SYSTEMS BIOLOGY AND SYNTHETIC BIOLOGY

系统与合成生物学

Edited by
Pengcheng Fu
Sven Panke

WILEY

SYSTEMS BIOLOGY AND SYNTHETIC BIOLOGY

Edited by

Pengcheng Fu

China University of Petroleum Beijing, China

Sven Panke

ETH Zurich, Institute of Process Engineering Zurich, Switzerland

WILEY

A JOHN WILEY & SONS, INC., PUBLICATION

Published by John Wiley & Sons, Inc., Hoboken, New Jersey
Published simultaneously in Canada

Library of Congress Cataloging-in-Publication Data:

Systems biology and synthetic biology / [edited by] Pengcheng Fu, Sven Panke.
p. ; cm.
Includes bibliographical references.
ISBN 978-0-471-76778-7 (cloth)
1. Biotechnology. 2. Genetic engineering. 3. Biological systems. I. Fu, Pengcheng. II. Panke, Sven.
[DNLM: 1. Genetic Engineering–methods. 2. Molecular Biology–methods. 3. Genes, Synthetic. 4. Genomics–methods. 5. Models, Genetic. 6. Systems Biology–methods. QU 450 S995 2009]
TP248.2.S97 2009
660.6 –dc22

2008027981

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

SYSTEMS BIOLOGY AND SYNTHETIC BIOLOGY

To my wife Juan Huang and sons, Eugene and Edgar

FOREWORD

The popular use of the term "systems biology" arose following the appearance of the first full genome sequences. These genome sequences suggested that we would have a full delineation of the molecular components of an organism. Expression profiling and proteomic data then could tell us when these components were actually used in a context-specific manner.

The need to track the interrelationship of all such components created the need to develop networks of the interactions of such components. Protein–protein interaction maps are one manifestation of this need, stoichiometric models are another; they are, however, amenable to rigorous mathematical analysis and prospective uses. Network reconstruction took center stage in systems biology, as networks describe the interactions between the gene products and the chemical compounds they make, provide context for high-throughput data mapping, and give the basis for mechanistic models that can compute phenotypic functions.

Having molecular manipulation tools and mathematical models in turn provides tools that allow the synthesis of biological components and biological functions. We thus witnessed the emergence of "synthetic biology." It is practiced on multiple scales, from component design, that is akin to classical molecular biology, to design of whole cell functions, such as metabolic engineering.

Thus, in retrospect we can state that genomics gave rise to systems, and systems biology in turn gave rise to synthetic biology. This of course is a simplified view, but provides a first-order approximation to the historical origin and appearance of these popularly used terms. This volume contains a series of chapters that highlight the development and status of the various aspects of systems and synthetic biology.

BERNHARD Ø. PALSSON

University of California, San Diego, CA

CONTRIBUTORS

Wenlong Cheng, Department of Biological and Environmental Engineering, Cornell University, Ithaca, New York 14853

Hidde de Jong, INRIA, Grenoble-Rhône-Alpes, 655 avenue de l'Europe, Montbonnot, 38334 Saint Ismier Cedex, France

Liang Ding, Department of Biological and Environmental Engineering, Cornell University, Ithaca, New York 14853

Pengcheng Fu, Faculty of Chemical Science and Engineering, China University of Petroleum, 18 Fuxue Road, Changping District, Beijing 102249, People's Republic of China

Hisakage Funabashi, Department of Biological and Environmental Engineering, Cornell University, Ithaca, New York 14853

Martin Fussenegger, Department of Biosystems Science and Engineering, ETH Zurich, Mattenstrasse 26, CH-4058 Basel, Switzerland

Johannes Geiselmann, Laboratoire Adaptation et Pathogénie des Micro-organismes, Université Joseph Fourier, UMR CNRS 5163, Bâtiment Jean Roget, Faculté de Médecine-Pharmacie, Domaine de la Mercie, La Tronche 38700, France

Marcus Graf, GENEART AG, Josef-Engert-Str. 11, Regensburg 93053 Germany

David Greber, Department of Biosystems Science and Engineering, ETH Zurich, Mattenstrasse 26, CH-4058 Basel, Switzerland

