cycle comprises a series of ordered events by which cell replication and division take place. Studying cell cycle regulation provides useful insights in cancer growth and spread. The relationship between cell cycle and cancer has been a focus of many studies [30, 31]. In Sigal et al. [32], a proteomics approach was applied to investigate cell cycle mechanisms. The approach is based on the use of time-lapse microscopy to study protein dynamics. The study identified cell cycle-dependent changes in protein localization, where 40 percent of the investigated nuclear proteins demonstrated cell cycle dependence. Another challenging problem is to find patterns of polarized growth in cells where such growth is viewed as an important process in organisms. In order to investigate the biological problem, Narayanaswamy et al. [33] conducted a study by using budding yeast as the model system. The proposed computational method is based on the use of microarray image analysis and a machine learning technique, namely, naive Bayes algorithm. The study found 74 localized proteins including previously uncharacterized proteins and observed novel patterns of cell polarization in budding yeast.

In a recent study [34], a computational technique is presented for predicting peptide retention times. The method is at the intersection of two machine learning approaches, namely, neural networks and genetic algorithms. In order to predict the retention times, an artificial neural network is trained and the predicted values are further optimized

by using a genetic algorithm. The method was successfully applied to *Arabidopsis* proteomics data.

1.4 COMPUTATIONAL SYSTEMS BIOLOGY AND AGING

Aging is a complex phenomenon that has not been well understood. In aging, we witness gradual diminishing/decreasing functions at different levels, including organs and tissues. Cell division has been viewed as a key process in aging since long [35, 36]. Recently, de Magalhaes and Faragher [37] have elucidated that aging might be affected by variations in cell division. Hazard rates and nutrition may be the key factors that influence the longevity of cellular organisms [38]. There are a number of theories that describe how aging occurs. Kirkwood [38] listed five different theories that are as follows:

- Somatic mutation theory
- Telomere loss theory
- Mitochondrial theory
- Altered proteins and waste accumulation theory
- Network theory

Aging has been extensively studied in *Caenorhabditis elegans* (nematode), mice, humans, and fruit flies. A number of genes that extend organisms' life span have been discovered. Several studies on aging found that genetic mutations could increase longevity [39–41]. Furthermore, aging genes with their associated pathways may influence the variations in aging between different species but may not have any affect on the differences in aging within a particular specie [42]. Gene expression and pathway analysis can provide useful means to identify aging-related similarities and differences between various species [43], where the efficacy of DNA microarray technology, in studying aging, is significant [44]. In a recent study on aging, DNA microarray experiments were utilized to show that aging in *C. elegans* is influenced by GATA transcriptional circuit [45].

Advances in computational systems biology have led to the development of tools and methods for solving highly complex problem of aging. For example, Xue et al. [46] addressed the key issue regarding aging by applying an analytic method to human/fruit fly protein–protein interaction network, namely, NP analysis [47]. The method is based on the identification of active modules in network, where the chosen module comprised of protein–protein interaction subnetwork between genes that show (positive or negative) correlation during aging. The application of the method to human brain aging identified four modules. Among these modules, the two showed transcriptionally anticorrelation with each other. The other two modules comprised of immunity genes and translational genes, respectively. In order to study correlation between genes in other species during aging, the method was applied to fruit fly interactome. The results of the study showed that in addition to two transcriptionally anticorrelated

genes modules, there were two other modules that demonstrated such anticorrelation. On the basis of these findings, the authors suggest that only a few modules are associated with aging. The other key result of the study is the identification of the influence of module connecting genes on aging.

In another study, Garan et al. [48] presented a computational systems biology framework for studying neuroendocrine aging. The framework allows fusion of heterogeneous data from different disciplines such as endocrinology, cell biology, genetics, and so on. The method can be effective in identifying underlying relationship between the components that define aging.

Machine learning provides useful approaches and techniques to conduct studies on aging. In Swindell et al. [49], a number of machine learning methods were used to predict mouse life span. Twenty-two learning algorithms were applied to the problem, where the results demonstrated usefulness of support vector machines (SVMs), stabilized linear discriminant analysis, and nearest shrunken centroid in solving the problem, hence establishing the efficacy of machine learning technique for aging research. Agent-based modeling techniques have also been used to understand the biological processes of aging. The study published by Krivenko and Burtsev [50] is indicative of the success of such approaches for aging related studies. The authors applied their technique to simulate evolution and studied important factors including kin recognition and aggression.

