

Chapter 2

Bioinformatics: Applications and Implications

Kijpokin Kasemsap

Suan Sunandha Rajabhat University, Thailand

ABSTRACT

This chapter describes the overview of bioinformatics; bioinformatics, data mining, and data visualization; bioinformatics and secretome analysis; bioinformatics, mass spectrometry, and chemical cross-linking reagents; bioinformatics and Software Product Line (SPL); bioinformatics and protein kinase; bioinformatics and MicroRNAs (miRNAs); and clinical bioinformatics and cancer. Bioinformatics is the application of computer technology to the management and analysis of biological data. Bioinformatics is an interdisciplinary research area that is the interface between biology and computer science. The primary goal of bioinformatics is to reveal the wealth of biological information hidden in the large amounts of data and obtain a clearer insight into the fundamental biology of organisms. Bioinformatics entails the creation and advancement of databases, algorithms, computational and statistical techniques, and theory to solve the formal and practical problems arising from the management and analysis of biological data.

INTRODUCTION

Bioinformatics involves algorithms to represent, store, and analyze the biological data (Li, 2015), including DNA (deoxyribonucleic acid) sequence, RNA (ribonucleic acid) expression, protein, and small-molecule abundance within cells (Altman, 2012). Bioinformatics is an interdisciplinary approach utilizing both data collection and information modeling to organize, analyze, and visualize the biological data (Al-Ageel, Al-Wabil, Badr, & AlOmar, 2015) based on the optimal use of big data gathered in genomics, proteomics, and functional genomics research (Song, Kim, Zhang, Ding, & Chambers, 2014).

Big data contains the very large sets of data that are produced by people using the Internet, and that can only be stored, understood, and utilized with the help of special tools and methods (Kasemsap, 2016a). Big data can enhance the significant value by making information transparent and usable at much higher frequency in modern operations (Kasemsap, 2017a). The important goal of bioinformatics is to

facilitate the management, analysis, and interpretation of biological data from biological experiments and observational studies (Moore, 2007).

With the increasingly accumulated data from high-throughput technologies, study on biomolecular networks has become one of the most important perspectives in systems biology and bioinformatics (Zhang, Jin, Zhang, & Chen, 2007). Many bioinformatics tools developed for DNA microarrays can be reused in proteomics, however, the uniquely quantitative nature of proteomics data also offers the novel analysis possibilities, which directly enhance the biological mechanisms (Kumar & Mann, 2009). DNA microarray enables investigators to simultaneously study the gene expression profile and gene activation of thousands of genes and sequences (Chen, Weixing, Sheng, & Zhilong, 2013). Health care-related bioinformatics databases are increasingly offering the possibility to maintain, organize, and distribute the DNA data (Dalpé & Joly, 2014).

Recently, the interests in proteomics have been increased, and the proteomic methods have been widely applied to many problems in cell biology (Haga & Wu, 2014). Proteomics is the large-scale study of proteins, particularly their structures and functions (Anderson & Anderson, 1998), and has become one of the most important approaches to analyzing and understanding biological systems (Oveland et al., 2015). Protein-protein interactions have been at the focus of computational biology in recent years (Hooda & Kim, 2012). Solving the problem of predicting protein interactions from the genome sequence leads to the obvious understanding of complex networks, evolution, and human disease (Reimand, Hui, Jain, Law, & Bader, 2012). Nibbe et al. (2011) stated that protein interaction databases play an increasingly important role in systems biology approaches to the study of disease.

Bioinformatics tools for proteomics, also known as proteome informatics tools, span a large panel of very diverse applications, ranging from simple tools to compare protein amino acid compositions to sophisticated software for the large-scale protein structure determination (Palagi, Hernandez, Walther, & Appel, 2006), and deal with the analysis of protein sequences (one dimensional pattern) and structures (three-dimensional pattern) (Mardia, 2013). Courcelles et al. (2011) indicated that the novel and improved computational tools are required to transform large-scale proteomics data into valuable information of biological relevance.

This chapter is based on a literature review of bioinformatics. The extensive literature of bioinformatics provides a contribution to practitioners and researchers by describing the applications and implications of bioinformatics in order to maximize the scientific impact of bioinformatics in the fields of biology and computer science.

BACKGROUND

In recent years, there has been an explosion of biological data stored in the central databases, tools to handle the data, and educational programs to train scientists in utilizing bioinformatics resources (Shachak, 2006). Bioinformatics is a significant discipline that performs the analysis, modeling, and simulation of complex biological systems by using computer science approach, dealing with large amounts of biological data (Merelli, Pérez-Sánchez, Gesing, & D'Agostino, 2014). Bioinformatics enables biomedical investigators to exploit the existing and emerging computational technologies to effectively store, mine, retrieve, and analyze the biological data from genomics and proteomics technologies (Fenstermacher, 2005). Modern technologies have rapidly transformed biology into a data-intensive discipline (Strizh, 2006).

20 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/bioinformatics/182940

Related Content

Clustering Genes Using Heterogeneous Data Sources

Erliang Zeng, Chengyong Yang, Tao Liand Giri Narasimhan (2010). *International Journal of Knowledge Discovery in Bioinformatics* (pp. 12-28).

www.irma-international.org/article/clustering-genes-using-heterogeneous-data/45163

Implementation of n-gram Methodology for Rotten Tomatoes Review Dataset Sentiment Analysis

Prayag Tiwari, Brojo Kishore Mishra, Sachin Kumarand Vivek Kumar (2017). *International Journal of Knowledge Discovery in Bioinformatics* (pp. 30-41).

www.irma-international.org/article/implementation-gram-methodology-rotten-tomatoes/178605

Evaluation of Coupled Nuclear and Cytoplasmic p53 Dynamics

Tingzhe Sun, Meihong Cai, Jun Cuiand Pingping Shen (2013). *Bioinformatics: Concepts, Methodologies, Tools, and Applications* (pp. 1184-1195).

www.irma-international.org/chapter/evaluation-coupled-nuclear-cytoplasmic-p53/76114

The Present and the Future Perspectives of Biological Network Inference

Paola Leccaand Alida Palmisano (2012). *Systemic Approaches in Bioinformatics and Computational Systems Biology: Recent Advances* (pp. 118-140).

www.irma-international.org/chapter/present-future-perspectives-biological-network/60831

Collaborations for Innovation in the Bio-Pharmaceutical Industry: An Exploratory Analysis on the Role of Platform Biotech Firms

Mattia Bianchi, Lorenzo Boscherini, Davide Chiaroni, Federico Frattiniand Vittorio Chiesa (2012). *Pharmacoinformatics and Drug Discovery Technologies: Theories and Applications* (pp. 72-91).

www.irma-international.org/chapter/collaborations-innovation-bio-pharmaceutical-industry/64067