Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

CHAPTER

22

In silico disease model: from simple networks to complex diseases

Debmalya Barh[1], *Eugenia Ch. Yiannakopoulou*[2], *Emmanuel O. Salawu*[3,4], *Atanu Bhattacharjee*[5], *Sudhir Chowbina*[6], *Joseph J. Nalluri*[7], *Preetam Ghosh*[7] *and Vasco Azevedo*[8]

[1]Centre for Genomics and Applied Gene Technology, Institute of Integrative Omics and Applied Biotechnology (IIOAB), Nonakuri, Purba Medinipur, West Bengal, India [2]Department of Basic Medical Lessons Faculty of Health and Caring Professions, Technological Educational Institute of Athens, Athens, Greece [3]Institute of Bioinformatics and Structural Biology, National TsingHua University, Taipei, Taiwan [4]Institute of Information Science, Academia Sinica, Taipei, Taiwan [5]Bioinformatics Laboratory, Department of Biotechnology and Bioinformatics, North Eastern Hill University, Shillong, India [6]Advanced Biomedical Computing Center, SAIC-Frederick, Inc., Frederick National Laboratory for Cancer Research, National Cancer Institute, Frederick, MD, United States [7]Department of Computer Science, Virginia Commonwealth University, Richmond, VA, United States [8]Department of General Biology, Federal University of Minas Gerais, Belo Horizonte, Brazil

Summary

In this chapter we consider in silico modeling of diseases starting from some simple to some complex (and mathematical) concepts. Examples and applications of in silico modeling for some important categories of diseases (such as for cancers, infectious diseases, and neuronal diseases) are also given.

What you can expect to know

Recent advances in bioinformatics and systems biology enable modeling and simulation of subcellular, cellular processes, and disease, using primary methods from dynamical systems theory. In this approach, all interactions among all components in a system are described mathematically and computed models are established. These in silico models encode and test hypotheses about mechanisms underlying the function of cells, the pathogenesis, and pathophysiology of disease and contribute to the identification of new drug targets and drug design. The development of in silico models is facilitated by rapidly advancing experimental and analytical tools that generate information-rich, high-throughput biological data. Bioinformatics providing tools for pattern recognition, machine learning, statistical modeling, and data extraction from databases contribute to in silico modeling. Dynamical systems theory is the natural language to investigate complex biological systems demonstrating nonlinear spatiotemporal behavior. Most of the in silico models aim to complement and not replace experimental research. Experimental data are needed for parameterization, calibration, and validation of in silico models. Typical examples in biology are models for molecular networks, where the behavior of cells is expressed in terms of quantitative changes in the levels of transcripts and gene products as well as models of the cell cycle. In medicine, in silico models of cancer, immunological disease, lung disease, and infectious diseases complement conventional research with in vitro models, animal models, and clinical trials. This chapter presents basic concepts of bioinformatics, systems biology, and their applications in in silico modeling and reviews its applications in biology and disease.

Animal Biotechnology.
DOI: https://doi.org/10.1016/B978-0-12-811710-1.00020-3

© 2020 Elsevier Inc. All rights reserved.

Bioinformatics in animal biotechnology

Biotechnology is the most promising life science frontiers for the next decade. Together with informatics, biotechnology is leading to revolutionary changes in our society and economy. This genomic revolution is global and is creating new prospects in all biological sciences including medicine, human health, disease, and nutrition, agronomy, and animal biotechnology.

Animal biotechnology is the source of innovation in production and processing, profoundly impacting the animal husbandry sector which seeks to improve animal product quality, health, and well-being. Biotechnological research products, such as vaccines, diagnostics, in vitro fertilization, transgenic animals, stem cells, and a number of other therapeutic recombinant products are now commercially available. In view of the immense potential of biotechnology in the livestock and poultry sectors, the interest in animal biotechnology has increased over the years.

The fundamental requirement to modern biotechnology projects is the ability to gather, store, classify, analyze, and distribute biological information derived from genomics projects. Bioinformatics deals with methods for storing, retrieving, and analyzing biological data and protein sequences, structures, functions, pathways, and networks, and recently in silico disease modeling and simulation using systems biology. Bioinformatics encompasses both conceptual and practical tools for the propagation, generation, processing, and understanding of scientific ideas and biological information (Rhee, 2005).

Genomics is the scientific study of structure, function, and interrelationships of both individual genes and the genome. Recently, genomics research has facilitated the whole-genome mapping of various animals and understanding of the building blocks of biology. This has enabled researchers to decipher the fundamental cellular functions at the DNA level such as gene regulation or protein–protein interactions (PPIs) and thus discover molecular signatures (cluster of genes, proteins, metabolites, etc.), which are characteristic of a biological process or of a specific phenotype. Bioinformatics methods and databases can be developed to provide solutions to the challenges of handling massive amounts of data.

The history of animal biotechnology with bioinformatics is to make a strong research community that will build the resources and support the veterinary and agricultural research. There are some technologies that were used dating back to 5000 BCE. Many of these techniques are still being used today. Hybridizing animals by crossing specific strains of animals so that they can create greater genetic varieties. The offspring of some of these crosses is afterward bred selectively to produce the most desirable traits in those specific animals.

There has been significant interest in the complete analysis of the genome sequence of the farm animals such as chicken, pig, cattle, sheep, fish, and rabbit. The genomes of farm animals have been altered to search for preferred phenotypic traits and then selecting for better-quality animals to continue into the next generation. Access to these sequences has given rise to genome array chips and a number of web-based mapping tools and bioinformatics tools required to make sense of the data. In addition, the organization of gigabytes of sequence data requires an efficient bioinformatics database. Fadiel et al. provide a nice overview of resources related to the farm animal bioinformatics and genome projects (Fadiel et al., 2005).

With farm animals consuming large amounts of genetically modified crops, such as modified corn and soybean crops, it is a good question as to effect this will have on the meats. Some of the benefits of this technology are that what once took many years of trial and error is not completed in just months. The meats that are being produced are coming from animals that are better nourished by the use of biotechnology. Biotechnology and conventional approaches are benefiting both poultry and livestock producers. This will give a more wholesome affordable product that will meet the growing population demands.

Moreover, bioinformatics methods devoted to investigating the genomes of farm animals can bring eventual economic benefits, such as ensuring food safety and better food quality in the case of beef. Recent advances in high-throughput DNA sequencing techniques, microarray technology, and proteomics have led to effective research in bovine muscle physiology, to improve beef quality either by breeding or rearing factors. Bioinformatics is a key tool to analyze the huge data sets obtained from these techniques. The computational analysis of global gene expression profiling at the mRNA or protein level has shown that previously unsuspected genes may be associated either with muscle development or growth and may lead to the development of new molecular indicators of tenderness (Hocquette et al., 2007). Gene expression profiling has been used to document changes in gene expression, for example, following infection by pathological organisms (Meade et al., 2006), during the metabolic changes imposed by lactation in dairy cows (Loor et al., 2005), in cloned bovine embryos (Somers et al., 2006) and in various other models.

Bioinformatics enrichment tools are playing an important role in facilitating the functional analysis of large gene lists from various high-throughput biological studies. Huang et al. discuss 68 bioinformatics enrichment tools which help us understand their algorithms and details of a particular tool (Huang et al., 2009). However in biology, genes do not act

independently, but in a highly coordinated and interdependent manner with each other. In order to understand the biological meaning, one needs to map these genes into gene-ontology (GO) categories or metabolic and regulatory pathways. Different bioinformatics approaches and tools are employed for this task, starting from GO-ranking methods, pathway mappings, and biological network analysis (Werner, 2008). The awareness of these resources and methods is essential to make the best choices for our research interests.

The knowledge of bioinformatics tools will facilitate its wide application in the field of animal biotechnology. Bioinformatics is the computational data management discipline that helps us gather, analyze, and represent this information in order to educate ourselves, understand biological processes in the healthy and disease states, and facilitate the discovery of better animal products. Continued efforts are required to develop cost-effective and efficient computation platforms that can retrieve, integrate, and interpret the knowledge behind the genome sequences. The application of bioinformatics tools for biotechnology research will have significant implications in life sciences and the betterment of human lives. Bioinformatics is being adopted worldwide by academic groups, companies, and national and international research groups, and it should be thought of as an important pillar of current and future biotechnology, without which a rapid progress in the field is not possible. Systems approaches in combination with genomics, proteomics, metabolomics, and kinomics data have tremendous potential in providing insights into various biological mechanisms including the most important human diseases.

Bioinformatics and systems biology

We are witnessing the birth of a new era in biology. The ability to decipher the genetic code of living organisms promises to improve the quality of human life has dramatically changed the landscape of the biological and biomedical sciences and has brought with it new challenges.

One such challenge is that recent and novel technologies produce biological data sets of ever-increasing size, including genomic sequences, RNA, and protein abundances, their interactions with each other, and the identity and abundance of other biological molecules. The storage and compilation of such quantities of biological data is a challenge: the human genome, for example, contains three billion chemical units of DNA, whereas a protozoan genome has 670 billion units of DNA. Data management and interpretation require the development of newly sophisticated computational methods based on research in biology, medicine, pharmacology, and agricultural studies and using methods from computer science and mathematics—in other words, the multidisciplinary subject of bioinformatics.

Bioinformatics enables researchers to store large data sets in a standard computer database format and provides tools and algorithms scientists use to extract integrated information from the databases and use it to create hypotheses and models. Bioinformatics is a growth area because almost every experiment now involves multiple sources of data, requiring the ability to handle those data and to draw out inferences and knowledge. After 15 years of rapid evolution, the subject is now quite ubiquitous.

Due to the tremendous surge in the availability of biological data, the term "big data" is often used to describe the nature, scale, and dimension of these data. Moreover, the upcoming technologies must address and account for new schemas for the complicated and unstructured nature of biological data while designing data repositories. Currently, data repositories like the National Center for Biotechnology (Sayers et al., 2011), GenBank (Benson et al., 2013), Gene Expression Omnibus (Edgar et al., 2002), and Protein Data Bank (Berman et al., 2000) are some of the major bioinformatics resources for research in this domain.

Another challenge lies in deciphering the integrated functions of thousands of genes or systems biology. Systems biology is a term used to describe a number of trends in bioscience research and a movement that draws on those trends. It can be described as a biology-based interdisciplinary field of study that focuses on complex interactions of biological systems. Those in the field claim that it represents a shift in perspective toward holism instead of reduction. Systems biology brings in the aspect of complex systems to cell biology. The fundamental guiding principle of the field of systems biology is that of studying the cell as a whole entity and not merely in parts or isolation. It treats the cell as one functional component with a highly efficient integrated machinery of molecular interactions causing activation of cellular interactions which lead to effects in tissues and organs that subsequently manifest in physiological functions. This integrated understanding makes systems biology truly interdisciplinary combing fields of biology, mathematics, physics, computer science, and electrical engineering.