Analysis of pathways for aging can also facilitate the understanding of complex diseases such as cancer. The probability of the occurrence of a cancer can be substantially lowered by downregulating the aging pathways [39]. Recently, Bergman et al. [51] investigated longevity genes. They conducted an extensive study by using more than 1200 subjects. On the basis of system-based analysis, the authors recommend that the investigation of genetic pathways can lead to the development of strategies that may regulate age-related diseases and disorders.

1.5 COMPUTATIONAL SYSTEMS BIOLOGY IN DRUG DESIGN

Millions of people are suffering from fatal diseases such as cancer, AIDS, and many other bacterial and viral illnesses. Computational systems biology approaches can provide a solution to the key issue that is how to design lifesaving and cost-effective drugs so that the diseases can be cured and prevented. Pharmaceutical companies view that systems-based computational techniques will be highly useful in designing effective therapeutic drugs [52–54]. Furthermore, advanced and sophisticated methods will accelerate drug discovery and development. In 2007, FDA approved only 17 new drugs [55] and approximately 50 drugs in 2008 (`http://www.fda.gov/`).

It is believed that the association between systems-based biological methods and drug design is age-old. Herbal drugs were developed by observing the diseases; hence, today's drug design has been (directly/indirectly) influenced by such early attempts [56]. Computational systems biology approaches may revolutionize therapeutic intervention in clinical medicine [2]. Effective systems-based drug design techniques can be developed by exploiting the knowledge of the robustness of biological systems [57].

An overview of a number of computational methods' (Petri nets, cellular automata techniques, hybrid methods, pi calculus, agent systems, and differential equations-based methods) application to the task of drug design can be found in Materi and Wishart [52].

Identification of novel drug targets in diseases is a key problem. In order to solve such problems, Chu and Chen [58] recently presented a systems-based approach for the identification of apoptosis drug targets. The selection of the drug targets by utilizing the approach can be viewed as a multistage discovery process. In the first stage, a protein–protein interaction network is constructed by a number of datasets and online interactome databases. In the second stage, a stochastic model of protein–protein interactions is constructed. In order to refine the model, false protein interactions are removed by utilizing an information theoretic measure, namely, Akaike's information criterion to microarray data. Finally, drug targets are identified by conducting a network-level comparison between normal and cancer cells.

Transcription factors-based methods can play an important role in devising an effective therapeutic and preventive interventions strategy for diseases. In Rosenberger et al. [59], the role of activating transcription factor 3 (ATF3) was investigated for murine cytomegalovirus (MCMV) infection. Mouse was used as the model system. The study demonstrated negative regulation of interferon-gamma (IFN-γ) expression caused by ATF3 in natural killer cells. The mice that had zero ATF3 exhibited high resistance to MCMV infection.

In another study, Nelander et al. [60] introduced a computational systems biology methodology for the prediction of pathway responses to combinatorial drug perturbations or drug combinations. The method is based on the use of multiple input–output model. Given that the linear models are not able to capture crucial information required for the task at hand, the authors presented nonlinear multiple input–output model. The approach was applied to analyze perturbations in MCF7 human breast carcinoma cells, where a number of compounds including rottlerin, rapamycin, and and so on were selected as perturbants. The leave-one-out cross-validation results showed the efficacy of the method.

Genetic causes of diseases can provide information that is crucial to design effective therapeutic approaches. A network that illustrates the association between diseases and their related genes can be highly informative. The human disease network presented in Goh et al. [61] is an attempt in this direction. The graph theoretic framework is based on the construction of a network to analyze and investigate the association between phenotypes and disease genes. In the constructed bipartite graph, one set of nodes represents genetic disorders and the second set denotes known disease genes in human genome. The edge between the disease and a gene represents the mutation in gene caused by the disease. The network provides a means to study novel patterns of gene disease associations.

