

# Preface

The arrival of gene expression microarray has forever changed research in bioinformatics and computational biology. Biologists and clinical researchers now care about how genes and proteins are related to the biological pathways they spent their whole lives studying. Now, the scientific community believes that genes or proteins interact in a complex manner, and that the systematic decoding of their interaction will lead to a better understanding of the mechanism of many diseases and, subsequently, to improved human health and patient care. Generally speaking, the study of systems biology attempts to identify individual components and molecules within the vast networks created by biological molecules that regulate and control life. The algorithms and techniques of computational systems bioinformatics described in this book are consistent with those of systems biology in terms of the structure (e.g. gene regulatory and biochemical networks), dynamics, and design methods of the biological system to judge how much we understand the system. This book discusses these issues from data modeling and computational viewpoints.

The rapid advances in biotechnology — such as gene expression microarray, protein profiling mass spectrometry, and high-content cellular analysis — allow the collection of large amounts of heterogeneous data at multiple scales of biology within a very short period of time. Much of these data are new to bioinformatics and computational biology communities, let alone the integration of them in modeling complex biological systems. These new advances have accelerated the need for developing new analytic and modeling techniques from the systems perspective. Although many books on the

subject of bioinformatics have been published, they often focus on one particular technique or modality, such as gene sequence analysis or microarray analysis. This new book on bioinformatics will not cover traditional bioinformatics techniques such as gene sequence, motif analysis, or protein structure databasing; instead, this book attempts to keep up with the quick pace of change of this field from the systems perspective, and covers three topics from computational systems biology and their applications in systems biology.

First, we focus on the prevalent high-throughput techniques used in computational systems bioinformatics, e.g. gene expression microarray studies. The topics vary from gene preprocessing to gene selection, gene classification, gene prediction, and inference of gene regulatory networks. Second, we discuss how to combine DNA sequence information and gene microarray information in order to study gene regulation. After that, we move the discussion from DNA and RNA to proteins; in particular, we provide application examples of protein structure prediction and proteomics biomarker discovery. Third, we review the emerging subject of image informatics for molecular and cellular imaging, which connects molecular- and cellular-level information, generates objective quantitative phenotypes, and provides the scientific community with a new and powerful approach to study cells and cell interaction at a systems level.

There are many individuals we would both like to thank, without whose efforts this volume would not have become a reality. First and foremost, our thanks go to all of the present and past members of the HCNR-BWH Center for Bioinformatics, Harvard Medical School. Their expertise and dedication in using the systems approach in studying bioinformatics and computational biology problems motivated us to write this book volume. In particular, we would like to thank Dr Kuang-Yu Liu of the Brigham and Women's Hospital (BWH), Harvard Medical School, for his contribution in Chapters 11 and 12. We would also like to express our great thanks to Dr Liu, Ms Baillie Yip, and Mr Nicholas Young for their review of the book volume in making this book's contents as strong as they are. We would also like to acknowledge the continued support of the National Institutes

of Health (NIH), the National Science Foundation (NSF), and the Harvard Center for Neurodegeneration and Repairs (HCNR) over the years that have led to many original research results described in this book. Finally, we thank our publisher, World Scientific Publishing, for inviting us to write this book and tirelessly pushing us to finish it on schedule.