

Preface

I am pleased to offer the research community my second book-length contribution to the field of bioinformatics. My first book, *Advanced Analysis of Gene Expression Microarray Data*, was published in 2006 by World Scientific as part of its Science, Engineering, and Biology Informatics (SEBI) series. I first became involved in the study of bioinformatics in 1998 and, over the ensuing decade, have been struck by the enormous quantity of data being generated and the need for effective approaches to its analysis.

The analysis of protein–protein interactions (PPIs) is fundamental to the understanding of cellular organizations, processes, and functions. It has been observed that proteins seldom act as single isolated species in the performance of their functions; rather, proteins involved in the same cellular processes often interact with each other. Therefore, the functions of uncharacterized proteins can be predicted through comparison with the interactions of similar known proteins. A detailed examination of a PPI network can thus yield significant new insights into protein functions. These interactions have traditionally been examined via intensive small-scale investigations of a small set of proteins of interest, each yielding information about a limited number of PPIs. The existing databases of PPIs have been compiled from such small-scale screens, presented in individual research papers. Because these data were subject to stringent controls and evaluation in the peer-review process, they can be considered to be fairly reliable. However, each experiment observes only a few interactions and yields a data set of very limited size. Recent large-scale investigations of PPIs using such techniques as two-hybrid systems, mass spectrometry, and protein microarrays have enriched the available protein interaction data and facilitated the construction of integrated PPI networks. The resulting large volume of PPI data has posed a challenge to experimental investigation. Consequently, computational analysis of the networks has become a necessary tool for the determination of functionally associated proteins.

This book is intended to provide a comprehensive understanding of the computational methods available for the analysis of PPI networks. It offers an in-depth survey of a range of approaches to this analysis, including statistical, topological, data-mining, and ontology-based methods. The fundamental principles underlying each of

these approaches are discussed, along with their respective benefits and drawbacks. Suggestions for future research are also offered. In total, this book is intended to offer bioinformatics researchers a comprehensive and practical guide to the analysis of PPI networks, which will assist and stimulate their further investigation.

Some knowledge on the part of the reader in the fields of molecular biology, data mining, and statistics is assumed. Apart from this, the book is designed to be self-contained, as it includes introductions to the fundamental concepts underlying data generation and analysis. Thus, this book is expected to be of interest to a variety of researchers. It can be used as a textbook for advanced graduate courses in bioinformatics, and most of its content has been tested in the author's graduate-level course in this field. In addition, it can serve as a resource for graduate students seeking topics for investigation. The book will also be useful to researchers involved in computational biology in universities, organizations, and industry. For this audience, it will provide guidance on the techniques available for analysis of PPI networks. Research professionals interested in expanding their knowledge base can draw upon the material presented here to gain an understanding of principles and methods involved in this growing and highly significant field.

ACKNOWLEDGMENTS

I would like to express my deepest thanks to my doctoral students, Pritam Chanda, Young-rae Cho, Woo-chang Hwang, Taehyong Kim, and Lei Shi, for their excellent technical contributions. I am also highly appreciative of the editorial work of Rachel Ramadhyani.

The inspiration for this book was an invitation from Ms. Lauren Cowles, a senior editor from Cambridge University Press. I would like to express my special thanks to her.

Aidong Zhang
Buffalo, New York