Screening toxic compounds is a key issue in drug design and development. In Amini et al. [62], a novel computational methodology was introduced as an accurate means of predicting toxicity of compounds. The technique integrates two machine learning approaches, namely, SVMs [63] and inductive logic programming (ILP), and is termed support vector inductive logic programming (SVILP). The method works

by obtaining a set of rules from an ILP system, hence mapping the compounds into relational ILP space. The induced rules are then applied to compute the similarity between two compounds by the use of a novel kernel function. The function, given by an inner product in relational ILP space, is a weighted sum over all the common hypothesized rules. The ILP kernel is used in conjunction with SVMs to compute toxicity. The authors applied their method to a diverse and broad ranging toxicity dataset, namely, DSSTox [64]. The effectiveness of the method was established by using a cross-validation experimental methodology to predict the toxicity of the compounds. The results of the study confirmed the efficacy of the method for drug design and development. In Lodhi et al. [65], the method is extended to classify mutagens and recognize protein folds. The extended method learns a multiclass classifier by using a divide-and-conquer reduction strategy that divides multiclasses into binary groups and solves each individual problem by inducing an SVILP. The extended multiclass SVILP was successfully applied to classify compounds.

The database storing detailed kinetic knowledge can be a useful resource as it can provide information that is required to build models of biological processes. In order to provide such a knowledge base, a database of kinetic data, namely, KDBI, has been developed [66]. The database contains various types of data, including protein–protein interactions and protein–small molecule interactions. It includes 19,263 records, where 2635 entries belong to protein–protein interactions and 11,873 records contain information regarding protein–small molecule interactions. The database also comprises ordinary differential equations-based pathways models.

1.6 SOFTWARE TOOLS FOR SYSTEMS BIOLOGY

In this section, we will very briefly describe software tools that are designed for modeling, simulating, and analyzing complex biological processes. Bioconductor is a project that provided a number of useful tools for conducting systems biology-based studies. The design of effective infrastructure is crucial for the development of efficient and user-friendly tools. Software infrastructures may be developed by using only a basic computer language and generator (a software tool) [67]. Chapter 15 provides an in-depth description of a text mining tool for systems biology. Table 1.1 summarizes a number of software packages for studying and investigating biological systems.

SQUAD [68] is an example of modeling tools for systems biology. It constructs dynamic models of signaling networks, where the unavailability of kinetic data do not hinder its performance. The underlying methodology of the systems is based on the integration of Boolean and continuous modeling techniques. The implementation is written in Java, whereas C++ has been used to code algorithms for the computation of steady states. SQUAD supports a number of input formats, including NET (text file), MML (xml file), and SMBL (systems biology markup language). The system performs simulations as follows: It takes as input a directed graph representing the structure of the network. The steady states of the graph are identified by

Table 1.1 Software tools for systems biology

Tools	Biological systems	Input format	Platform
	Modeling		
SQUAD	Signaling and regulatory networks	XML, MML, and NET	Windows and Linux
CellNetAnalyzer	Metabolic, signaling, and regulatory networks	Network Composer and ASCII	All platforms (approximately)
BioTapestry	Signaling and regulatory networks	CSV and tabular	Linux, Mac, and Windows
	Sensitivity Analysis		
SBML-SAT	Signaling, regulatory and metabolic network	SBML	Linux, Mac, and Windows
	Visualization		
Cytoscape	Molecular interaction networks	MS Excel, SIF, and so on	All platforms (approximately)
CellProfiler	Cell images	DIB	Linux, Mac, and Windows

using a Boolean algorithm. Then, a dynamic model is constructed. Finally, a user can perform simulations. SQUAD has a user-friendly graphical interface and can be downloaded from `http://www.enfin.org/dokuwiki/doku.php?id=squad:start`.

CellNetAnalyzer [69] is a related software tool for modeling and analyzing biological process. It can be applied to analyze signaling, regulatory, and metabolic networks. The software tool is implemented in MATLAB, and C has been used to code some underlying techniques. The input data can be provided to CellNetAnalyzer by using Network Composer or ASCII file. It is available at `http://www.mpi-magdeburg.mpg.de/projects/cna/cna.html`.

BioTapestry [70] is another biological modeling tool. It can perform analysis and modeling of large biological networks. Linux, Windows, and Mac are supported platforms. BioTapestry is available at `http://www.biotapestry.org/`.

Sensitivity analysis is an important aspect of computational modeling for systems biology. SBML-SAT [71] performs sensitivity analysis of biological systems, and the systems are represented in the form of ordinary differential equations. It incorporates and implements a number of well-known sensitivity analysis techniques. Windows, Mac, and Linux are supported platforms. SBML-SAT is implemented in MATLAB, where the input data need to be coded in SBML format. It is available at `http://sysbio.molgen.mpg.de/SBML-SAT/`.

