ANIL AGGRAWAL'S

Internet Journal of Book Reviews

A sister journal of AAIJFMT · Anil Aggrawal's Internet Journal of Forensic Medicine and Toxicology

VOL. 3, NO. 1 · JANUARY – JUNE 2004

Downloaded from www.anilaggrawal.com · ISSN 0972-9674 (Online) · https://www.anilaggrawal.com/br/vol_003_no_001/reviews/ts/page001.html

Technical Books Section

THE ITEM

Published 2002
Pages 542
ISBN-10 0-262-10092-4
Price $55.00

HOW TO CITE

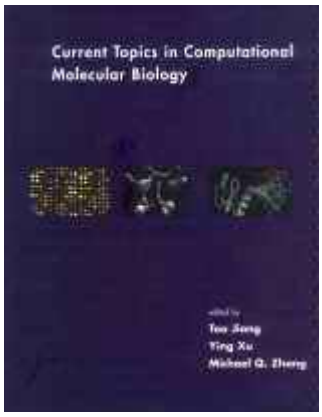
Thomas S. Review of *Current Topics in Computational Molecular Biology*. *Anil Aggrawal's Internet Journal of Book Reviews*. 2004;3(1).

Ref: Thomas S.S. Current Topics in Computational Molecular Biology (Book Review). *Anil Aggrawal's Internet Journal of Book Reviews, 2004*; Vol. 3, No. 1 (January - June 2004): ; Published March 30, 2004, (Accessed:

Email Dr. Thomas by clicking here

PANORAMIC VIEW OF BIOINFORMATICS

Current Topics in Computational Molecular Biology *edited by* Tao Jiang, Ying Xu, Michael Q. Zhang, hard cover, 7"x9"
The MIT Press, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142 (A Bradford Book): xiv + 542 Pages: Publication Date: 2002: ISBN 0-262-10092-4: Price - $55.00



'Science is advanced by new observations and technologies'. This book is one amongst a series of Bradford books on Computational Molecular Biology which adequately tries to cover the rapid developments of large-scale and high-throughput biotechnologies in the field of Bioinformatics (Computational Molecular Biology). It is an up-to-date survey of current topics in this field, which has a growing impact on health and medicine.

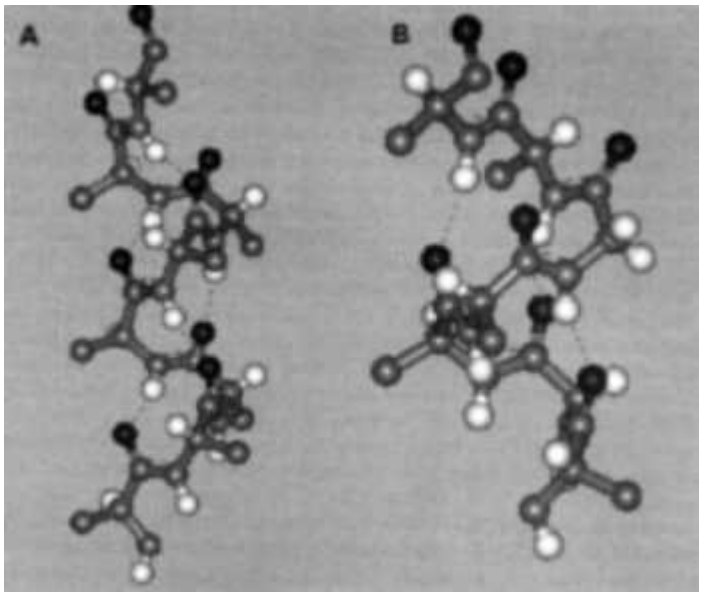
Computational molecular biology draws on the disciplines of biology, mathematics, statistics, physics, chemistry, computer science, and engineering. The Human Genome Project has led to a massive outpouring of genomic data, which has resulted in a revolution that is transforming the entire biomedical research field into a new systems level of genomics, transcriptomics, and proteomics, fundamentally changing how biological science and medical research are done. This revolution would not have been possible without the parallel emergence of the new field of computational molecular biology or bioinformatics, as many people would call it. It is one of the top strategic growing areas throughout academic as well as industrial institutions because of its vital role in genomics and proteomics.