Systems biology has great potential to facilitate the development of drugs to treat specific diseases. The drugs currently on the market can target only those proteins that are known to cause disease. However with the human genome now completely mapped, we can target the interaction of genes and proteins at a systems biology level. This will enable the pharmaceutical industry to design drugs that will only target those genes that are diseased, improving healthcare in

the United States. Like two organs in one body, systems analysis and bioinformatics are separate but interdependent.

Common computational methods in systems biology

Computational methods take an interdisciplinary approach, involving mathematicians, chemists, biologists, biochemists, and biomedical engineers. The robustness of data sets related to gene interaction and co-operation at the system level requires multifaceted approaches to create a hypothesis that can be tested. Two approaches are used to understand the network interactions in systems biology, namely, Experimental and Theoretical and Modeling techniques (Choi, 2007). Below is a detailed overview of the different computational/ bioinformatics methods in modern systems biology.

Experimental methods in systems biology

Experimental methods utilize real situations to test the hypothesis of mined data sets. As such, living organisms are used whereby various aspects of genome-wide measurements and interactions are monitored. Specific examples on this point include the following.

Protein–protein interactions

PPIs predictions are methods used to predict the outcome of pairs or groups of protein interactions. These predictions are done in vivo and various methods can be used to carry out the predictions. Interaction prediction is important as it helps researchers make inferences of the outcomes of PPI. PPI can be studied by phylogenetic profiling, identifying structural patterns and homologous pairs, intracellular localization, and posttranslational modifications among others (Choi, 2007). A survey of available tools and web servers for analysis of PPIs is provided by Tuncbag et al. (2009).

Transcriptional control networks

Within biological systems, several activities involving the basic units of a gene take place. Such processes as DNA replication, RNA translation, and transcription into proteins must be controlled; otherwise, the systems could yield numerous destructive or useless gene products. Transcriptional control networks, also called gene regulatory networks, are segments within the DNA that govern the rate and product of each gene.

Bioinformatics has devised methods to look for destroyed, dormant, or unresponsive control networks. The discovery of such networks helps in corrective therapy, hence, the ability to control some diseases resulting from such control networks breakdown (Choi, 2007). There has also been rapid progress in the development of computational methods for the genome-wide "reverse engineering" of such networks. ARACNE is an algorithm to identify direct transcriptional interactions in mammalian cellular networks and promises to enhance our ability to use microarray data to elucidate cellular processes and to identify molecular targets of pharmacological drugs in mammalian cellular networks (Margolin et al., 2006). In addition to methods like ARACNE, we need systems biology approaches that incorporate heterogeneous data sources, such as genome sequence and protein–DNA interaction data. The development of such computational modeling techniques to include diverse types of molecular biological information clearly supports the gene regulatory network inference process and enables the modeling of the dynamics of gene regulatory systems (Hecker et al., 2009).

Signal transduction networks

Signal transduction is how cells communicate with each other. Signal transduction pathways involve interactions between proteins, micro and macromolecules and DNA. A breakdown in signal transduction pathways could lead to detrimental consequences within the system due to the lack of integrated communication. Correction of broken signal transduction pathways is a therapeutic approach researched for use in many areas of medicine.

The high-throughput and multiplex techniques for quantifying signaling and cellular responses are being increasingly available and affordable. A high-throughput quantitative multiplex kinase assay (Janes et al., 2003), mass spectrometry-based proteomics (Ong and Mann, 2005), and single-cell proteomics (Irish et al., 2006) are few of the experimental methods to elucidate signal transduction mechanisms of cells. These large-scale experiments are generating large data sets on protein abundance and signaling activity. Data-driven modeling approaches such as clustering, principal components analysis, and partial least squares need to be developed to derive biological hypothesis. The potential of data-driven models to study large-scale data sets quantitatively and

comprehensively will make sure that these methods will emerge as standard tools for understanding signal transduction networks (Janes and Yaffe, 2006).

Mathematical modeling techniques

Modeling biological systems is a major task of systems biology and mathematical biology. Computational systems biology intends to develop and use efficient algorithms, data structures, visualization, and communication tools with the goal of computer modeling of biological systems. A mathematical model can provide new insights into a biological model of interest and help in generating testable predictions.

Modeling or simulation can be viewed as a way of creating an artificial biological system in vitro whose properties can be changed or made dynamic. By externally controlling the model, new data sets can be created and implemented at the system level to create novel insights in treating gene-related problems. In modeling and simulation, sets of differential equations and logic clauses are used to create a dynamic systems environment that can be tested.

Mathematical models of biochemical networks (signal transduction cascades, metabolic pathways, and gene regulatory networks) are a central component of modern systems biology. The development of formal methods adopted from theoretical computing science is essential for the modeling and simulation of these complex networks (de Jong, 2002; Breitling et al., 2008). The computational methods that are being employed in mathematical biology and bioinformatics are (1) directed graphs, (2) Bayesian networks, (3) Boolean networks and their generalizations, (4) ordinary and partial differential equations, (5) qualitative differential equations, (6) stochastic equations, and (7) rule-based formalisms. Below are a few specific examples of the applications of these methods.

Mathematical models can be used to predict drug response or causes of drug resistance under a given set of conditions based on specific tumor properties. This integration can help in the development of tools that aid in the diagnosis, prognosis, and thus improve treatment outcome in patients with cancer. For example, in breast cancer being one of the well-studied diseases over the last decade serves as a model disease. One can thus apply the principles of molecular biology and pathology in designing new predictive mathematical frameworks that can unravel the dynamic nature of the disease. Genetic mutations of BRCA1, BRCA2, TP53, and PTEN significantly affect disease prognosis and increases the likelihood of adverse reactions to certain therapies. These mutations enable normal cells to become self-sufficient in survival in a stepwise process. Enderling et al. (2006) have modeled this mutation and expansion process by assuming that mutations in two tumor suppressor genes are sufficient to give rise to cancer. They modify the earlier model of Enderling et al. which is based on an established partial differential equation model of solid tumor growth and invasion (Anderson et al., 2006). The stepwise mutations from a normal breast stem cell to a tumor cell have been described using a model consisting of four differential equations.

Lauffenburger has applied a novel graphical modeling methodology known as Bayesian network analysis to model discovery and model selection for signaling events that direct mouse embryonic stem, an important preliminary step in hypothesis testing, in protein signaling networks. The model predicts bidirectional dependence between the two molecules ERK and FAK. It is interesting to appreciate that the apparent complexity of these dynamic ERK-FAK interactions is quite likely responsible for the difficulty in determining clear "upstream" versus "downstream" influence relationships by means of standard molecular cell biology methods. Bayesian networks determine the relative probability of statistical dependence models of arbitrary complexity for a given set of data (Woolf et al., 2005). This method offers further clues to apply Bayesian approaches to cancer biology problems.

The cell cycle is a process in which cells proliferate while collectively performing a series of coordinated actions. Cell-cycle models also have an impact on drug discovery. Chassagnole et al. use a mathematical model to simulate and unravel the effect of multitarget kinase inhibitors of cyclin-dependent kinases. They quantitatively predict the cytotoxicity of a set of kinase inhibitors based on the in vitro IC50 measurement values. Finally, they assess the pharmaceutical value of these inhibitors as anticancer therapeutics (Chassagnole et al., 2006).

In cancer, avascular tumor growth is characterized by localized, benign tumor growth where the nearby tissues consume most of the nutrients. Mathematical modeling of avascular tumor growth is important to understand the advanced stages of cancer. Kiran et al. have developed a spatial-temporal mathematical model classified as different zone model for avascular tumor growth based on the diffusion of nutrients, their consumption, and it includes key mechanisms in the tumor. The diffusion and nutrient consumption are represented using partial differential equations. This model predicts that the onset of necrosis occurs when the concentrations of vital nutrients are below critical values and also the overall tumor growth based on the size effects of proliferation zone, quiescent zone, and necrotic zone (Kiran et al., 2009).

The mathematical approaches used to model the three natural scales of interest: subcellular, cellular, and tissue, as discussed above. A challenge is the development of models that predict effects across biological scales. The long-term goal is to build a "virtual human made up of mathematical models with connections at the different biological scales (from genes to tissue to the organ)" (Brook et al., 2011).

Concept of modeling

A model is an optimal mix of hypotheses, evidence, and abstraction to explain a phenomenon. The hypothesis is a tentative explanation for an observation, phenomenon, or scientific problem that can be tested by further investigation. Evidence describes the information, that is, experimental data that help in forming a conclusion or judgment. Abstraction is an act of filtering out the required information to focus on a specific property only. For example, archiving books based on the year of publication, irrespective of the author name, would be an example of abstraction. In this process, we lose some detail and gain some. Through modeling, predictions are made, that may be tested by experiment. A model may be simple, for example, the logistic equation describing how a population of bacteria grows or the model may be complicated. Models may be mathematical or statistical (Coveney and Fowler, 2005).

Mathematical models make predictions, whereas statistical models enable us to draw statistical inferences about the probable properties of a system. In other words, models can be deductive or inductive. If the prediction is necessarily true given that the model is also true, then the model is a deductive model. On the other hand, if the prediction is statistically inferred from observations, then the model is inductive. Deductive models contain a mathematical description, for example, the reaction–diffusion equations that make predictions about reality. If these predictions do not agree with experiment, then the validity of the entire model may be questioned. Mathematical models are commonly applied in physical sciences. On the other hand, inductive models are mostly applied in the biological sciences. In biology, models are used to describe, simulate, analyze, and predict the behavior of biological systems. Modeling in biology provides a framework that enables description and understanding of biological systems through building equations that express biological knowledge. Modeling enables the simulation of the behavior of a biological system by performing in silico experiments, that is, numerically solving the equations/rules that describe the model. The results of these in silico experiments become the input for further analysis, e.g., identification of key parameters or mechanisms, interpretation of data, or comparison of the ability of different mechanisms to generate observed data.

In particular, systems biology employs an integrative approach to characterize biological systems, in which interactions among all components in a system are described mathematically to establish a computable model. These in silico models complement traditional in vivo animal models and can be applied to quantitatively study the behavior of a system of interacting components. The advent of high-throughput experimental tools has allowed for the simultaneous measurement of thousands of biomolecules, opening the way for in silico model construction of increasingly large and diverse biological systems. Integrating heterogeneous dynamic data into quantitative predictive models holds great promise to significantly increase our ability to understand and rationally intervene in disease-perturbed biological systems. This promise—particularly with regard to personalized medicine and medical intervention—has motivated the development of new methods for systems analysis of human biology and disease. Such approaches offer the possibility of gaining new insights into the behavior of biological systems, of providing new frameworks for organizing and storing data and performing statistical analyses, of suggesting new hypotheses and new experiments, and even of offering a "virtual laboratory" to supplement in vivo and in vitro work.