We now briefly describe Cytoscape [72] that facilitates the visualization and analysis of biological networks. It also allows data integration. The supported input formats are delimited text files, MS Excel, SIF (simple interaction format), SMBL, GO (gene

association), and so on. It enables the identification of active modules in biological networks. Cytoscape also allows export of network structures as images in different formats. Cytoscape is available at `http://www.cytoscape.org/`.

The development of CellProfiler [73, 74] is an attempt to study complex biological processes by using image analysis software packages. The tool comprises two components, namely, CellProfiler and CellProfiler Analyst. The images are processed by using CellProfiler. CellProfiler Analyst is applied to analyze the processed data produced by CellProfiler. The tool can analyze hundreds and thousands of images. It is characterized by its capability of recognizing nonmammalian cells and quantification of phenotypes. It supports processing and analysis of multidimensional images and can perform illumination correction and cell identification by using standard and advanced methods. The tool is implemented in MATLAB and is available for Windows, Unix, and Mac platforms. The software tool is available at `http://www.cellprofiler.org/`.

1.7 CONCLUSION

The review presented in the chapter shows that computational systems biology encompasses a range of complex problems and methodologies. We are witnessing a rapid development in the field that will revolutionize and give answers to unsolved questions in biology. Biotechnology will be on the forefront due to the influence of systems-based approaches on medicine, agriculture, and so on [75, 76]. We believe that the growing popularity of systems-based computational techniques to studying and analyzing biological processes will foster collaboration between researchers from diverse disciplines and will lead to significant development and progress in the field of computational systems biology.

REFERENCES

1. P. K. Sorger. A reductionist's systems biology: opinion. *Curr. Opin. Cell Biol.*, 17(1):9–11, 2005.
2. E. Klauschen, B. R. Angermann, and M. Meier-Schellersheim. Understanding disease by mouse click: the promise and potential of computational approaches in systems biology. *Clin. Exp. Immunol.*, 149:424–429, 2007.
3. H. Kitano. Computational systems biology. *Nature*, 420:206–210, 2002.
4. D. Noble. *The Music of Life*. Oxford University Press, 2006.
5. H. Kitano. Systems biology: a brief overview. *Science*, 295:1662–1664, 2002.
6. O. Wolkenhauer. Systems biology: the reincarnation of systems theory applied in biology. *Brief. Bioinform.*, 2(3):258–270, 2001.
7. N. Wiener. *Cybernetics: or Control and Communication in the Animal and the Machines*. MIT Press, 2000 (first edition 1948).
8. A. Bellouquid and M. Delitala. *Mathematical Modeling of Complex Biological Systems: A Kinetic Theory Approach*. Birkhauer, 2006.

