

Preface

Those who are involved with mapping the genomic landscapes are participating in one of the most exciting frontiers of science. We have the opportunity to reverse engineer the blueprints and the control systems of living organisms. Computational tools are key enablers in the deciphering process. Thus, this book provides an in-depth presentation of some of the important computational biology approaches to genomic sequence analysis. The first part of the book discusses methods for discovering patterns in DNA and RNA. This is followed by the second part that reflects on methods in various ways, including performance, usage and paradigms.

Part I, *Pattern Discovery Methods*, provides a collection of computational methods and tools. Chapter 1, "*Large-Scale Gene Regulatory Motif Discovery with NestedMICA*," presents an algorithmic approach, describes usage of the tool based on the algorithm, and illustrates its usage via a detailed case study. In Chapter 2, "*R'MES: A Tool to Find Motifs with a Significantly Unexpected Frequency in Biological Sequences*," the authors describe a software tool that contains rigorous statistical models of DNA words. "*An Intricate Mosaic of Genomic Patterns at Mid-range Scale*," Chapter 3 of the book, focuses on intricate mosaics found in genomes; a number of specific patterns are identified. The fourth chapter, "*Motif Finding from Chips to ChIPs*," provides a comprehensive survey of methods for the *de novo* discovery of putative over-represented transcription factor binding sites in nucleotide sequences. Part I concludes with a chapter that considers the discovery of RNA structural motifs: "*A New Approach to the Discovery of RNA Structural Elements in the Human Genome*."

The second part, *Performance and Paradigms*, consists of chapters that contemplate the effectiveness of relevant computational biology techniques. Chapter 6, "*Benchmarking of Methods for Motif Discovery in DNA*," presents a variety of metrics for assessing the performance of the class of methods described in Part I. In "*Encyclopedias of DNA Elements for Plant Genomes*," the application of methods is illustrated with case studies. The topic of scalable algorithmic approaches is considered in

Chapter 8, "*Manycore High-Performance Computing in Bioinformatics*." Chapter 9, "*Natural Selection and the Genome*," discusses evolution of genomic sequences and the role that natural selection plays in directing genome evolution. It also provides a conceptual framework for better understanding of the evolutionary implications and insights that are generated through genomic sequence analyses, and emphasizes the critical role of purifying selection.