A team of renowned experts who have been actively working at the each forefront of major area of the field have covered most of the important topics in computational molecular biology, ranging from traditional ones such as protein structure modeling and sequence alignment, to the recently emerged ones such as expression data analysis and comparative genomics.

This book is unique as it covers a wide spectrum of topics (including a number of new ones not covered in existing books, such as gene expression analysis and pathway databases) and it combines algorithmic, statistical, database, and AI-bases methods for biological problems. It also contains a general introduction to the field, as well as a chapter or general statistical modeling and computational techniques in molecular biology. Each chapter is a self-contained review of a specific subject. It typically starts with a brief overview of a particular subject, then describes in detail the computational techniques used and the computational results generated, and ends with open challenges. Hence the reader need not read the chapters sequentially.

The book would be useful to a broad readership, including students, nonprofessionals, and bioinformatic experts who want to brush up topics related to their own research areas.

Figure 1: The book is full of illustrative diagrams such as this. This one appearing on page 369 of this book illustrates helices with internal hydrogen bonds in proteins (A)3_{10} - helix; (B) alpha-helix



There are 19 chapters which are grouped into 4 sections -

I) Introduction

II) Computational methods for comparative sequence and genome analyses

III) Computational methods for mining biological data and discovering patterns hidden in the data

IV) Computational approaches for structure prediction and modeling of macromolecules

I) Introduction

This section comprises of comprises of just one chapter, entitled "The Challenges Facing Genomic Informatics" by Temple F. Smith:

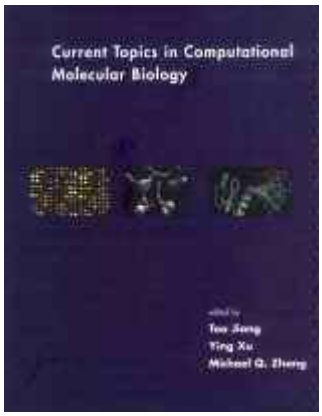
1) The Challenges Facing Genomic Informatics

- Sets bioinformatics into a useful historical context

- Data explosion in terms of the type and sheer volume resulted in the birth of bioinformatics

- This interdisciplinary area provides the data and computational support for functional genomics research domain focused on linking the behavior of cells, organisms, and populations to the information encoded in the genomes.

- This section gives an interesting description of the beginning of modern biology in 1921 at Toronto, with a paper read by Hermann Muler who stated that the gene was a physical particle of complex structure; unique from its product; normally duplicated unchanged, but once mutated, the new form is in turn duplicated faithfully.



" ..This book is unique as it covers a wide spectrum of topics (including a number of new ones not covered in existing books, such as gene expression analysis and pathway databases) and it combines algorithmic, statistical, database, and AI-bases methods for biological problems. It also contains a general introduction to the field, as well as a chapter or general statistical modeling and computational techniques in molecular biology. ... "

- By the mid-1950s motivation was in place to find genes and their products and the beginning of modern molecular biology was on its way.

- The 1st major database of protein sequences was created by mid - 1960s by Margaret Dayhoff (1966).

- By 1982, larger datasets (Gen Bank at Los Alamos) was started and a marriage between sequence analysis and computer science emerged naturally.

- By 1990, nearly all of the comparative sequence analysis methods had been refined and applied many times and the human genome project had been formally initiated.