9. H. Lodhi and S. Muggleton. Modelling metabolic pathways using stochastic logic programs-based ensemble methods. In V. Danos and V. Schachter, editors. *Second International Conference on Computational Methods in System Biology (CMSB-04)*, LNCS. Springer, 2004, pp. 119–133.
10. J. Southern, J. Pitt-Francis, J. Whiteley, D. Stokeley, H. Kobashi, R. Nobes, Y. Kadooka, and D. Gavaghan. Multi-scale computational modelling in biology and physiology. *Prog. Bio. Mol. Biol.*, 96:60–89, 2008.
11. D. Noble. Modeling the heart. *Physiology*, 19(4):191–197, 2004.
12. P. J. Hunter, E. J. Crampin, and P. M Nielsen. Bioinformatics, multiscale modeling and the IUPS Physiome Project. *Brief. Bioinform.*, 9(4):333–343, 2008.
13. L. Zhang, C. A. Athale, and T. S. Deisboeck. Development of a three-dimensional multiscale agent–based tumor model: simulating gene-protein interaction profiles, cell phenotypes and multicellular patterns in brain cancer. *J. Theor. Biol.*, 244(1):96–107, 2007.
14. C. Athale, Y. Mansury, and T. Deisboeck. Simulating the impact of a molecular 'decision-process' on cellular phenotype and multicellular patterns in brain tumors. *J. Theor. Biol.*, 233(4):469–481, 2005.
15. J. J. Tyson and B. Novak. Regulation of the eukaryotic cell cycle: molecular antagonism, hysteresis, and irreversible transitions. *J. Theor. Biol.*, 210(2):249–263, 2001.
16. T. Alacron, H. M. Byrne, and P. K. Maini. A mathematical model of the effects of hypoxia on the cell-cycle of normal and cancer cells. *J. Theor. Biol.*, 229(3):395–411, 2004.
17. J. Mora, N. K. Cheung, and W. L. Gerald. Genetic heterogeneity and clonal evolution in neuroblastoma. *Br. J. Cancer*, 85(2):182–189, 2001.
18. S. A. Hill, S. Wilson, and A. F. Chambers. Clonal heterogeneity, experimental metastatic ability, and p21 expression in H-ras-transformed NIH 3T3 cells. *J. Natl. Cancer Inst.*, 80(7):484–490, 1988.
19. L. Zhang, C. G. Strouthos, Z. Wang, and T. S. Deisboeck. Simulating brain tumor heterogeneity with a multiscale agent-based model: linking molecular signatures, phenotypes and expansion rate. *Math. Comput. Model.*, 49(1–2):307–319, 2009.
20. S. H. Robertson, C. K. Smith, A. L. Langhans, S. E. McLinden, M. A. Oberhardt, K. R. Jakab, B. Dzamba, D. W. DeSimone, J. A. Papin, and S. M. Peirce. Multiscale computational analysis of *Xenopus laevis* morphogenesis reveals key insights of systems-level behaviour. *BMC Syst. Biol.*, 1(46), 2007.
21. E. Lee, A. Salic, R. Kruger, R. Heinrich, and M. W. Kirschner. The roles of APc and Axin derived from experimental and theoretical analysis of the Wnt pathway. *PLoS Biol.*, 1:116–132, 2003.
22. M. Nagel, E. Tahinci, K. Symes, and R. Winklbauer. Guidance of mesoderm cell migration in the *Xenopus gastrula* requires PDGF signalling. *Development*, 131:2727–2736, 2004.
23. A. C. Cuellar and P. Lloyd. An overview of CellML 1.1, a biological model description language. *Simulation: Trans. Soc. Model. Simul. Int.*, 79(12):740–747, 2003.
24. G. Shinar, R. Milo, M. R. Martinez, and U. Alon. Input–output robustness in simple bacterial signaling systems. *Proc. Natl. Acad. Sci. USA*, 104(50):19931–19935, 2007.
25. R. Aebersold and D. R. Goodlett. Mass spectrometry in proteomics. *Chem. Rev.*, 101(2):269–295, 2001.
26. A. D. Weston and L. Hood. Systems biology, proteomics and the future of health care: toward predictive, preventative and personalized medicine. *J. Proteome Res.*, 3(2):179–196, 2004.

27. X. Feng, K. Liu, Q. Luo, and B.-F. Liu. Mass spectrometry in systems biology: an overview. *Mass Spectrom. Rev.*, 27(6):635–660, 2008.

28. S. Rho, S. You, and D. Hwang. From proteomics towards systems biology: integration of different types of proteomics data into network models. *BMB Rep.*, 41(3):184–193, 2008.

29. X-M. Zhao, R-S. Wang, L. Chen, and K. Aihara. Uncovering signal transduction networks from high-throughput data by integer linear programming. *Nucleic Acids Res.*, 36(9):e48, 2008.

30. K. Collins, T. Jacks, and N. P. Pavletich. The cell cycle and cancer. *Proc. Natl. Acad. Sci. USA*, 94:2776–2778, 1997.

31. M. Macaluso, G. Russo, C. Cinti, V. Bazan, N. Gebbia, and A. Russo. Ras family genes: an interesting link between cell cycle and cancer. *J. Cell. Physiol.*, 192:125–130, 2002.

32. A. Sigal, R. Milo, A. Cohen, N. Geva-Zatrosky, I. Alaluf, N. Swerdlin, N. Perzov, T. Danon, Y. Liron, T. Raveh, A. E. Carpenter, G. Lahav, and U. Alon. Dynamic proteomics in individual human cells uncovers widespread cell-cycle dependence of nuclear proteins. *Nat. Methods*, 3(7):525–531, 2006.

33. R. Narayanaswamy, E. K. Moradi, W. Niu, G. T. Hart, M. Davis, K. L. McGray, A. D. Ellington, and E. M. Marcotte. Systematic definition of protein constituents along the major polarization axis reveals an adaptive reuse of the polarization machinery in pheromone-treated budding yeast. *J. Proteome Res.*, 8:6–19, 2009.

34. K. Shinoda, M. Tomita, and Y. Ishihama. Aligning LC peaks by converting gradient retention times to retention index of peptides in proteomic experiments. *Bioinformatics*, 24(14):1590–1595, 2008.