However in silico modeling in the life sciences is far from straightforward and suffers from a number of potential pitfalls. Thus mathematically sophisticated but biologically useless models often arise because of a lack of biological input, leading to models that are biologically unrealistic or address a question of little biological importance. On the other hand, models may be biologically realistic but mathematically intractable. This problem usually arises because biologists unfamiliar with the limitations of mathematical analysis want to include every known biological effect in the model. Even if it were possible to produce such models they would be of little use since their behavior would be as complex to investigate as the experimental situation. These problems can be avoided by formulating clear explicit biological goals before attempting to construct a model. This will ensure that the resulting model is biologically sound, can be experimentally verified, and will generate biological insight or new biological hypotheses. The aim of a model should not simply be to reproduce the biological data, and indeed often the most useful models are those that exhibit discrepancies from the experiment. Such deviations will typically stimulate new experiments or hypotheses. An iterative approach has

been proposed, starting with a biological problem, developing a mathematical model, and then feeding back into the biology. Once established, this collaborative loop can be traversed many times, leading to ever-increasing understanding.

The ultimate goal of in silico modeling in biology is the detailed understanding of the function of molecular networks as they appear in metabolism, gene regulation, or signal transduction. This is achieved by using a level of mathematical abstraction that needs a minimum of biological information to capture all physiologically relevant features of a cellular network. For example, ideally, for in silico modeling of a molecular network, knowledge of the network structure, of all reaction rates, concentrations, and spatial distributions of molecules at any time point is needed. Unfortunately, such information is unavailable even for the best-studied systems. In silico simulations thus always have to use a level of mathematical abstraction, which is dictated by the extent of our biological knowledge, by molecular details of the network, and by the specific questions that are addressed. Understanding the complexity of the diseases and its biological significance in health can be achieved by integrating data from the different functional genomics experiments with medical, physiological, and information on environmental factors and computed mathematically. The advantage of mathematical modeling of disease lies in the fact that such models not only shed insight as to how a complex process works, which could be very difficult to infer an understanding of each component of this process but also predict what may follow as time evolves or as the characteristics of particular system components are modified. Mathematical models have generally been utilized in association with an increased understanding of what models can offer in terms of prediction and insight.

Models have two distinct roles, prediction and understanding, related to the model properties of accuracy, transparency, and flexibility. Prediction of the models should be accurate including all the complexities and population-level heterogeneity having an additional use as a statistical tool. It also provides an understanding of how the disease spreads in the real world and how the complexity affects the dynamics. Understanding the model helps to develop sophisticated predictive models and gather more relevant epidemiological data. A model should be suited for its purpose that is, it should be as simple as possible, but no simpler—having an appropriate balance of accuracy, transparency, and flexibility (Keeling and Rohani, 2008). The model built should be helpful in understanding the behavior of the disease and able to simplify the other disease condition.

In silico models of cells

Several projects are proceeding along these lines such as E-CELL (Tomita, 2001), Virtual Cell (Resasco et al., 2012), and CellDesigner (Funahashi et al., 2003) and simulations of biochemical pathways (Palsson, 2000). Whole-cell modeling integrates information from metabolic pathways, gene regulation, and gene expression. In addition, several software applications model specific aspects of cellular interactions, such as COPASI (Hoops et al., 2006) for simulation and analysis of biochemical networks, PhysioDesigner 1.2 (Asai et al., 2012) for modeling physiological systems, etc. Three elements are needed for constructing a good cell model: precise knowledge of the phenomenon, an accurate mathematical representation, and a good simulation tool (Tomita, 2001).

A cell represents a dynamic environment of interaction among nucleic acids, proteins, carbohydrates, ions, pH, temperature, pressure, and electrical signals. Many cells with similar functionality form tissue. In addition, each type of tissue uses a subset of this cellular inventory to accomplish a particular function. For example, in neurons, electrochemical phenomena take precedence over cell division, in which cell division is a fundamental function of skin, lymphocytes, and bone marrow cells. Thus an ideal virtual cell not only represents all the information but also exhibits the potential to differentiate into neuronal or epithelial cell. The first step in creating a whole-cell model is to divide the entire network into pathways and pathways into individual reactions. Any two reactions belong to a pathway if they share a common intermediate. In silico modeling consists not only of decomposing events into manageable units but also of assembling these units into a unified framework. In other words, mathematical modeling is an art of converting biology into numbers.

For whole-cell modeling, a checklist of biological phenomena that call for mathematical representation is needed. Biological phenomena taken into account for in silico modeling of whole cells are the following:

- DNA replication and repair
- translation
- transcription and regulation of transcription
- energy metabolism
- cell division
- chromatin modeling
- signaling pathways
- membrane transport (ion channels, pump, nutrients)
- intracellular molecular trafficking
- cell membrane dynamics
- metabolic pathways.

The whole-cell metabolism includes enzymatic and nonenzymatic processes. Enzymatic processes cover most of the metabolic events, while nonenzymatic processes include gene expression and regulation, signal transduction, and diffusion.

In silico modeling of whole cells not only requires precise qualitative and quantitative data but also an appropriate mathematical representation of each event. For metabolic modeling, the data input consists of kinetics of individual reactions and also effects of cofactors, pH, and ions on the model. The key step in modeling is to choose an appropriate assumption. For example, a metabolic pathway may be a mix of forward and reverse reactions. Furthermore inhibitors that are part of the pathway may influence some reactions. At every step, enzymatic equations are needed that best describe the process. In silico models are built because they are easy to understand, controllable, and can store and analyze large amounts of information. A well-built model has diagnostic and predictive abilities. A cell by itself is a complete biochemical reactor that contains all the information one needs to understand life. Whole-cell modeling enables investigation of the cell cycle, physiology, spatial organization, and cell–cell communication. Sequential actions in whole-cell modeling are the following:

- catalog all the substances that make up a cell
- make a list of all the reactions, enzymes, and effectors
- map the entire cellular pathways: gene regulation, expression, metabolism, etc.
- build a stoichiometric matrix of all the reactions versus substances (for qualitative modeling)
- add rate constants, concentration of substances, strength of inhibition
- assume appropriate mathematical representations for individual reactions
- simulate reactions with suitable simulation software
- diagnose the system with system analysis software
- perturb the system and correlate its behavior to an underlying genetic and/or biochemical
- phenomenon using a hypothesis generator.

In silico metabolic modeling

Metabolic modeling is important to simulate the cell's function and phenotype. In recent years, metabolic modeling is frequently used in the field of drug discovery, clinical trial, and precision medicine. Computational and mathematical techniques are essential for metabolic modeling and several approaches and computational tools are currently available for this purpose. Essential features of metabolic modeling include metabolic systems biology, genome-scale metabolic models, constraint-based modeling and flux balance analysis, and multiomic flux balance analysis among others. Apart from these applications, metabolic modeling is also important for developing various disease models for tissue- and patient-specific insights into human diseases such as diabetes, neurodegenerative diseases, and cancers (Angione, 2019). Among the several approaches, machine and deep learning is applied for genome-scale metabolic modeling and Python-based models are available for cellular metabolic functional and phenotype or disease analysis (Cardoso et al., 2018; Zampieri et al., 2019). Several tools are nowadays available for metabolic modeling and analysis such as "Cameo" for in silico design of cell factories (Cardoso et al., 2018), "Escher-FBA" for flux balance analysis (Rowe et al., 2018), and "COBRAme" for genome-scale models of metabolism and gene expression (Lloyd et al., 2018). The area is evolving and several new approaches and tools are emerging.

In silico modeling of disease: in practice

In silico modeling has been applied in cancer, systemic inflammatory response syndrome (SIRS), immune disease, neuronal disease, and infectious disease. In silico models of disease can contribute to a better understanding of the pathophysiology of the disease, suggest new treatment strategies, and provide insight into the design of experimental and clinical trials for the investigation of new treatment modalities.

In silico modeling of disease combines the advantages of both in vivo and in vitro experimentation, without subjecting itself to the ethical considerations and lack of control associated with in vivo experiments. Unlike in vitro experiments, which exist in isolation, in silico models allow the researcher to include a virtually unlimited array of parameters, which render the results more applicable to the organism as a whole (Colquitt et al., 2011).

In silico modeling of disease is quite challenging. Attempting to incorporate every single known interaction rapidly leads to an unmanageable model. Further parameter determination in such models can be a frightening experience. Estimates come from diverse experiments, which may be elegantly designed and well executed but can still give rise to widely differing values for parameters. Data can come from both in vivo and in vitro experiments and results that hold in one medium may not always hold in the other. Further despite the many similarities between mammalian systems, significant differences do exist and so results obtained from experiments using animal and human tissue may not always be consistent.

In silico models of cancer

In silico modeling of cancer has become an interesting alternative approach to traditional cancer research. In silico models of cancer are expected to predict the complexity of cancer at multiple temporal and spatial resolutions, with the aim of supplementing diagnosis and treatment by helping plan more focused and effective therapy via surgical resection, standard chemotherapy, novel treatments. In silico models of cancer include (1) statistical models of cancer, such as molecular signatures of perturbed genes and molecular pathways, and statistically inferred reaction networks; (2) models that represent biochemical, metabolic, and signaling reaction networks important in oncogenesis, including constraint-based and dynamic approaches for the reconstruction of such networks; and (3) models of the tumor microenvironment and tissue-level interactions for microenvironment-tissue level (Edelman et al., 2010; Araujo and McElwain, 2004; Byrne et al., 2006; Bellomo et al., 2008; van Riel, 2006; Rejniak and Anderson, 2011).

Statistical models of cancer can be broadly divided into those that employ unbiased statistical inference and those that also incorporate a priori constraints of specific biological interactions from data. Statistical models of cancer biology at the genetic, chromosomal, transcriptomic, and pathway levels provide insight about molecular etiology and the consequences of malignant transformation, despite incomplete knowledge of underlying biological interactions. These models are able to identify molecular signatures that can inform diagnosis and treatment selection, for example, with molecularly targeted therapies such as Imatinib (Gleevec) (Edelman et al., 2010).

However in order to characterize specific biomolecular mechanisms that drive oncogenesis, genetic, and transcriptional activity must be considered in the context of cellular networks that ultimately drive cellular behavior. In microbial cells, network inference tools have been developed and applied for the modeling of diverse biochemical, signaling, and gene expression networks. However due to the much larger size of the human genome compared with microbes, and the substantially increased complexity of eukaryotic genetic regulation, the inference of transcriptional regulatory networks in cancer presents increased practical and theoretical challenges (Edelman et al., 2010).