Claes Gustafsson, DNA2.0, Inc., 1430 O'Brien Drive, Menlo Park, California 94025

Matthias Heinemann, ETH Zurich, Institute of Molecular Systems Biology, Wolfgang-Pauli-Str. 16, 8093 Zurich, Switzerland

Philipp Holliger, MRC Laboratory of Molecular Biology, Cambridge CB2 2QH, UK

Cliff Hooker, Faculty of Education and Arts, School of Humanities and Social Science, The University of Newcastle, NSW, Australia

Mitsuhiro Itaya, Laboratory of Genome Designing Biology, Institute for Advanced Biosciences, Keio University, 403-1 Nipponkoku, Daihoji, Tsuruoka, Yamagata 997-0017, Japan

Andrew R. Joyce, Bioinformatics Program, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093

Hyun Uk Kim, Department of Chemical and Biomolecular Engineering (BK21 Program), Center for Systems and Synthetic Biotechnology, Institute for the BioCentury, KAIST, Daejeon 305-701, Korea

Jin Sik Kim, Department of Chemical and Biomolecular Engineering (BK21 Program), Center for Systems and Synthetic Biotechnology, Institute for the BioCentury, KAIST, Daejeon 305-701, Korea

Tae Yong Kim, Department of Chemical and Biomolecular Engineering (BK21 Program), Center for Systems and Synthetic Biotechnology, Institute for the BioCentury, KAIST, Daejeon 305-701, Korea

Hiroaki Kitano, Sony Computer Science Laboratories, Inc., 3-14-13 Higashi-Gotanda, Shinagawa, Tokyo 141-0022, Japan; The Systems Biology Institute, 6-31-15 Jingumae, Shibuya, Tokyo 150-0001, Japan; Department of Cancer Systems Biology, Cancer Institute of Japanese Foundation for Cancer Research, Tokyo, Japan

Sang Yup Lee, Department of Chemical and Biomolecular Engineering (BK21 Program), Center for Systems and Synthetic Biotechnology, Institute for the BioCentury, KAIST, Daejeon 305-701, Korea; Department of Bio and Brain Engineering, BioProcess Engineering Research Center and Bioinformatics Research Center, KAIST, Daejeon 305-701, Korea

James C. Liao, Chemical and Biomolecular Engineering Department, University of California, Box 951592, Los Angeles, California 90095

David Loakes, MRC Laboratory of Molecular Biology, Cambridge CB2 2QH, UK

Dan Luo, Department of Biological and Environmental Engineering, Cornell University, Ithaca, New York 14853

Juanita Mathews, Department of Molecular Biosciences and Bioengineering, University of Hawaii at Manoa, Honolulu, Hawaii 96822

Jens Nielsen, Department of Chemical and Biological Engineering, Chalmers University of Technology, Kemivägen 10, SE-412 96, Gothenburg, Sweden

Gordon S. Okimoto, Bioinformatics Shared Resource, Cancer Research Center of Hawaii, University of Hawaii at Manoa, 1236 Lauhala Street, Honolulu, Hawaii 96813

Bernhard Ø. Palsson, Department of Bioengineering, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093

Sven Panke, ETH Zurich, Institute of Process Engineering, Universitätsstr. 6, 8092 Zurich, Switzerland

Nokyoung Park, Department of Biological and Environmental Engineering, Cornell University, Ithaca, New York 14853

Delphine Ropers, INRIA Grenoble-Rhône-Alpes, 655 avenue de l'Europe, Montbonnot, 38334 Saint Ismier Cedex, France

Thomas Schoedl, GENEART AG, Josef-Engert-Str. 11, Regensburg 93053, Germany

Huidong Shi, Department of Pathology and Anatomical Sciences, Ellis Fischel Cancer Center, University of Missouri School of Medicine, M263 Medical Science Building, One Hospital Drive, Columbia, Missouri 65203; Department of Biochemistry and Molecular Biology, Medical College of Georgia, MCG Cancer Center CN4130, 1120 15th Street, Augusta, Georgia 30912