35. L. Hayflick. *How and Why We Age*. Ballantine Books, 1994.

36. J. Campisi. Replicative senescene: an old lives' tale. *Cell*, 84:497–500, 1996.

37. J. P. de Magalhaes and R. G. A. Faragher. Cell divisions and mammalian aging: integrative biology insights from genes that regulate longevity. *Bioessays*, 30(6):567–578, 2008.

38. T. B. L. Kirkwood. Understanding the odd science of aging. *Cell*, 120, 2005.

39. V. D. Longo, M. R. Leiber, and J. Vijg. Turning anti-aging genes against cancer. *Mol. Cell Biol.*, 9:903–910, 2008.

40. V. D. Longo and C. E. Finch. Evolutionary medicine: from dwarf model systems to healthy centenarians. *Science*, 299:1342–1346, 2003.

41. C. Kenyon. The plasticity of aging: insights from long-lived mutant. *Cell*, 120(4):449–460, 2005.

42. J. P. de Magalhaes and G. M. Church. Analyses of human–chimpanzee orthologous gene pairs to explore evolutionary hypotheses of aging. *Mech. Ageing Dev.*, 128:355–364, 2007.

43. S. K. Kim. Common aging pathways in worms, flies, mice and humans. *J. Exp. Biol.*, 210(9):1607–1612, 2007.

44. S. K. Kim. Genome-wide views of aging gene networks. *Molecular Biology of Aging*. Cold Spring Harbor Laboratory Press, 2008.

45. Y. V. Budovskaya, K. Wu, L. K. Southworth, M. Jiang, P. Tedesco, and T. E. Johnson. An elt-3/elt-5/elt-6 GATA transcription circuit guides aging in *C. elegans. Cell*, 134:1–13, 2008.

46. H. Xue, B. Xian, D. Dong, K. Xia, S. Zhu, Z. Zhang, L. Hou, Q. Zhang, Y. Zhang, and J.-D. J. Han. A modular network model of aging. *Mol. Syst. Biol.*, 3(147), 2007.

47. K. Xia, D. Dong, H. Xue, S. Zhu, J. Wand, Q. Zhang, L. Hou, H. Chen, R. Tao, Z. Huang, Z. Fu, Y. G. Chen, and J. D. Han. Identification of the proliferation/differentiation switch in the cellular network of multicellular organisms. *PLoS Comput. Biol.*, 2(11):e145, 2006.
48. S. A. Garan, W. Freitag, V. Caspo, P. Chrysler, B. Rizvi, and N. Shewaramani. A computational systems biology approach to neuroendocrine aging: initial results. *Exp. Gerontol.*, 42:142, 2007.
49. W. R. Swindell, J. M. Harper, and R. A. Miller. How long will my mouse line? Machine approaches for prediction of mouse life span. *J. Gerontol.*, 63A(9):895–906, 2008.
50. S. Krivenko and M. Burtsev. Simulation of the evolution of aging: effects of aggression and kin-recognition. In *Advances in Artificial Life, 9th European Conference, ECAL. Notes in Computer Science* 4648, 2007, pp. 84–92.
51. A. Bergman, G. Atzmon, K. Ye, T. MacCarthy, and N. Barzilai. Buffering mechanisms in aging: a systems approach toward uncovering the genetic component of aging. *PLoS Comput. Biol.*, 3(8):e170, 2007.
52. W. Materi and S. Wishart. Computational systems biology in drug discovery and development: methods and applications. *Drug Discov. Today*, 12(7–8):295–303, 2007.
53. C. R. Cho, M. Labow, M. Reinhardt, J. van Oostrum, and M. C. Peitsch. The application of systems biology to drug discovery. *Curr. Opin. Chem. Biol.*, 10(4):294–302, 2006.
54. N. Kumar, B. S. Hendriks, K. A. Janes, D. de Graaf, and D. A. Lauffenburger. Applying computational modeling to drug discovery and development. *Drug Discov. Today*, 11(17–18):806–811, 2006.
55. M. L. Billingsley. Druggable targets and targeted drugs: enhancing the development of new therapeutics. *Pharmacology*, 82:239–244, 2008.
56. E. C. Butcher, E. L. Berg, and E. J. Kunkel. Systems biology in drug discovery. *Nat. Biotechnol.*, 22(10):1253–1259, 2004.
57. H. Kitano. A robustness-based approach to systems-oriented drug design. *Nat. Rev. Drug Discovery*, 6:202–210, 2007.
58. L.-H. Chu and B.-S. Chen. Construction of a cancer-perturbed protein–protein interaction network for discovery of apoptosis drug targets. *BMC Syst. Biol.*, 2(56), 2008.
59. C.M. Rosenberger, A. E. Clark, P. M. Treuting, C. D. Jhonson, and A. Aderem. ATF3 regulates MCMV infection in mice by modulating IFN-γ expression in natural killer cells. *Proc. Natl. Acad. Sci. USA*, 105(7):2544–2549, 2008.
60. S. Nelander, W. Wang, B. Nilsson, C. Pratilas Q.-B. She, N. Rossen, and P. Gennemark. Models from experiments: combinatorial drug perturbations of cancer cells. *Mol. Syst. Biol.*, 4(216), 2008.
61. K.-II. Goh, M. C. Cusick, D. Valle, B. Childs, M. Vidal, and A.-L Barabasi. The human disease network. *Proc. Natl. Acad. Sci. USA*, 104(21):8685–8690, 2007.
62. A. Amini, S. Muggleton, H. Lodhi, and M.J.E. Sternberg. A novel logic-based approach for quantitative toxicology prediction. *J. Chem. Inform. Model.*, 47(3):998–1006, 2007.
63. V. Vapnik. *The Nature of Statistical Learning Theory*. Springer, New York, 1995.
64. A.M. Richard and C.R. Williams. Distributed structure-searchable toxicity (DSSTox) public database network: a proposal. *Mutat. Res.*, 499:27–52, 2000.
65. H. Lodhi, S. Muggleton and M. J. E Sternberg. *Learning Large Margin First Order Decision Lists for Multi-Class Classification*. In J. Gama, V. S. Costa, A. M. Jorge and P. B. Brazdil, editors. Discovery Science (DS) 2009, LNCS (LNAI) 5808, Springer, 168–183, 2009.