Biochemical reaction networks are constructed to represent explicitly the mechanistic relationships between genes, proteins, and the chemical interconversion of metabolites within a biological system. In these models, network links are based on preestablished biomolecular interactions rather than statistical associations; significant experimental characterization is thus needed to reconstruct biochemical reaction networks in human cells. These biochemical reaction networks require, at a minimum, knowledge of the stoichiometry of the participating reactions. Additional information such as thermodynamics, enzyme capacity constraints, time-series concentration profiles, and kinetic rate constants can be incorporated to compose more detailed dynamic models (van Riel, 2006; Edelman et al., 2010).

Microenvironment-tissue level models of cancer apply an "engineering" approach that views tumor lesions as complex micro-structured materials, where three-dimensional tissue architecture ("morphology") and dynamics are coupled in complex ways to cell phenotype, which in turn is influenced by factors in the microenvironment. Computational approaches of in silico cancer research include continuum models, discrete models, and hybrid models.

In continuum models, extracellular parameters can be represented as continuously distributed variables to mathematically model cell−cell or cell−environment interactions in the context of cancers and the tumor microenvironment. Systems of partial differential equations have been used to simulate the magnitude of interaction between these factors. Continuum models are suitable for describing the individual cell migration, change of cancer cell density, the diffusion of chemoattractants, heat transfer in hyperthermia treatment for skin cancer, cell adhesion, and the molecular network of a cancer cell as an entire entity. However this type of in silico models has limited ability in investigating single-cell behavior and cell−cell interaction.

On the other hand, "discrete" models, that is, cellular automata models represent cancer cells as discrete entities of defined location and scale, interacting with one another and external factors in discrete time intervals according to predefined rules. Agent-based models expand the cellular automata paradigm to include entities of divergent functionalities interacting together in a single spatial representation, including different cell types, genetic elements, and environmental factors. Agent-based models have been used for modeling three-dimensional tumor cell patterning, immune system surveillance, angiogenesis, and the kinetics of cell motility.

Hybrid models have been created which incorporate both continuum and agent-based variables in a modular approach. Hybrid models are ideal for examining direct interactions between individual cells and between the cells and their microenvironment, but they also allow us to analyze the emergent properties of complex multicellular systems (such as cancer). Hybrid models are often multiscale by definition integrating processes on different temporal and spatial scales, such as gene expression, intracellular pathways, intercellular signaling, cell growth, or migration. There are two

general classes of hybrid models, those that are defined upon a lattice and those that are off lattice.

The classification of hybrid models on these two classes depends on the number of cells these models can handle and the included details of each individual cell structure, that is, models dealing with large-cell populations but with simplified cell geometry, and those that model small colonies of fully deformable cells. Interested readers can find details on hybrid models of cancer in the interesting recent review of Rejniak and Anderson (2011).

For example, a hybrid model investigated the invasion of healthy tissue by a solid tumor. The model focused on four key parameters implicated in the invasion process: tumor cells, host tissue (extracellular matrix), matrix-degradative enzymes, and oxygen. The model was considered to be hybrid since the latter three variables were continuous (i.e., concentrations) and the tumor cells were discrete (i.e., individuals). This hybrid model aimed to investigate how individual-based cell interactions (with one another and the matrix) can affect the tumor shape (Anderson, 2005). Another model incorporated a continuous model of a receptor signaling pathway, an intracellular transcriptional regulatory network, cell-cycle kinetics, and three-dimensional cell migration in an integrated, agent-based simulation of solid brain tumor development (Zhang et al., 2007). The interactions between cellular and microenvironment states have also been considered in a multiscale model that predicts tumor morphology and phenotypic evolution in response to such extracellular pressures (Anderson et al., 2006).

In silico models of tumor microenvironment integrate information about the biological context in which cancers develop. Multiple factors involved in the development of an intrinsically complex tumor microenvironment have been studied including extracellular biomolecules, vasculature, and the immune system. However rarely these methods have been integrated with a large cell–cell communication network in a complex tumor microenvironment. Recently, an interesting effort of in silico modeling was described, in which the investigators integrated all the intercellular signaling pathways known to date for human glioma and generated a dynamic cell–cell communication network associated with the glioma microenvironment. Then, they applied evolutionary population dynamics and the Hill functions to interrogate this intercellular signaling network and execute an in silico tumor microenvironment development. The observed results revealed a profound influence of the microenvironmental factors on tumor initiation and growth and suggested new options for glioma treatment by targeting cells or soluble mediators in the tumor microenvironment (Wu et al., 2012).

In silico models and inflammatory response syndrome in trauma and infection

Trauma and infection elicit an acute inflammatory response. In certain circumstances, the degree of the acute inflammatory response may result in pathologic manifestations, namely, SIRS, sepsis, and multiple organ failure. Further research is needed for the appropriate management of these states. Despite longstanding efforts, there has been uniform difficulty in translating the results of basic science research into effective therapeutic regimes. It has been suggested that this difficulty is due in part to a failure to account for the complex, nonlinear nature of the inflammatory process of which SIRS/MOF represents a disordered state. In silico modeling seems to be a promising research approach in this field. Indeed, in silico modeling of inflammation has been applied in an effort to bridge the gap between basic science and clinical trials. Specifically, both agent-based modeling and equation-based modeling have been utilized (Vodovotz et al., 2008; Geris et al., 2010). Equation-based modeling encompasses primarily ordinary differential equations (ODEs) and partial differential equations (PDEs). Initial modeling studies were focused on the pathophysiology of the acute inflammatory response to stress, and these studies suggested common underlying processes generated in response to infection, injury, and shock. Later, mathematical models included the recovery phase of injury and gave insight into the link between the initial inflammatory response and the subsequent healing process. The first mathematical model of wound healing dates back to the 1980s and early 1990s. These models and others developed in the 1990s investigated epidermal healing, repair of the dermal extracellular matrix, wound contraction, and wound angiogenesis. Most of these models were deterministic and formulated using differential equations. In addition, recent models have been formulated using differential equations to analyze different strategies for improved healing, including wound VACs, commercially engineered skin substitutes, and hyperbaric oxygen. In addition, agent-based models have been used in wound healing research. For example, Mi et al. (2007) have developed an agent-based model to analyze different treatment strategies with wound debridement and topical administration of growth factors. Their model produced the expected results of healing when analyzing for different treatment strategies including debridement, release of PDGF, reduction in tumor necrosis factor-α, and increase of TGF-β1. The investigators suggested that a drug company should use a mathematical model to test a new drug before going through the expensive process of basic science testing, toxicology, and clinical

trials (Mi et al., 2007). Indeed, clinical trial design can be improved by prior in silico modeling. For example, in silico modeling has led to the knowledge that patients who suffered from the immune-suppressed phenotype of late-stage multiple organ failure and were susceptible to usually trivial nosocomial infections demonstrated sustained elevated markers of tissue damage and inflammation through 2 weeks of simulated time. However anticytokine drug trials with treatment protocols of only one dose or 1 day had not incorporated this knowledge into their design, with subsequent failure of candidate treatments.

Applications of in silico disease modeling

The application of mathematical modeling has been successfully deployed in representing several pathophysiological disorders and is known as disease modeling. Several dynamics of pathophysiological and cell-signaling pathways have been studied using disease modeling. Disease modeling has also been used to study the spread of diseases and epidemics (Dezső and Barabási, 2002; Wang et al., 2016).

Infectious diseases

By now the reader is expected to be familiar with the meaning and the basics of in silico model (ing). In this section, we discuss the application of in silico modeling in the understanding of infectious diseases and in the proposition/development of better treatments for infectious diseases. In fact, it is worthy of note that the applications of in silico modeling can help far beyond just the understanding of the dynamics (and sometimes, statics) of infectious diseases, and far beyond the proposition/development of better treatments for infectious diseases. In silico modeling can be helpful even in the understanding of better prevention of infectious diseases.

The process of infection is defined as the level of a pathogen within the host which in turn is determined by the growth rate of the pathogen and its interaction with the host's immune response. Initially, no pathogen is present but just a low-level nonspecific immunity within the host. On infection, the parasite grows abundantly over time with the potential to transmit the infection to other susceptible individuals.

Triad of infectious diseases as the source of parameters for in silico modeling of infectious diseases

To comprehensively understand in silico modeling in the domains of infectious diseases, one should first understand the "triad of infectious diseases," and the characteristics of "infectious agent," "host," and "environment" on which the models and always based. In fact, modeling of infectious diseases is just impossible without this triad, after all, the model would be built on some parameters (also called variables, in a more general language), and those parameters always have their origin from the so-called "triad of infectious diseases." At this point, a good question would be: What is "triad of infectious diseases?"

By "triad of infectious diseases," we mean the interactions between (1) agent, which is the disease-causing organism—the pathogen; (2) host, which is the infected organism, or in the case of preinfection, the organism to be infected is the host—thus in this case host is the animal the agent infects; and (3) environment, which is a kind of the link between the agent and the host—essentially, the environment is an umbrella word for the entirety of the possible media through which the agent reaches the host (Park, 2009).

Now that we have set the direction, and we (now) know what in silico modeling of infectious diseases are fundamentally based on, let us proceed gradually and get a better understanding of the parameters on which most in silico infectious disease models are based. To discuss the parameters in an orderly manner, we just categorize them under each of the three components of the "triad of infectious diseases" and summarize them in the next subsection. It must be emphasized at this point that (1) even though all the possible parameters for in silico modeling of infectious diseases can be successfully categorized under the characteristics of one of any of the three components of the "triad of infectious diseases"—agent, host, and environment, (2) the parameters discussed in the next subsection are by no mean the entirety of all the possible parameters that can be included in in silico modeling of infectious diseases—in fact, several parameters exist that this section cannot possibly enumerate them all, and that is why we have discussed the parameters using categorical approach.

Parameters for in silico modeling of infectious diseases

Parameters derived from characteristics of agent

Some of the parameters for in silico modeling of infectious diseases are essentially a measure of infectivity (ability to enter the host), pathogenicity (ability to cause divergence from homeostasis/disease), virulence (degree of divergence from homeostasis caused/ability to cause death), antigenicity (ability to bind to mediators of host's adaptive immune system), and immunogenicity (ability to trigger adaptive immune response) (Scott and De Groot, 2010) of the concerned infectious agent. The exact measure (and thus the

units) used can vary markedly depending on the intentions for which the in silico infectious diseases model is built, as well as the assumptions on which the in silico disease model is based. From the knowledge of agent's characteristics, one should know that unlike parameters related to the other characteristics of agent, the parameters related to infectivity find their most important use only in the modeling of preinfection stage in infectious disease modeling.