Seung Bum Sohn, Department of Chemical and Biomolecular Engineering (BK21 Program), Center for Systems and Synthetic Biotechnology, Institute for the BioCentury, KAIST, Daejeon 305-701, Korea

Dorothea K. Thompson, Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907

Soong Ho Um, Department of Biological and Environmental Engineering, Cornell University, Ithaca, New York 14853

Jordi Vallverdú, Philosophy Department, Universitat Autònoma de Barcelona, E-08193 Bellaterra (BCN), Catalonia

Goutham Vemuri, Center for Microbial Biotechnology, Denmark Technical University, Biocentrum-DTU, Lyngby DK2800, Denmark

Ralf Wagner, University of Regensburg, Molecular Microbiology and Gene Therapy, Franz-Josef-Strauss Allee 11, Regensburg 93053, Germany

Xiu-Feng Wan, Department of Microbiology, Miami University, Oxford, Ohio 45056

Guangyi Wang, Department of Oceanography, University of Hawaii at Manoa, 1000 Pope Road, Honolulu, Hawaii 96811

Michael X. Wang, Department of Pathology and Anatomical Sciences, Ellis Fischel Cancer Center, University of Missouri School of Medicine, M263 Medical Science Building, One Hospital Drive, Columbia, Missouri 65203

Wilson W. Wong, Chemical and Biomolecular Engineering Department, University of California, Box 951592, Los Angeles, California 90095

Jianfeng Xu, Department of Biological and Environmental Engineering, Cornell University, Ithaca, New York 14853

Hongseok Yun, Department of Chemical and Biomolecular Engineering (BK21 Program), Center for Systems and Synthetic Biotechnology, Institute for the BioCentury, KAIST, Daejeon 305-701, Korea

CONTENTS

1

INTRODUCTION

Pengcheng Fu

Faculty of Chemical Science & Engineering, China University of Petroleum, Beijing, 18 Fuxue Road, Changping District, Beijing 102249, China

In the twentieth century, engineering sciences have inspired numerous successful applications in the fields of manufacturing, electronics, communications, transportation, computer and networks, and so on. Compared to the engineering systems, biological systems are more complex and their mechanisms are less known. Historically, biological questions have been approached by a reductionist paradigm that is completely different from methodologies being applied to engineering systems. This reductionist way of thinking was based on the assumption that by unraveling the function of all the different components the information gained could be used to piece together the puzzle of complex cellular networks [1]. The research paradigm has dominated mainstream biology with enormous progresses in accumulating biological information at genetic and protein levels. However, this is a slow and exhaustive process that fails to adequately approach the true complexities of living phenomena and is of limited relevance to biological systems as a whole.

The fast-growing applications of genomics and high-throughput technologies have led to recognition of the limitations of the reductionist/atomistic view of the world. It is realized that a new systems biology paradigm is needed for the next level of understanding of the functions of the genes and proteins, and the regulation of intracellular networks that cannot be obtained by studying the individual constituents on a part-by-part basis. It is also realized that there is great similarity between biology and engineering at the system level, despite their obviously different physical implementation, and that important research challenges in biology may have parallels with those in complicated engineering systems [2]. This similarity forms a basis for the

Systems Biology and Synthetic Biology Edited by Pengcheng Fu and Sven Panke

introduction of synthetic biology or the engineering applications within biological systems.

Systems biology attempts to investigate the behavior and relations of all the elements in a particular biological system while it is functioning [3,4]. It aims at system-level understanding of biological processes and biochemical networks as a whole. This "system-oriented" new biology is shifting our focus from examining particular molecular details to studying the information flows at all biological levels: Genomic DNA, mRNA, proteins, informational pathways, and regulatory networks. Systems biology approaches seek to study the complexity of life to help in understanding how the cellular networks work together. To this end, the approach emphasizes the investigation of biological phenomena by considering system structures, system dynamics, control methods, and design methods [5,6]. It requires a broad interdisciplinary integration of molecular and cell biology, biochemistry, informatics, mathematics, computing, and engineering.

