

Preface

Over the past few decades, developments in genomic, molecular research, as well as information technologies have produced a tremendous amount of information related to molecular biology. This made computing skills become essential in molecular biology research. Computational methods have evolved to analyze and interpret various types of data, such as nucleotide and amino acid sequences, protein domains, and protein structures, etc. Bioinformatics is the name given to these mathematical and computing approaches used for understanding biological processes and becomes an important part of many areas of biology. Bioinformatics not only provides tools we can use to understand the basic aspect of biology, including development, metabolism, adaptation to the environment, genetic variations of individual, and evolution, but also facilitates our understanding of the diseases processes through the analysis of molecular sequence data.

This book is a handbook of methods and protocols for biologists. It aims at undergraduate, graduate students and researchers originally trained in biological or medical sciences who need to know how to access the data archives of genomes, proteins, metabolites, gene expression profiles and the questions these data and tools can answer, for example, how to draw inferences from data archives and how to make connections among them to deduce useful predictions. For each chapter, the conceptual and experimental background is provided, together with specific guidelines for handling raw data, including preprocessing and analysis.

The book is structured in three parts. Part I introduces the basic knowledge about popular bioinformatics tools, databases and web resources, including online sequence databases, sequence alignment, predicting DNA and protein function from sequence, protein structure prediction and analysis, molecular phylogeny and evolution. Part II presents examples of Omics bioinformatics applications, including genetic variation and human disease, gene expression profile and data management, qualitative and quantitative proteomics, bioinformatics for metabolomics, and integrating Omics data for pathways and interaction networks. Part III provides basic statistical analysis skills and programming skills needed to handle and analyze Omic datasets.

Bioinformatics is an interdisciplinary field involving molecular biology and genetics, computer science, mathematics, and statistics. It is too far broad to be understood by one person. Thus, this book is written by multiple authors, each of whom brings a deeper knowledge of the subject. I wish to express my gratitude to all authors for their dedication in providing excellent chapters, some of the data and examples presented in the book are the results of their own research. As for any omissions or errors, the responsibility is mine.

In preparing the book, we read many textbooks and published papers and viewed many websites, we sincerely apologize to those authors and researchers whose work we did not cite.

Enjoy reading.

Wang Ziling

College of Life Sciences and Bioengineering,
Beijing Jiaotong University, Beijing

October, 2017