Finally, some of the agent-related parameters of great importance in in silico modeling of infectious diseases are the concentration of agent's antigen–host's antibody complex, case fatality rate, the strain of the agent, other genetic information of the agent, etc.

Parameters derived from characteristics of host

The parameters originating from characteristics of host can also be so diversified and be based on the intentions for which the in silico infectious diseases model is built and the assumptions on which the in silico disease model is based, but then the parameters could be grouped and explained under host's genotype (the allele at host's specified genetic locus), immunity/health status (biological defenses to avoid infection), nutritional status (feeding habits/food intake characteristics), gender (often categorized as male, or female), age, and behavior (host's behaviors that affect its resistance to homeostasis disruptors).

Typical examples of host-related parameters are the alleles at some specifically targeted genetic loci; the total white blood cell counts; differential white blood cell counts, and/or much more sophisticated counts of specific blood cell types; blood levels of some specific cytokines, hormone, and/or neurotransmitters; daily calories, protein, and/or fat intake; daily amount of energy expended and/or duration of exercise; etc.

Parameters derived from characteristics of environment

At first, parameters originating from the environment might seem irrelevant to the in silico modeling of infectious diseases, but they are. Even after the preinfection stage, the environment still modulates the host–agent intersections. For example, the ability (and thus the related parameters) of the agent to multiply and/or harm the host are continually influenced by the host's environmental conditions, and in a similar way, the host's defense against the adverse effects of the agents are modulated by host's environmental conditions. But somehow, not so many of these parameters have been included in in silico infectious disease models in the recent past. A few examples of these parameters are host's ambient temperature, host's ambient atmospheric humidity, altitude, host's light–dark cycle, etc.

Infectious diseases in silico model proper, a typical approach/scenario

Now that we know the parameters for in silico infectious disease modeling, the next reasonable question would be "What form does a typical in silico infectious disease model take?" So, this subsection attempts to answer this very important question.

To answer this question, we start by employing the reader to view in silico model as a system of well-integrated functional equations/formulae. Then, these integrated functional equations/formulae should be seen as been so highly integrated that, ultimately, they could altogether be seen as a single gigantic functional equation/formula. From this big single functional equation, it is then possible (at least through the aid of a computer) to make any of the contained variables the subject of the equation depending on what one wants to obtain from the model. This is essentially an in silico model, and it is basically characterized by the possibility of computationally working with any (and sometimes, close to infinite) number of possible data points within the reasonable limits one set.

So the equations behind a typical infectious disease in silico model could take the form

$$H = \beta(link\,function)\, f(A)(link\,function)\, g(E) + \varepsilon\ldots \quad (22.1)$$

where H is the output from a smaller equation that is based on host parameters, β is a constant, f and g are link functions which may be same as or different from each other and other link functions in this system of equations, A is the output from a smaller equation that is based on agent parameters, g is a link function which may be same or different from other link functions in this system of equations, E is the output from a smaller equation that is based on environment parameters, and $\mathcal{E}$ is a random error parameter.

Readers should know that we use the term "link function" to refer to any of the various possible forms of mathematical operations or functions. Which means that based on the complexity of the model, a particular "link function" might be as simple as a mere addition or as complex as several combinations of an operator with high-degree polynomials.

H in Eq. (22.1) could have resulted from a smaller model/function of the form

$$H = \beta_h\ (\text{link function})\ f_{h1}(h_1)\ (\text{link function})\ f_{h2}(h_2)\ (\text{link function})\ldots f_{hx}(h_x)\ (\text{link function}) + \varepsilon\ldots \quad (22.2)$$

where β_h is a constant; f_{h1}, f_{h2}, … f_{hx} are link functions that may be the same or different (individually) from (every) other link functions in this system of equations; h_1, h_2, … h_x are a set of host's parameters (e.g., age, gender, white blood cell count, cytokine level, etc.); and $\mathcal{E}$ is a random error parameter.

A in Eq. (22.1) could have resulted from a smaller model/function of the form

$$A = \beta_a(\text{link function})\, f_{a1}(a_1)\,(\text{link function})\, f_{a2}(a_2)\,(\text{link function}) \ldots f_{ax}(a_x)\,(\text{link function}) + \varepsilon \ldots \quad (22.3)$$

where β_a is a constant; f_{a1}, f_{a2}, … f_{ax} are link functions which may be the same or different (individually) from (every) other link functions in this system of equations; a_1, a_2, … a_x are a set of agent's parameters (e.g., case fatality rate, agent's genotype, etc.); and $\mathcal{E}$ is a random error parameter.

In a similar way, E in Eq. (22.1) could have resulted from a smaller model/function of the form

$$E = \beta_e\,(\text{link function})\, f_{e1}(e_1)\,(\text{link function})\, f_{e2}(e_2)\,(\text{link function}) \ldots f_{ex}(e_x)\,(\text{link function}) + \varepsilon \ldots \quad (22.4)$$

where β_e is a constant; f_{e1}, f_{e2}, … f_{ex} are link functions which may be the same or different (individually) from (every) other link functions in this system of equations; e_1, e_2, … e_x are a set of environmental parameters (e.g., host's ambient temperature, host's ambient atmospheric humidity, etc.); and $\mathcal{E}$ is a random error parameter.

Specific examples of infectious diseases in silico model

Muñoz-Elías et al. (2005) have documented (through their paper Replication dynamics of *Mycobacterium tuberculosis* in chronically infected mice) a successful in silico modeling of infectious diseases (specifically, tuberculosis). In their in silico modeling of tuberculosis in mice, the researchers investigated both the static and dynamic of host–pathogen/agent equilibrium (i.e., mice-mycobacterium tuberculosis static and dynamic equilibrium). The rationale behind their study was that a better understanding of host–pathogen/agent interactions would make possible the development of better antimicrobial drugs for the treatment of tuberculosis (as well as provide similar a understanding for the cases of other chronic infectious diseases). They modeled different types of host–pathogen/agent equilibriums (ranging from completely static equilibrium, all the way through semidynamic, down to completely dynamic scenarios) by varying the rate of multiplication/growth and the rate of death of the pathogen/agent (*M. tuberculosis*) during the infection's chronic phase. Through their in silico study (which was also verified experimentally), Muñoz-Elías et al. (2005) documented a number of remarkable findings. For example, they established that "viable bacterial counts and total bacterial counts in the lungs of chronically infected mice do not diverge over time" and explained that "rapid degradation of dead bacteria is unlikely to account for the stability of total counts in the lungs over time, because treatment of mice with isoniazid for 8 weeks led to a marked reduction in the viable counts without reducing the total count."

Readers who are interested in the further details on the generation of this in silico model for the dynamics of *M. tuberculosis* infection, as well as the complete details of the parameters/variables considered, and the comprehensive findings of the study should refer to the article of Muñoz-Elías et al. published in infection and immunity (Muñoz-Elías et al., 2005).

Another one of the many other notable works in the domain of infectious diseases in silico modeling is the study by Navratil et al. (2011). Using PPIs data that authors obtained from available literature and public databases, the authors (after first curating and validating the data) computationally (in silico) re-examined virus–human protein interactome. Interestingly, the authors were able to show that onset and the pathogenesis of some disease conditions (most especially, chronic disease conditions) that are often believed to be of genetic, lifestyle, or environmental origin, are in fact modulated by infectious agents.

The reader may also be interested in a few other studies or reviews such as Chavali et al. (2008) and Watterson and Ghazal (2010) which involve applications of in silico techniques in the better understanding of infectious agents. But we warn that these two articles do not deal completely/exactly with in silico modeling of infectious disease(s).

Model of bacterial and viral dynamics

Models have been constructed to simulate bacterial dynamics, such as growth under various nutritional and chemical conditions (Brookmeyer et al., 2005), chemotactic response (Andrews et al., 2006), and interaction with host immunity (Ben-David et al., 2005). Clinically important models of bacterial dynamics relating to peritoneal dialysis (Hotchkiss et al., 2004), pulmonary infections (Henson, 2003), and particularly of antibiotic treatment and bacterial resistance (Hupert et al., 2002) have also been developed.

Baccam et al. (2006) utilized a series of mathematical models of increasing complexity, which incorporate target cell limitation and the innate interferon response, the model is applied to examine influenza A virus kinetics in the upper respiratory tracts of experimentally infected adults showing the models to be applicable to improve the understanding of influenza A virus infection and estimated that during an upper respiratory tract infection, influenza virus initially spreads rapidly with one cell, on average, infecting ~20 others (Daun and Clermont, 2007).

Model parameter and spread of disease: model parameters are one of the main challenges in mathematical modeling since all models do not have a physiological meaning. Sensitivity analysis and bifurcation analysis give us the opportunity to understand how model outcome and model parameters are correlated, how the sensitivity of the system is with respect to certain parameters and the uncertainty in the model outcome yielded by the uncertainties in the parameter values (Bernard et al., 2004). Uncertainty and sensitivity analysis was used to evaluate the input parameters play on the basic productive rate (*Ro*) of the severe acute respiratory syndrome (Chowell et al., 2006) and tuberculosis (Sanchez and Blower, 1997). The control of the outbreak depends on identifying the disease parameters that are likely to lead to a reduction in *R*.

Challenges in in silico modeling of infectious diseases

Difficulty in finding the most appropriate set of parameters for the in silico modeling of infectious diseases is often a challenge. But this challenge hopes to subside with the advancement in infectonomics and high-throughput technology. However another important challenge lies in the understanding (and the provision of reasonable interpretations for) the results from all the complex interaction of parameters considered.

Neuronal diseases

In this subsection, we focus on the application of in silico modeling to improve knowledge of neuronal diseases and thus improve the applications of neurological knowledge in the solving of neuronal health problems. It is not an overstatement to say that one of the many aspects of life sciences where in silico disease modeling would have the biggest applications is in the better understanding of the pathophysiology of nervous system (neuronal) diseases. This is basically because of the inherently delicate nature of the nervous system and the usual extra need to be much sure of how to proceed prior to attempting to treat neuronal disease conditions. By these, we mean that the need to first model neuronal disease conditions in silico prior to deciding on or suggesting (for example) a treatment plan is, in fact, rising. This is not unexpected, after all, it is better to much sure of what would work (say through in silico modeling) than to try what would not work.

Pathophysiology of neuronal diseases as the source of parameters for in silico modeling of neuronal diseases

Obtaining appropriate parameters for the in silico modeling of a nervous system (neuronal) disease is rooted in a good understanding of the pathophysiology of such neuronal disease. But since comprehensive details of pathophysiology of neuronal diseases are beyond the scope of this book, we only present the basic idea that would allow the reader to understand how in silico modeling of a nervous system (neuronal) disease can be done. We encourage readers who are interested in more details of the pathophysiology of neuronal disease to proceed to available textbooks and articles (de la Torre, 2008).