Synthetic biology is a recently emerging field that applies engineering formalisms to design and construct new biological parts, devices, and systems for novel functions or life forms that do not exist in nature. This "engineering" biology relies on and shares tools from genetic engineering, bioengineering, systems biology, and many other engineering disciplines. Synthetic biology is also different from these subjects, in both insights and approach. The synthetic biology study will not only investigate the effects of genetic and pathway modification or the cellular responses on genetic variation/environmental perturbation, but also design and build biological systems with novel cellular functions, combining *in silico* and *in vivo* experimental approaches.

Recently, synthetic biology has been redefined as (1) the design and construction of new biological parts, devices, and systems that do not already exist in the nature and (2) the redesign of existing, natural biological systems for useful purposes (http://syntheticbiology.org/). More specifically, synthetic biology aims to design and build engineered biological systems that process information, manipulate chemicals, fabricate materials, produce energy, provide food, and maintain and enhance human health and our environment (see the Wikipedia: http://en.wikipedia.org/wiki/Synthetic_biology).

Synthetic biology is a "bottom-up" approach, in which basic functional elements of replication, self-assembly, growth, metabolism, repair, signaling, and regulation are defined and assembled into life forms or biomaterials with new properties and behavior. Synthetic biology makes use of and is complementary to systems biology, which focuses "from the top down" on the fully integrated networks of function and control in living cells, and their responses to various perturbations. While preceded by some pioneering work, systems biology and synthetic biology are the new subdisciplines that were invisible 20 years ago. A Google search under these categories now yields more than a million web pages and systems and synthetic biology research departments have grown from scratch to 100 + staff in a decade or less. The aspiration of systems biology is no less than the full mastery of cellular and multicellular dynamics, both epistemic and technical. This has made enormous implications for health, through the development of new treatments and (as importantly) new preventive measures, and for agriculture and ecology more generally through the

development of new strains, disease controls, and management processes. It also promises/threatens a host of new security and military applications. The complementary mission of synthetic biology is no less than the systematic extension of engineering, in all its aspects, so as to encompass the biological realm and, conversely, the systematic integration of biological elements into engineered systems. This encompasses practical applications from biofuel to beer production, from sensors to cyborgs.

There are several foundational and illustrative examples regarding applications of systems biology and synthetic biology available in the literature.

A gene deletion perturbation experimental paradigm has emerged in systems biology beginning with research conducted by Trey Ideker, Timothy Galitski, and Leroy Hood [3]. The authors used a systems approach to explore, expand, and refine the understanding of the yeast galactose utilization (GAL) system. A regulatory network model was used to predict changes in gene expression. The authors perturbed the galactose pathway by deleting each of the nine galactose genes of interest. They carried out replicate hybridizations in four different DNA microarrays for each perturbed condition to obtain robust estimates of how the gene expression profile of each knockout strain differed from that of the wild type. Expression data for each perturbation were visually superimposed on the metabolic network. The predicted and observed cellular responses to the perturbation were found to be consistent for most cases. The authors then used the discrepancies between the predicted and observed expression responses to suggest possible refinement to the model [3].

As another example, Pamela Silver, Professor in the Department of Systems Biology at Harvard Medical School, and her group have successfully created a "memory" loop in yeast cells, using two synthesized, transcription factor coding genes. The first gene, which was designed to switch on when exposed to galactose, created a transcription factor that grabbed on to, and thus activated, the second gene. Cellular memory may enable the yeast cells to remember their functional network states. In addition to the reconstruction of the dynamics of cell memory, the researchers have also established a mathematical model for the prediction on the functional outputs of the transcription factor to control how much of a particular protein the gene should make (see BiomimicryNews.com: Scientists synthesize memory in yeast cells, 9/15/2007, http://www.biomimicrynews.com/research/Scientists_synthesize_memory_in_yeast_cells.as). This application exemplifies the core concept that systems biology and synthetic biology should be integrated for understanding life phenomena and creating novel biological modules to modify an existing biological system.

