Book review on Bioinformation Discovery: Data to knowledge in Biology

Subhash Janardhan Bhore

Department of Biotechnology, Faculty of Applied Sciences, AIMST University, Bedong-Semeling Road, Bedong, 08100, Kedah Darul Aman, Malaysia; Bhore Subhash J - Email: subhashbhore@yahoo.com; *Corresponding author

Researchers in almost all disciplines of Biology use mathematical, statistical models and computer programs to analyze and validate biological data. Systematic data mining is frequently completed in recent years using Bioinformatics soft-wares. The available Bioinformatics techniques and tools help annotate functions for newly generated data in biological investigation. The book (ISBN 978-1-4419-0518-5 e-ISBN 978-1-4419-0519-2) entitled "Bioinformation Discovery: Data to knowledge in Biology" by Dr. Pandjassarame Kangueane is published by Springer USA. This book describes the use of Biological data and Bioinformatics techniques and tools in Biological knowledge discovery. The book contains ten chapters. The first chapter describes concepts, principles and components of Bioinformatics with its various applications in agriculture, healthcare and industrial biotechnology sectors. The author precisely highlights that a sound knowledge in the basic concepts of bioinformatics is the real foundation for Bioinformation discovery from data. In the same chapter he also discussed the importance of Bioinformatics in Drug Discovery. The second chapter describes biological datasets and methods to create data subsets from primary databases like GenBank, EMBL, DDBJ and Protein databank (PDB). The provided dataset examples illustrate the importance of specialized datasets in Biological knowledge discovery.

There are number of Bioinformatics tools and techniques freely available online for data mining and knowledge discovery in specific areas of Biology. In this context, Chapter 3 (Tools and Techniques) gives an account of some of the commonly used tools frequently cited in this book. The ongoing trends in biological sciences are to understand the protein-protein interaction within and across cells. The chapter 4 (Protein Subunits Interaction) highlights the use of 3D structure datasets to derive information and geometrical (interface size, planarity, sphericity, and complementarity) and chemical (the types of amino acid chemical groups, hydrophobicity, electrostatic interactions, and H-bonds) properties for protein subunit interactions. The formation and mechanism of folding for homodimers (dimers of identical subunits) is intriguing. The fifth chapter relates structural features to homodimer folding and binding using structural datasets. The provided schematic illustrations help in understanding the homodimer folding and binding mechanism. In Chapter 6 (Fusion Proteins), the author discusses on fusion proteins in one species mimic operons, protein-protein interaction, alternative splicing and multiple functions in another species. This chapter help relate gene fusion in establishing optimal dynamics and kinetics.

The study of genetic variation among ethnic groups (American Indian, Australian aboriginal, Black, Caucasoid, Oriental, Hispanic, Mixed race, and Pacific Islander) is fascinating. Chapter 7 (Major Histocompatibility Complex (MHC) and Peptide Binding) discusses major histocompatibility complex (MHC) genes and their sequence polymorphism (sequence level variation). The discussion on issues in human MHC genes and methods for MHC peptide binding prediction helps to get insights into gene polymorphism and its impact in population based medicine. Chapter 8 (HLA Supertypes) gives the precise account of the human leukocyte antigen (HLA) alleles that demonstrates functional overall lap among HLA alleles for population coverage. To take readers further, in Chapter 9 (T-Epitope Designer) the author described a web server named 'T-Epitope Designer' which can be used to facilitate HLA-peptide binding prediction for peptide vaccine design. The last Chapter (Chapter 10: Eukaryotic Genes, Functions, Genomes, Design, and Evolution) has highlighted the main issues associated with the study of eukaryotic genes, genomes, function, design, and their evolution. Database such as U-Genome (database of relevant genome information in unicellular eukaryotes) which plays an important role in the understanding of unicellular genomes, their design and evolution is also discussed along with some other databases like ExInt (database of exon–intron gene structure in eukaryotic genes). This chapter gives a summary of several issues (exons, introns, single exon genes, intron organization and chromosome content) in comparative and functional genomics.

In summary, this is a very useful book. It is written meticulously by an expert in the emerging field of Bioinformation discovery. This book provides plenty of interesting tips to use data in knowledge discovery. Each chapter concludes with a set of exercise problems for mastery of the subject. The language used in the book is extremely lucid, and hence book is very useful for undergraduate. This book is highly recommended for post-graduate students in Biology and Life Sciences. It should be noted that the book is available in both hard bound and electronic versions.