To give a generalized explanation and still concisely present the basic ideas underlying the pathophysiology of neuronal diseases, we proceed by systematically categorizing the mediators of neuronal diseases pathophysiology under (1) nervous cell characteristics, (2) signaling chemicals and body electrolytes, (3) host/organism factors, and (4) environmental factors. Readers need to see all these categories as been highly integrated pathophysiologically rather than see them as been spate entities—and that we have only grouped them this way to make simple the explanation of how the parameters for in silico modeling of neuronal diseases are generated.

When something goes wrong with (or there is a marked deviation from equilibrium in) a component of any of the four categories above, the other components (within and/or outside the same category) try hard to make adjustments so as to annul/compensate for the undesired change. For example, if the secretion of a chemical signal suddenly becomes abnormally low, the target cells for the chemical signal may develop mechanisms to use the signaling chemical more efficiently and the degradation rate of the signaling chemical may be reduced considerably. Through these, the potentially detrimental effects of reduced secretion of the chemical signal are annulled, through the compensation from the other components. This is just a very simple example—much complex regulatory and homeostasis mechanisms exist in the neuronal system. But despite the robustness of those mechanisms, things still go out of hand sometimes, and disease conditions result. The exploration of what happens in/to each and all of the components of this giant system in disease conditions is called the pathophysiology of the neuronal disease, and it this pathophysiology that "provides" parameters for the in silico modeling of neuronal diseases.

Parameters for in silico modeling of neuronal diseases

Parameters derived from characteristics of nervous cell

Some of the important parameters (that are of nervous cell characteristics origin) for a typical in silico

modeling of a neuronal disease [e.g., Alzheimer's disease (AD)] are the population (or relative population) of specific neuronal cells (such as glial cells—microglia, astrocytes, etc.), motion of specific neuronal cells (e.g., microglia), amyloid production, aggregation, and removal of amyloid (Edelstein-Keshet and Spiros, 2002), morphology of specific neuronal cells, status of neuronal cell receptors, generation/regeneration/degeneration rate of neuronal cells, status of ion neuronal cells' channels, etc. Based on their relevance to the pathophysiology of the neuronal disease being studied, many of these parameters are often considered in the in silico modeling of the neuronal disease. And more importantly, their spatiotemporal dynamics are often seriously considered.

Parameters derived from characteristics of signaling chemicals and body electrolytes

The importance of signaling chemicals and electrolytes in the nervous system makes parameters related to them very important. The secretion, uptake, degradation (Edelstein-Keshet and Spiros, 2002), and diffusion rates, of various neurotransmitters and cytokines are often very important parameters in the in silico modeling of neurodiseases. Other very important parameters are the concentration gradients of the various neurotransmitters and cytokines, the availability and concentration of second messengers, and the cells'/system's electrolyte status/balance. The spatiotemporal dynamics of all of these are also often seriously considered.

Parameters derived from host/organism factors

The parameters under host/organism factors can be highly varied depending on the intentions and the assumptions governing the in silico disease modeling. Nonetheless, one could basically group and list the parameters collectively under genotype (based on allele at specified genetic locus), nutritional status (feeding habits/food intake characteristics, e.g., daily calories, protein, etc., intake), gender (male or female), age, and behavior (host's behaviors/lifestyle that influences homeostasis and/or responses to stimuli).

Parameters derived from environmental factors

A few examples of these parameters are ambient temperature, altitude, light–dark cycle, social network, type of influences from people in network, etc.

Neuronal disease in silico model proper, a typical approach/scenario

Just like other in silico models, the neuronal disease in silico model is also based on what could be viewed as a single giant functional equation, which is composed of highly integrated simpler functional equations.

So, the equations behind a typical neuronal disease in silico model could take the form

$$N = \beta\ (\text{link function})\ f(C)\ (\text{link function})\ g(S)\ (\text{link function})\ j(H)\ (\text{link function})\ k(E) + \varepsilon \ldots \tag{22.5}$$

where N could be a parameter that is directly a measure of the disease manifestation; β is a constant; f, g, j, and k are link functions which may be the same or different from other link functions in this system of equations; C, S, H, and E are the outputs from smaller equations that are based on parameters from neuronal cell characteristics, signaling molecules and electrolyte parameters, host parameters, and environment parameters, respectively; and $\mathcal{E}$ is a random error parameter.

Reader should know that each of N, C, S, H, and E could have resulted from smaller equations that could take forms similar to those (Eqs. (22.2)–(22.4)) described under in silico modeling of infectious diseases (previous subsection).

Specific examples of neuronal disease in silico model

In their work, Edelstein-Keshet and Spiros (2002) used in silico modeling to study the mechanism and/formation of AD. The target of their in silico modeling was to explore and demystify how various parts implicated in the etiology and pathophysiology of AD work together as a whole. Employing the strength of in silico modeling, the researchers were able to transcend the difficulty of identifying detailed disease progression scenarios, and they were able to test a wide variety of hypothetical mechanisms, at various levels of detail.

Readers who may be interested in the complete details of the assumptions that govern in silico modeling of AD, the various other aspects of the model, and more detailed account of the findings would love to look at the article of Edelstein-Keshet and Spiros.

Several other interesting studies have applied in silico modeling techniques to investigate various neuronal diseases. A few examples include the work of Altmann and Boyton (2004) who investigated multiple sclerosis (a very common disease resulting from demyelination in the central nervous system) using in silico modeling techniques; Lewis et al. (2010) who used in silico modeling to study the metabolic interactions between multiple cell types in AD condition; and Raichura et al. (2006) who applied in silico modeling techniques to dynamically model alpha-synuclein processing (in normal and) in Parkinson's disease state(s).

A more specific example of a molecular level in silico AD model can be found in Ghosh et al. (2010, 2012). Among the amyloid proteins, amyloid-β (Aβ) peptides

(Aβ42 and Aβ40) are known to form aggregates that deposit as senile plaques in the brains of AD patients. The process of Aβ aggregation is strongly nucleation-dependent and is inferred by the occurrence of a "lag-phase" prior to fibril growth that shows a sigmoidal pattern. Ghosh et al. (2010) dissected the growth curve into three biophysically distinct sections to simplify the modeling and to allow the data to be experimentally verifiable. Stage I is where the prenucleation events occur whose mechanism is largely unknown. The prenucleation stage is extremely important in dictating the overall aggregation process where critical events such as conformation change and concomitant aggregation take place, and it is also the most experimentally challenging to decipher. In addition to mechanistic reasons, this stage is also physiologically important as low-molecular weight species are implicated in AD pathology. The rate-limiting step of nucleation is followed by growth (stage II, Fig. 22.1). The overall growth kinetics and structure and shape of the fibrils are mainly determined by the structure of nucleating species. An important intermediate along the aggregation pathway, called "protofibrils" have been isolated and characterized that have propensities to both elongate (by monomer addition) as well as to laterally associate (protofibril−protofibril association) to grow into mature fibrils (stage III in the growth curve).

Simulation of the fibril growth process in Aβ42 aggregation. Ghosh et al. (2010) generated an ODE-based molecular simulation (using mass-kinetics methodology) of this fibril growth process to estimate the rate constants involved in the entire pathway. The dynamics involved in the protofibril elongation stage of the aggregation (stage III of the process) were estimated and validated by in vitro biophysical analysis.

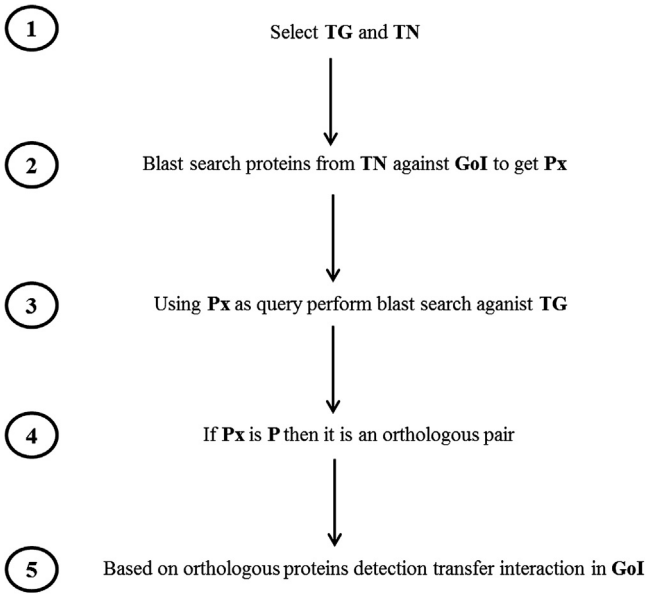

FIGURE 22.1 Flow Chart showing use of bio-informatics in disease modeling.

Preliminary identification of nucleation mass: Ghosh et al. (2012) next used the rate constants identified from stage III to create a complete aggregation pathway simulation (combining stages I, II, and III) to approximately identify the nucleation mass involved in Aβ aggregation.

In order to model the Aβ system, one needs to estimate the rate constants involved in the complete pathway and the nucleation mass itself. It is difficult to iterate through different values for each of these variables to get close to the experimental plots (fibril growth curves measured via fluorescence measurements with time) due to the large solution space and finding the nucleation phase cannot be done independently without estimating the rate constants alongside. However having separately estimated the postnucleation stage rate constants as mentioned above reduces the overall parameter estimation complexity.

The complete pathway simulation in Ghosh et al. (2012) was used to study the lag times associated with the aggregation pathway and hence predict possible estimates of the nucleation mass. The following strategy was used: estimate the prenucleation rate constants that give the *maximum lag times* for each possible estimate of the nucleation mass. This led to four distinctly different regimes of possible nucleation masses corresponding to four different pairs of rate constants for the prenucleation phase (regime 1: $n = 7,8,9,10,11$; regime 2: $n = 12,13,14$; regime 3: $n = 15,16,17$; regime 4: $n = 18,19,20,21$). However it was experimentally observed that the semilog plot of the lag times against the initial concentration of Aβ is linear and this characteristic was used to figure out what values of nucleation mass are most feasible for the Aβ42 aggregation pathway. The simulated plots show a more stable relationship between the lag times and the initial concentrations, and the best predictions for the nucleation mass was reported to be in the range 10,11,. . .,16.

Such molecular pathway level studies are extremely useful in understanding the pathogenesis of AD in general and can motivate drug development exercises in the future. For example, the characterization of the nucleation mass is important as it has been observed that various fatty acid interfaces can arrest the fibril growth process (by stopping the reactions beyond the prenucleation stage). Such in-depth modeling of the aggregation pathway can suggest what concentrations of fatty acid interfaces should be used (under a given Aβ concentration in the brain) to arrest the fibril formation process leading to direct drug dosage and interval prediction for AD patients.