More recently, Craig Venter and his synthetic biology research team have constructed a largest man-made DNA structure on the Earth out of laboratory chemicals. They used the bacterium *Mycoplasma genetalium* as the template to build the synthetic chromosome that contains 381 genes, that is, 582,970 bp long. This wholly artificial genetic circuit was then transplanted into a living *M. genetalium* cell. The resultant "*Mycoplasma laboratorium*" is expected to be able to replicate itself with its synthetically reconstructed DNA, making it the most fully synthetic organism to date, although the molecular machinery and chemical environment that would allow it to replicate would not be synthetic (http://www.guardian.co.uk/science/2007/oct/06/

genetics.climatechange). The building blocks of DNA—adenine (A), guanine (G), cytosine (C), and thiamine (T)—are not easy chemicals to artificially synthesize into chromosomes. A key point for the Craig Venter Institute to achieve their goal of synthetic biology is that they could use homologous recombination (a process that cells use to repair damage to their chromosomes) in the yeast *Saccharomyces cerevisiae* to rapidly build the entire bacterial chromosome from large subassemblies (http://www.syntheticgenomics.com/press/2008-01-24.htm). The team attempts to continue to work for the creation of a living bacterial cell based entirely on the synthetically made genome. It is obvious that using the same approach, existing life forms can be modified by adding components to them or by taking components away from them.

These examples illustrate that we are able to not only "read" the genetic code to understand living systems but also "write" the message for the creation of new life forms. All this fuels the need to frame these latest developments that promise to revolutionize our understanding of biology, blur the boundaries between the living and the engineered in a vital new bioengineering, and transform our daily relationship to the living world.

Systems biology and synthetic biology are emerging as two complementary approaches that embody the breakthrough in biology and invite application of engineering principles. Although systems biology and synthetic biology approach biological problems with different emphasis, they are indeed the two sides of the same coin. Borrowing from ancient Chinese intellectual thinking, we can see systems biology and synthetic biology as Yin and Yang of a research and development framework in the new biological paradigm. For systems biology, all of the "omics" experiments are used to discover life phenomena and to empower scientists to increase fundamental understanding of complex living systems. For synthetic biology, existing life forms can be modified based on the existing body of knowledge. Since systems biology and synthetic biology are like Yin and Yang in nature, we will keep the mutuality of the two approaches in mind when writing the book. We will demonstrate that systems biology relies on synthetic biology technologies to perturb and monitor responses of the biological systems, while synthetic biology depends on the knowledge obtained from systems biology approaches for design and implementation. The development of

systems biology and synthetic biology is thus cyclic; understanding (by systems biology) and creation (by synthetic biology) will continuously enhance and transform each other and thus converge in an iterative fashion.

We have assembled a group of investigators/educators who are at the cutting edge of biological research in bioinformatics, functional genomics, genome-scale modeling, and systems biology and synthetic biology. This book will benefit undergraduate and graduate students, scientific researchers, university lectures, and those with larger managerial responsibilities. It should be of especial interest to business development professionals working in biotechnological and pharmaceutical companies since novel products/functions are the goals of systems biology and synthetic biology. The first of its kind, this book aims to become an important reference book for researchers in various engineering fields, such as chemical engineering, mechanical engineering, and civil engineering, who consider systems analysis and engineering applications in biology as their next frontier.

ORGANIZATION OF THE BOOK

This book is organized into modular, stand-alone topics related to systems biology and synthetic biology. Chapter 2, by Michael Wang and Huidong Shi, provides an overview of the modern molecular biology to the readers from backgrounds other than biology. Chapter 3, by Xiu-Feng Wan and Dorothea Thompson, introduces recent advances in high-throughput technologies and functional genomics, with application examples to illustrate how the "omics" data can be used in aid of the establishment of the linkages between transcriptome profiling and biological functions. In Chapter 4, Gordon Okimoto explores and discusses mathematical modeling for "omics" data fusions to predict the global behavior of complex biological systems as networks of interacting genes, proteins, and metabolites. They use algorithms for cancer diagnosis and identification of subcategories. Chapter 5, by Mitsuhiro Itaya, provides a novel approach where a whole genome is cloned into another species. In Chapter 6, Andrew Joyce and Bernhard Palsson present and discuss *in silico* genome-scale metabolic models for microorganisms, such as *Escherichia coli* and the yeast *S. cerevisiae*, using the constraints-based approach. The constraints-based modeling approach they describe, using flux balance analysis and linear programming, is currently the only methodology capable of delivering a high, indeed a surprisingly high, degree of correlation between the predictions of genome-scale models and independently obtained experimental data.

