Applied Bioinformatics and Biostatistics in Cancer Research

Series Editors
Jeanne Kowalski
John Hopkins University, Baltimore, MD, USA
Steven Piantadosi
Cedars Sinai Medical Center, Los Angeles, CA, USA
Tuan Pham
Editor

Computational Biology
Issues and Applications in Oncology

Springer
Computational biology is an interdisciplinary research that applies approaches and methodologies of information sciences and engineering to address complex problems in biology. With rapid developments in the omics and computer technologies over the past decade, computational biology has been evolving to cover a much wider research domain and applications in order to adequately address challenging problems in systems biology and medicine. This edited book focuses on recent issues and applications of computational biology in oncology. This book contains 11 chapters that cover diverse advanced computational methods applied to oncology in an attempt to find more effective ways for the diagnosis and cure of cancer.

Chapter 1 by Chen and Nguyen addresses an analysis of cancer genomics data using partial least squares weights for identifying relevant genes, which are useful for follow-up validations. In Chap. 2, Zhao and Yan report an interesting biclustering method for microarray data analysis, which can handle the case when only a subset of genes coregulates under a subset of conditions and appears to be a novel technique for classifying cancer tissues. As another computational method for microarray data analysis, the work by Lê Cao and McLachlan in Chap. 3 discusses the difficulties encountered when dealing with microarray data subjected to selection bias, multiclass, and unbalanced problems, which can be overcome by careful selection of gene expression profiles. Novel methods presented in these chapters can be applied for developing diagnostic tests and therapeutic treatments for cancer patients.

Ductal carcinoma in situ (DCIS) is known as the earliest possible clinical diagnosis of breast cancer and performed with screening mammography that has detected small areas of calcification in the breast. Chapter 4 by Macklin et al. presents a biophysics- and agent-based cellular model of DCIS, which is modular in nature and can be extended to incorporate more advanced biological hypotheses. Chapter 5 by Verma gives an overview of a state of the art for the classification of suspicious areas in digital mammograms and presents a multicluster class-based approach for classifying such areas in benign and malignant cases.

The work in Chap. 6 by To and Pham presents several methods using evolutionary computation algorithms for classification of oncology data. Evolutionary computation is effective in this study because it can offer efficiency in searching in high-dimension space, particularly in nonlinear optimization and hard optimization.
problems. In Chap. 7, Solé et al. provide a thorough review of genetic association studies on SNP-array analysis techniques as well as many existing bioinformatics tools for carrying out such analyses.

Image analysis of cancer cells and tumors is an important research area in computational biology and bioinformatics. Image-based study enables an efficient way for gaining better understanding of the angiogenesis and genetic basis of cancer. Vallotton and Soon report in Chap. 8 several image analysis tools and techniques for the automatic identification of objects in image sequences and quantification of their dynamic behaviors. Chapter 9 by Tran and Pham reports several recent developments in cell classification for high-content screening, which can be useful computerized tools for automated analysis of large image data and for aiding the process of drug discovery.

Chapter 10 by Daskalaki et al. reviews important cellular processes for cancer onset, progression, and response to anticancer drugs and provides a summary of existing databases and tools for computational models in oncology. This chapter also presents several frameworks for modeling cancer-related signaling pathways. Finally, Le et al. discuss in Chap. 11 laser speckle imaging for real-time monitoring of blood flows and vascular perfusion; with proper experimental setups and quantitative analyses, this technology can offer its potential applications for research and development in diagnostic radiology and oncology.

Besides the availability of genomic data, life-science researchers study proteomics to gain insight into the functions of cells by learning how proteins are expressed, processed, recycled, and their localization in cells. Functional proteomics involves the use of mass spectrometry data to study the regulation, timing, and location of protein expression. Interaction studies seek to understand how proteins pair between themselves and other cellular components interact to constitute more complex models of the molecular machines. Chapter 12 by Jin et al. gives an overview of bioinformatics algorithms for biomarker discovery, validation, clinical application of proteomic biomarkers, and related biological backgrounds.

Many thanks to all the authors for their timely effort in contributing chapters to this edited book. Gratitude is expressed to Rachel R. Warren, the Editor of Cancer Research, Springer Science, along with Jeanne Kowalski at Sidney Kimmel Cancer Center at Johns Hopkins and Steven Piantadosi at the Cedars-Sinai Medical Center, in acknowledgment of their invitation, encouragement, and support in making this work a valuable contribution to the endeavor of exploring advanced mathematics, statistics, computer science, information technology, and engineering computation for solving challenging problems in oncology.

Canberra, Australia

Tuan Pham
Contents

1 Identification of Relevant Genes from Microarray Experiments based on Partial Least Squares Weights: Application to Cancer Genomics ................................................ 1 Ying Chen and Danh V. Nguyen

2 Geometric Biclustering and Its Applications to Cancer Tissue Classification Based on DNA Microarray Gene Expression Data ......................................................... 19 Hongya Zhao and Hong Yan


4 Agent-Based Modeling of Ductal Carcinoma In Situ: Application to Patient-Specific Breast Cancer Modeling ......................... 77 Paul Macklin, Jahun Kim, Giovanna Tomaiuolo, Mary E. Edgerton, and Vittorio Cristini

5 Multicluster Class-Based Classification for the Diagnosis of Suspicious Areas in Digital Mammograms .................................................. 113 Brijesh Verma

6 Analysis of Cancer Data Using Evolutionary Computation .............. 125 Cuong C. To and Tuan Pham

7 Analysis of Population-Based Genetic Association Studies Applied to Cancer Susceptibility and Prognosis ............................... 149 Xavier Solé, Juan Ramón González, and Víctor Moreno

8 Selected Applications of Graph-Based Tracking Methods for Cancer Research ................................................................. 193 Pascal Vallotton and Lilian Soon
9 Recent Advances in Cell Classification for Cancer Research and Drug Discovery .............................................. 205
Dat T. Tran and Tuan Pham

10 Computational Tools and Resources for Systems Biology Approaches in Cancer ...................................................... 227
Andriani Daskalaki, Christoph Wierling, and Ralf Herwig

11 Laser Speckle Imaging for Blood Flow Analysis ................................................................. 243
Thinh M. Le, J.S. Paul, and S.H. Ong

12 The Challenges in Blood Proteomic Biomarker Discovery ................. 273
Guangxu Jin, Xiaobo Zhou, Honghui Wang, and Stephen T.C. Wong

Index ................................................................................................................................................. 301
Computational Biology
*Issues and Applications in Oncology*

Tuan Pham, *Editor*

*Computational Biology: Issues and Applications in Oncology* provides a comprehensive report on recent techniques and results in computational oncology essential to the knowledge of scientists, engineers, as well as postgraduate students working on the areas of computational biology, bioinformatics, and medical informatics.

With chapters timely prepared and written by experts in the field, this in-depth and up-to-date volume covers advanced statistical methods, heuristic algorithms, cluster analysis, data modeling, image and pattern analysis applied to cancer research. The literature and coverage of a spectrum of key topics in issues and applications in oncology make this a useful resource to computational life-science researchers wishing to enhance the most recent knowledge to facilitate their own investigations.