Possible limitations of in silico modeling of neuronal diseases

Despite that we have mentioned several possible parameters for in silico modeling of neurodiseases, it is worthy of note that finding a set of the most reasonable set of parameters for the modeling is, in fact, a big challenge. On the other hand, understanding (and thus finding reasonable biological interpretations for) the results from the complex interaction of all parameters considered is also a big challenge. In addition, a number of assumptions that models are sometimes based on still have controversial issues. Accurately modeling the spatiotemporal dynamics of neurons and neurotransmitters transmitters (and other chemicals/ligands) also constitutes a huge challenge (Edelstein-Keshet and Spiros, 2002).

Conclusion

Understanding the complex systems involved in a disease will make it possible to develop smarter therapeutic strategies. Treatments for existing tumors will use multiple drugs to target the pathways or perturbed networks that show an altered state of activity. In addition, models can effectively form the basis for translational research and personalized medicine.

Biological function arises as the result of processes interacting across a range of spatiotemporal scales. The ultimate goal of the applications of bioinformatics in systems biology is to aid in the development of individualized therapy protocols to minimize patient suffering while maximizing treatment effectiveness. It is now being increasingly recognized that multiscale mathematical and computational tools are necessary if we are going to be able to fully understand these complex interactions, for example, in cancer (Anderson and Quaranta, 2008) and heart diseases (Liang et al., 2009).

With these bioinformatics tools, computational theories, and mathematical models introduced in this article, readers should be able to dive into the exhilarating area of formal computational systems biology. Investigating these models and confirming their findings by experimental and clinical observations is a way to bring together molecular reductionist with quantitative holistic approaches and create an integrative mathematical view of disease progression. We hope to have shown that there are many interesting challenges yet to be solved and that a structured principled approach will be essential for tackling them.

Systems biology is an emerging field that aims in understanding the biological system at the system level with a high aspect of mathematical and statistical modeling methods. In silico modeling of infectious disease is a rich and growing field focused on modeling spread and containment infection with designs being a flexible enabling adaptation to new data types.

References

Altmann, D.M., Boyton, R.J., 2004. Models of multiple sclerosis. Drug. Discov. Today Dis. Model. 1 (4), 405–410.

Anderson, A.R., 2005. A hybrid mathematical model of solid tumour invasion: the importance of cell adhesion. Math. Med. Biol. 22, 163–186.

Anderson, A.R., Quaranta, V., 2008. Integrative mathematical oncology. Nat. Rev. Cancer 8 (3), 227–234.

Anderson, A.R., Weaver, A.M., Cummings, P.T., Quaranta, V., 2006. Tumor morphology and phenotypic evolution driven by selective pressure from the microenvironment. Cell 127, 905–915.

Andrews, B.W., Yi, T.M., Iglesias, P.A., 2006. Optimal noise filtering in the chemotactic response of Escherichia coli. PLoS. Comput. Biol. 2, e154.

Angione, C., 2019. Human systems biology and metabolic modelling: a review-from disease metabolism to precision medicine. Biomed. Res. Int. 2019, 8304260. Available from: https://doi.org/10.1155/2019/8304260.eCollection2019.

Araujo, R.P., McElwain, D.L.S., 2004. A history of the study of solid tumour growth: the contribution of mathematical modelling. Bull. Math. Biol. 66, 1039–1109.

Asai, Y., et al., 2012. Multilevel modeling of physiological systems and simulation platform: physiodesigner, flint and flint K3 service. Applications and the Internet (SAINT). In: 2012 IEEE/IPSJ 12th International Symposium on. IEEE.

Baccam, P., Beauchemin, C., Macken, C.A., Hayden, F.G., Perelson, A.S., 2006. Kinetics of influenza A virus infection in humans. J. Virol. 80, 7590–7599.

Bellomo, N., Li, N.K., Maini, P.K., 2008. On the foundation of cancer modelling: selected topics, speculations, and perspectives. Math. Model. Methods Appl. Sci. 18, 593–646.

Ben-David, I., Price, S.E., Bortz, D.M., Greineder, C.F., Cohen, S.E., Bauer, A.L., et al., 2005. Dynamics of intrapulmonary bacterial growth in a murine model of repeated microaspiration. Am. J. Respir. Cell Mol. Biol. 33, 476–482.

Benson, D.A., et al., 2013. GenBank. Nucleic Acids Res. 41. D1 D36–D42.

Berman, H.M., et al., 2000. The protein data bank. Nucleic Acids Res. 28 (1), 235–242.

Bernard, S., Belair, J., Mackey, M.C., 2004. Bifurcations in a white-blood-cell production model. C. R. Biol. 327, 201–210.

Breitling, R., Gilbert, D., et al., 2008. A structured approach for the engineering of biochemical network models, illustrated for signalling pathways. Brief. Bioinform. 9 (5), 404–421.

Brook, B.S., Kohl, P., et al., 2011. Towards the virtual physiological human: mathematical and computational case studies. Philos. Trans. A Math. Phys. Eng. Sci. 369 (1954), 4145–4148.

Brookmeyer, R., Johnson, E., Barry, S., 2005. Modelling the incubation period of anthrax. Stat. Med. 24, 531–542.

Byrne, H.M., Alarcon, T., Owen, M.R., Webb, S.D., Maini, P.K., 2006. Modelling aspects of cancer dynamics: a review. Phil. Trans. R. Soc. A. 364, 1563–1578.

Cardoso, J.G.R., Jensen, K., Lieven, C., Lærke Hansen, A.S., Galkina, S., Beber, M., et al., 2018. Cameo: a python library for computer aided metabolic engineering and optimization of cell factories. ACS Synth. Biol. 7 (4), 1163–1166. Available from: https://doi.org/10.1021/acssynbio.7b00423. Epub 2018 Apr 4.

Chassagnole, C., Jackson, R.C., et al., 2006. Using a mammalian cell cycle simulation to interpret differential kinase inhibition in anti-tumour pharmaceutical development. Biosystems 83 (2–3), 91–97.
Chavali, A.V., Whittemore, J.D., Eddy, J.A., Williams, K.T., Papin, J. A., 2008. Systems analysis of metabolism in the pathogenic trypanosomatid Leishmania major. Mol. Syst. Biol. 4, 177.
Choi, S., 2007. Introduction to Systems Biology. Humana Press Inc.
Chowell, G., Castillo-Chavez, C., Fenimore, P.W., Kribs-Zaleta, C.M., Arriola, L., Hyman, J.M., 2006. Model parameters and outbreak control for SARS. Emrging Infect. Dis. 10, 1258–1263.
Colquitt, R.B., Colquhoun, D.A., Thiele, R.H., 2011. In silico modelling of physiologic systems. Best. Pract. Res. Clin. Anaesthesiol. 25, 499–510.
Coveney, P., Fowler, P.H., 2005. Modelling biological complexity: a physical scientist's perspective. J. R. Soc. Interface 2, 267–280.
Daun, S., Clermont, G., 2007. In silico modelling in infectious disease. Drug. Discov. Today Dis. Model. 4 (3), 117–122.
de la Torre, J.C., 2008. Pathophysiology of neuronal energy crisis in Alzheimer's disease. Neurodegenerative Dis. 5, 126–132.
de Jong, H., 2002. Modeling and simulation of genetic regulatory systems: a literature review. J. Comput. Biol. 9 (1), 67–103.
Dezső, Z., Barabási, A.-L., 2002. Halting viruses in scale-free networks. Phys. Rev. E 65 (5), 055103.
Edelman, L.B., Eddy, J.A., Price, N.D., 2010. In silico models of cancer. Wiley Interdiscip. Rev. Syst. Biol. Med. 4, 438–459.
Edelstein-Keshet, L., Spiros, A., 2002. Exploring the formation of Alzheimer's disease senile plaques in silico. J. Theor. Biol. 216, 301–326.
Edgar, R., Domrachev, M., Lash, A.E., 2002. Gene expression omnibus: NCBI gene expression and hybridization array data repository. Nucleic acids Res. 30 (1), 207–210.
Enderling, H., Anderson, A.R.A., et al., 2006. Mathematical modelling of radiotherapy strategies for early breast cancer. J. Theor. Biol. 241 (1), 158–171.
Fadiel, A., Anidi, I., et al., 2005. Farm animal genomics and informatics: an update. Nucleic Acids Res. 33 (19), 6308–6318.
CellDesigner: a process diagram editor for gene-regulatory and biochemical networks. In: Funahashi, A., Tanimura, N., Morohashi, M., Kitano, H. (Eds.), BIOSILICO, 1.
Geris, L., Schugart, R., Van Oosterwyck, H., 2010. In silico design of treatment strategies in wound healing and bone fracture healing. Philos. Trans. A Math. Phys. Eng. Sci. 368, 2683–2706.
Ghosh, P., Kumar, A., Datta, B., Rangachari, V., 2010. Dynamics of protofibril elongation and association involved in Abeta42 peptide aggregation in Alzheimer's disease. BMC Bioinforma 11 (Suppl. 6), S24.
Ghosh, P., Datta, B., Rangachari, V., 2012. Computational predictions for the lag-times and nucleation mass involved in Aβ42 peptide aggregation. BIOSTEC Bioinform. 1–4.
Hecker, M., Lambeck, S., et al., 2009. Gene regulatory network inference: data integration in dynamic models—a review. Biosystems 96 (1), 86–103.
Henson, M.A., 2003. Dynamic modelling of microbial cell populations. Curr. Opin. Biotechnol. 14, 460–467.
Hocquette, J.F., Lehnert, S., et al., 2007. Recent advances in cattle functional genomics and their application to beef quality. Animal 1 (1), 159–173.
Hoops, S., et al., 2006. COPASI—a complex pathway simulator. Bioinformatics 22 (24), 3067–3074.
Hotchkiss, J.R., Hermsen, E.D., Hovde, L.B., Simonson, D.A., Rotschafer, J.C., Crooke, P.S., 2004. Dynamic analysis of peritoneal dialysis associated peritonitis. ASAIO J. 50, 568–576.
Huang da, W., Sherman, B.T., et al., 2009. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 37 (1), 1–13.
Hupert, N., Mushlin, A.I., Callahan, M.A., 2002. Modelling the public health response to bioterrorism: using discrete event simulation to design antibiotic distribution centers. Med. Decis. Mak. 22, S17–S25.
Irish, J.M., Kotecha, N., et al., 2006. Mapping normal and cancer cell signalling networks: towards single-cell proteomics. Nat. Rev. Cancer 6 (2), 146–155.
Janes, K.A., Albeck, J.G., et al., 2003. A high-throughput quantitative multiplex kinase assay for monitoring information flow in signaling networks: application to sepsis-apoptosis. Mol. Cell Proteom. 2 (7), 463–473.
Janes, K.A., Yaffe, M.B., 2006. Data-driven modelling of signal-transduction networks. Nat. Rev. Mol. Cell Biol. 7 (11), 820–828.
Keeling, M.J., Rohani, P. 2008 Modelling Infectious Diseases in Humans and Animals, first ed. Princeton University Press, Princeton, NJ. ASIN: B005N8W2IC.
Kiran, K.L., Jayachandran, D., et al., 2009. Mathematical modelling of avascular tumour growth based on diffusion of nutrients and its validation. Can. J. Chem. Eng. 87 (5), 732–740.
Lewis, N.E., Schramm, G., Bordbar, A., Schellenberger, J., Andersen, M.P., Cheng, J.K., et al., 2010. Large-scale in silico modelling of metabolic interactions between cell types in the human brain. Nat. Biotechnol. 28, 1279–1285.
Liang, F., Takagi, S., et al., 2009. Multi-scale modeling of the human cardiovascular system with applications to aortic valvular and arterial stenoses. Med. Biol. Eng. Comput. 47 (7), 743–755.
Lloyd, C.J., Ebrahim, A., Yang, L., King, Z.A., Catoiu, E., O'Brien, E. J., et al., 2018. COBRAme: a computational framework for genome-scale models of metabolism and gene expression. PLoS Comput. Biol. 14 (7), e1006302. Available from: https://doi.org/10.1371/journal.pcbi.1006302. eCollection 2018 Jul.
Loor, J.J., Dann, H.M., et al., 2005. Temporal gene expression profiling of liver from periparturient dairy cows reveals complex adaptive mechanisms in hepatic function. Physiol. Genomics 23 (2), 217–226.
Margolin, A.A., Nemenman, I., et al., 2006. ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. BMC Bioinforma. 7 (Suppl. 1), S7.
Meade, K.G., Gormley, E., et al., 2006. Gene expression profiling of peripheral blood mononuclear cells (PBMC) from Mycobacterium bovis infected cattle after in vitro antigenic stimulation with purified protein derivative of tuberculin (PPD). Vet. Immunol. Immunopathol. 113 (1–2), 73–89.
Mi, Q., Rivière, B., Clermont, G., Steed, D.L., Vodovotz, Y., 2007. Agent-based model of inflammation and wound healing: insights into diabetic foot ulcer pathology and the role of transforming growth factor-β1. Wound Repair. Regen. 15, 671–682.
Muñoz-Elías, E.J., Timm, J., Botha, T., Chan, W.T., Gomez, J.E., McKinney, J.D., 2005. Replication dynamics of mycobacterium tuberculosis in chronically infected mice. Infect. Immun. 73 (1), 546–551.
Navratil, V., de Chassey, B., Combe, C.R., Lotteau, V., 2011. When the human viral infectome and diseasome networks collide: towards a systems biology platform for the aetiology of human diseases. BMC Syst. Biol. 5, 13.
Ong, S.E., Mann, M., 2005. Mass spectrometry-based proteomics turns quantitative. Nat. Chem. Biol. 1 (5), 252–262.
Palsson, B., 2000. The challenges of in silico biology. Nat. Biotechnol. 18 (11), 1147–1150.
Park, K., 2009. Textbook of Preventive and Social Medicine., 20th ed. Banarasidas Bhanot, Jabalpur.
Raichura, A., Valia, S., Gorinb, F., 2006. Dynamic modelling of alpha-synuclein aggregation for the sporadic and genetic forms of Parkinson's disease. Neuroscience 142 (3), 859–870.
Rejniak, K., Anderson, A., 2011. Hybrid models of tumor growth. Wiley Interdiscip. Rev. SystBiol. Med. 3, 115–125.