Chapter 7, by Delphine Ropers, Hidde de Jong, and Johannes Geiselmann, discusses the application of mathematical modeling to study the genetic regulatory networks that control gene expression in an organism and the adaptation of its cells to the environment. Illustrative examples using *E. coli* show that many aspects of their structure and dynamics can be fruitfully compared with the principles governing man-made systems. In Chapter 8, Vallverdú and Gustafsson bring up bioethic issues caused by our systems biology and synthetic biology efforts. It is obvious that the benefits from these emerging fields will far outweigh the negative impacts. However, as they

point out, the technologies for systems biology and synthetic biology which were not thinkable a couple of decades ago have the potential for misuse. Systems biology and synthetic biology must be operated within a framework of safety, ethics and public acceptance. Chapter 9 is contributed by Goutham Vemuri and Jens Nielsen. In this chapter, the authors review the use of the yeast *S. cerevisiae* as a prototype for systems biology and synthetic biology research. In Chapter 10, Sang Yup Lee et al. examine the construction of genome-scale metabolic models in global understanding of metabolism and physiological characteristics, and also in designing metabolic engineering strategies for the enhanced production of various bioproducts. This chapter can be related in part as application examples to the genome-scale modeling framework outlined in Chapter 6.

In Chapter 11, Matthias Heinemann and Sven Panke sketch a draft picture of synthetic biology as a new bio-based discipline. In their essay, synthetic biology is put into perspective with its scientific counterpart, the field of systems biology, and also the fundamental differences to other "bioengineering" areas such as metabolic engineering or protein engineering. In Chapter 12, Marcus Graf, Thomas Schoedl, and Ralf Wagner present their work for rational gene design and *de novo* gene construction. Since such genes do not exist in nature, they have to be constructed and cloned *de novo* from synthetic oligonucleotides. Advanced synthetic biology approaches are thus needed for generation of proteins with novel functions, new metabolic pathways, or even artificial organisms.

Self-replication is defined as the ability of a system to direct the synthesis of accurate copies of itself from dispersed building blocks. Chapter 13, written by Philipp Holliger and David Loakes, describes the opportunities and challenges for the engineering and bottom-up assembly of artificial self-replicating systems with quasibiotic properties and their potential applications in molecular sensing, computing, and the manufacture of nanodevices. In Chapter 14, Wilson Wong and James Liao present and discuss synthetic approaches similar to the design of engineering machinery for construction of biological circuits based on physical concepts, guided by mathematical models, and constrained by biological and chemical realities. In Chapter 15, David Greber and Martin Fussenegger discuss the state of the art in the field of synthetic genetic networks with particular emphasis on relating network architecture and design to network characteristics. They examine engineered devices such as toggle switches, oscillating networks, and molecular sensors that possess increasingly sophisticated functionality.

Chapter 16 by Hiroaki Kitano focuses on biological robustness as a fundamental feature of living systems, whereby its relationship with evolution, trade-offs among robustness, fragility, resource demands, and performance provides a possible framework for how biological systems have evolved and become organized. In this way, the understanding of robustness and its intrinsic properties provides us with a deeper understanding of biological systems, their anomalies, and countermeasures to reduce these. In Chapter 17, Wenlong Cheng et al. address the properties and functions of oligonucleic acids used as generic, instead of genetic, materials via nucleic acid engineering to utilize a myriad of molecular tools, mostly enzymes, to design and build novel biological systems with desired functions. They also discuss various approaches