66. P. Kumar, B. C. Han, Z. Shi, J. Jia, Y. P. Wang, Y. T. Zhang, L. Liang, Q. F. Liu, Z. L. Ji, and Y. Z. Chen. Update of KDBI: kinetic data of bio-molecular interaction database. *Nucleic Acids Res.*, 37:D636–D641, 2009.

67. M. A. Swertz and R. S. Jansen. Beyond standardization: dynamic software infrastructure for systems biology. *Nat. Rev. Genet.*, 8:235–243, 2007.

68. A. D. Cara, A. Garg, G. D. Micheli, I. Xenarios, and L. Mendoza. Dynamic simulation of regulatory networks using SQUAD. *BMC Bioinform.*, 8(462), 2007.

69. S. Klamt, J. Saez-Rodriguez, and E. D. Gilles. Structural and functional analysis of cellular networks with CellNetAnalyzer. *BMC Syst. Biol.*, 1(2), 2007.

70. W. J. Longabaugh, E. H. Davidson, and H. Bolouri. Computational representation of developmental genetic regulatory networks. *Dev. Biol.*, 283(1):1–16, 2005.

71. Z. Zi, Y. Zheng, A. E. Rundell, and E. Klipp. SBML-SAT: a systems biology markup language (SBML) based sensitivity analysis tool. *BMC Bioinform.*, 9(342), 2008.

72. M. S. Cline, et al. Integration of biological networks and gene expression data using Cytoscape. *Nat. Protocols*, 2(10):2366–2382, 2007.

73. A. E. Carpenter, T. R. Jones, M. R. Lamprecht, C. Clarke, I. H. Kang, O. Friman, D. A. Guertin, J. H. Chang, R. A. Lindquist, J. Moffat, P. Golland, and D. M. Sabatini. CellProfiler: image analysis software for identifying and quantifying cell phenotype. *Genome Biol.*, 7(R:100), 2006.

74. M. P. Lamprecht, D. M. Sabatini, and A. E. Carpenter. CellProfiler: free, versatile software for automated biological image analysis. *Biotechniques*, 42(1):71–75, 2007.

75. T. Ideker, T. Galitski, and L. Hood. A new approach to decoding life: systems biology. *Annu. Rev. Genomics Hum. Genet.*, 2:343–372, 2001.

76. A. Aderem. Systems biology: its practices and challenges. *Cell*, 121:511–513, 2005.