Resasco, D.C., et al., 2012. Virtual cell: computational tools for modeling in cell biology. Wiley Interdiscip. Reviews: Syst. Biol. Med. 4 (2), 129–140.
Rhee, S.Y., 2005. Bioinformatics. Current limitations and insights for the future. Plant. Physiol. 138 (2), 569–570.
Rowe, E., Palsson, B.O., King, Z.A., 2018. Escher-FBA: a web application for interactive flux balance analysis. BMC Syst. Biol. 12 (1), 84. Available from: https://doi.org/10.1186/s12918-018-0607-5.
Sanchez, M.A., Blower, S.M., 1997. Uncertainty and sensitivity analysis of the basic reproductive rate. Tuberculosis as an example. Am. J. Epidemiol. 145, 1127–1137.
Sayers, Eric W., et al., 2011. Database resources of the national center for biotechnology information. Nucleic Acids Res. 39 (Suppl. 1), D38–D51.
Scott, D.W., De Groot, A.S., 2010. Can we prevent immunogenicity of human protein drugs? Ann. Rheum. Dis. 69, i72–i76.
Somers, J., Smith, C., et al., 2006. Gene expression profiling of individual bovine nuclear transfer blastocysts. Reproduction 131 (6), 1073–1084.
Tomita, M., 2001. Whole-cell simulation: a grand challenge of the 21st century. Trends Biotechnol. 19 (6), 205–210.
Tuncbag, N., Kar, G., et al., 2009. A survey of available tools and web servers for analysis of protein-protein interactions and interfaces. Brief. Bioinform. 10 (3), 217–232.
van Riel, N.A., 2006. Dynamic modelling and analysis of biochemical networks: mechanism based models and model based experiments. Brief. Bioinform. 7, 364–374.
Vodovotz, Y., Csete, M., Bartels, J., Chang, S., An, G., 2008. Translational systems biology of inflammation. PLoS Comput. Biol. 4, e1000014.
Wang, Wei, et al., 2016. Suppressing disease spreading by using information diffusion on multiplex networks. Sci. Rep. 6.
Watterson, S., Ghazal, P., 2010. Use of logic theory in understanding regulatory pathway signaling in response to infection. Future Microbiol. 5 (2), 163–176.
Werner, T., 2008. Bioinformatics applications for pathway analysis of microarray data. Curr. Opin. Biotechnol. 19 (1), 50–54.
Woolf, P.J., Prudhomme, W., et al., 2005. Bayesian analysis of signaling networks governing embryonic stem cell fate decisions. Bioinformatics 21 (6), 741–753.
Wu, Y., Lu, Y., Chen, W., Fu, J., Fan, R., 2012. In silico experimentation of glioma microenvironment development and anti-tumor therapy. PLoS Comput. Biol. 8, e1002355.
Zampieri, G., Vijayakumar, S., Yaneske, E., Angione, C., 2019. Machine and deep learning meet genome-scale metabolic modeling. PLoS Comput. Biol. 15 (7), e1007084. Available from: https://doi.org/10.1371/journal.pcbi.1007084. eCollection 2019 Jul.
Zhang, Y., Athale, C.A., Deisboeck, T.S., 2007. 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, 96–107.

Long answer questions

1. Explain the role of bioinformatics in animal biotechnology.
2. Explain the common computational methods in systems biology.
3. Explain the concept of in silico modeling.
4. Discuss the advantages, disadvantages, and ethical issues of in silico modeling.
5. What are the different application areas of in silico modeling? Discuss in detail how in silico modeling is applied in one application area.

Short answer questions

1. Describe the template-based methods to reconstruct transcriptional regulatory networks.
2. What is the goal of in silico modeling?
3. What are the challenges in in silico modeling of infectious diseases?
4. What are the three types of cancer models discussed in the chapter?
5. Discuss the parameters considered for in silico modeling of infectious diseases.

Answers to short answer questions

1. The template-based transcriptional control network reconstruction method exploits the principle that orthologous proteins regulate orthologous target genes. In this approach, regulatory interactions are transferred from a genome (such as a genome of a model organism or well-studied organism) to the new genome.
2. The ultimate goal of in silico modeling in biology is the detailed understanding of the function of molecular networks as they appear in metabolism, gene regulation, or signal transduction.
3. There are two major challenges in modeling infectious diseases:
   a. Difficulty in finding the most appropriate set of parameters for the in silico modeling of infectious diseases is often a challenge.
   b. Understanding the results from all the complex interactions of parameters considered.
4. There are three types of cancer models. *Continuum models:* In these models, extracellular parameters can be represented as continuously distributed variables to mathematically model cell–cell or cell–environment interactions in the context of cancers and the tumor microenvironment. *Discrete models:* These models represent cancer cells as discrete entities of defined location and scale, interacting with one another and external factors in discrete time intervals according to predefined rules. *Hybrid models:* These models incorporate both continuum and discrete variables in a modular approach.
5. There are three types of parameters considered for in silico modeling of infectious diseases:

a. *Parameters derived from characteristics of agent:* Examples: concentration of the agent's antigen–host antibody complex; case fatality rate; strain of the agent; other genetic information of the agent; etc.
b. *Parameters derived from characteristics of host:* Examples: the total white blood cell counts; differential white blood cell counts, and/or much more sophisticated counts of specific blood cell types; blood levels of some specific cytokines, hormones, and/or neurotransmitters; daily calories, protein, and/or fat intake; daily amount of energy expended and/or duration of exercise; etc.
c. *Parameters derived from characteristics of environment:* Examples: host's ambient temperature; host's ambient atmospheric humidity; altitude; host's light–dark cycle; etc.

Yes/no type questions

1. Does bioinformatics play a role in animal biotechnology?
2. Does systems biology require computational approaches?
3. Does network analysis is important in disease modeling?
4. Can we develop e-cell for cellular phenotype simulation?
5. Does computational modeling or organ require mathematics?
6. Does metabolic modeling play a role in disease modeling?
7. Is E-CELL a computational model of a cell and help in analysis of biological systems in a cell?
8. Is a specific cancer can be modeled in silico?
9. Is it true that in silico disease modeling is not helpful in drug discovery and clinical trial.
10. Is metabolic pathway modeling possible using genomic data?

Answers to yes/no type questions

1. Yes—Bioinformatics is now essential for molecular aspects to phenotype analysis and prediction in animal biotechnology
2. Yes—Computational and mathematical modeling is essential in systems biology.
3. No—Network analysis is important for identification of pathways and key molecules in a disease.
4. Yes—Using various in silico modeling approach of cell, we can develop e-cell and study the phenotype.
5. Yes—Mathematical approach is essential in computational modeling of organ and disease for fixing various parameters of the organ or disease.
6. Yes—Without metabolic modeling approach, in silico disease modeling is not possible.
7. Yes—It helps in the analysis of biochemical pathways and simulation.
8. Yes.
9. No—Computational modeling of a disease in commonly practiced in drug discovery studies to understand efficacy and toxicity of the drug.
10. Yes.