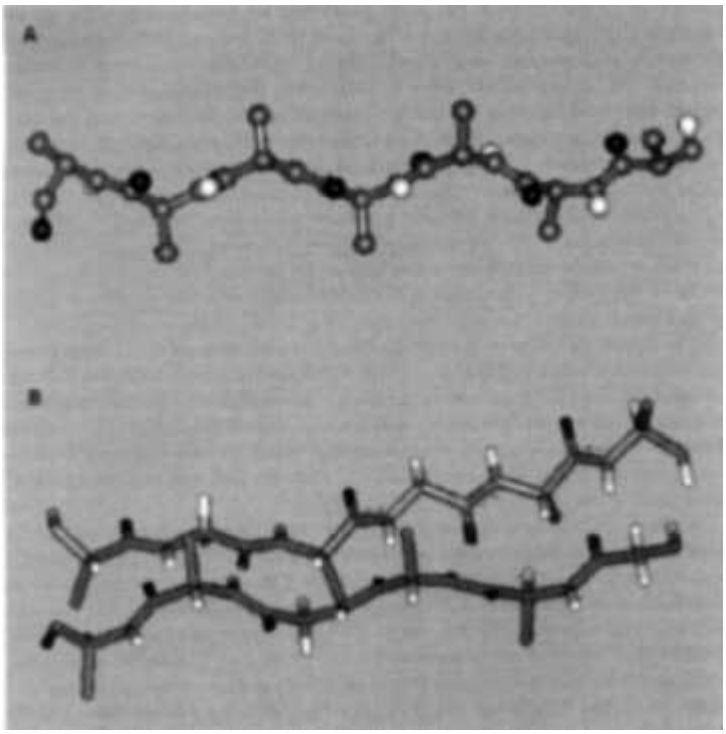
- The current challenge for the biological sciences is to begin to understand how the genome parts list encodes cellular function.

- Now we can generate treatable models or carry out large-scale exploratory experimental tests. The latter, forms the logic behind mRNA expression chips, whereas the former leads to experiments to test new regulatory network or metabolic pathway models. The design, analysis, and refinement of such complex models will surely require new computational approaches.

- The ultimate aim of functional genomics is the extraction of how the organism's range of behavior or environment responses is encoded in the genome. It is the data that has created the latest aspect of the biological revolution whereby we can measure not only a population's genetic variation, but nearly all the genes that might be associated with a particular environmental response.

II) Comparative Sequence and Genome Analysis

Figure 2: This figure appearing on page 372 of this book illustrates beta-structural confirmation (A) Beta-strand geometry (B) Interacting beta strands



This section comprises of following chapters (chapters 2 to 7):

2. Bayesian Modeling and Computation in Bioinformatics Research. Jun S. Liu.

3. Bio-Sequence Comparison and Applications. Xiaoqiu Huang.

4. Algorithmic Methods for Multiple Sequence Alignment. Tao Jiang and Lusheng Wang.

5. Phylogenetics and the Quartet Method. Paul Kearney.

6. Genome Rearrangement. David Sankoff and Nadia EL-Mabrouk.

7. Compressing DNA sequences. Ming Li.

2) Bayesian Modeling and Computation in Bioinformatics Research

Among available computational methods, those that are developed based on explicit statistical methods play an important role in the field and are the main focus of this chapter. The emphasis is on the application of the Bayesian methodology where a comprehensive probabilistic model is employed to describe relationships among various quantities under consideration - data and knowledge, scientific hypotheses and a scaffold.

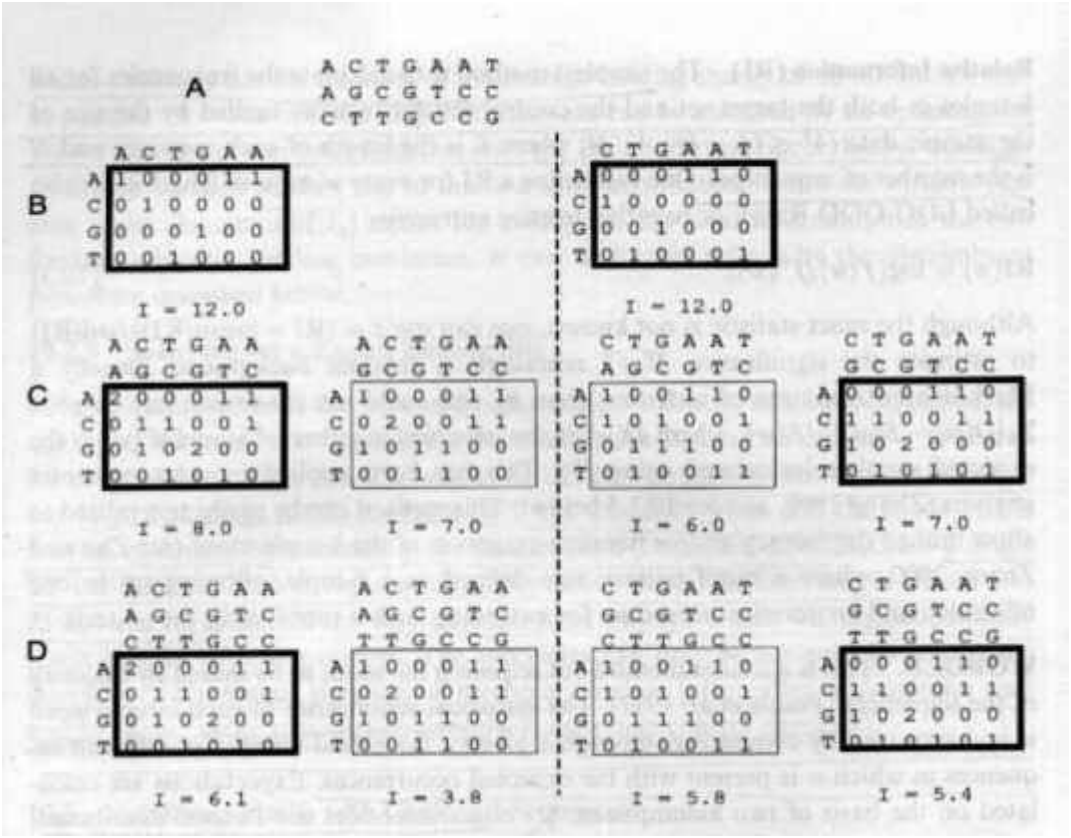
The chapter is organized as follows:

2.1 Introduction

2.2 Importance of formal statistical modeling and overview of two main approaches to statistical inference; the frequentist and Bayesian.

2.3 Bayesian procedure

Figure 3: This figure appearing on page 256 of this book depicts a diagram of the algorithm in CONSENSUS



2.4 Several popular algorithms (EM, Metropolis and Gibbs sampler)

2.5 Use of Bayesian method (to study sequence composition problem).

2.6 Use of Bayesian method to find repetitive motifs in a DNA sequence.

2.7 Discussion.

3) Bio-Sequence Comparison and Applications

The chapter focuses on methods for comparing two sequences which often serve as a basis for multiple sequence comparison methods. The first part describes algorithms for comparing two sequences that are entirely similar (global alignment algorithm and local alignment algorithm) as well as efficient computational techniques for comparing two sets of sequences. In the second part, every sequence in one set is compared with every sequence in the other set with the help of an algorithm. In the third part, usefulness of sequence alignment programs for the analysis of DNA and protein sequences is illustrated. The last part provides two directions for developments of new and improved sequence comparison methods.

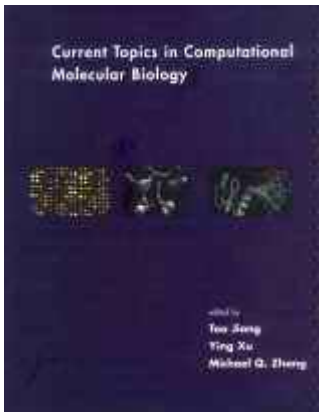
4) Algorithmic Methods for Multiple Sequence Alignment

Algorithms for multiple sequence alignment are routinely used to find conserved regions in biomolecular sequences, to construct family and superfamily representations of sequences, and to reveal evolutionary histories of species (or genes). The profile or consensus sequence obtained from a multiple alignment can be used to characterize a family or super family of species. Computer programs for multiple sequence alignment are becoming critical tools for biological sequence analysis, and can help extract and represent biologically important commonalities (conserved motifs, conserved characters in DNA or protein, common secondary or tertiary structures, etc) from a set of sequences.

5) Phylogenetics and the Quartet Method

This chapter focuses on the development of computational techniques for the analysis of gene sequences evolution by point mutation. Phylogenetics is the design and development of computational and statistical methods for evolutionary analyses. The quartet method is a paradigm for developing phylogenetic methods. A quartet is a set of four sequences, and a quartet topology is an evolutionary tree for a set of four sequences.

Quartet puzzling, introduced by Strimmer and von Haeseler (1996), is currently the most widely used quartet method.



" ..Chapter 5 focuses on the development of computational techniques for the analysis of gene sequences evolution by point mutation. Phylogenetics is the design and development of computational and statistical methods for evolutionary analyses. The quartet method is a paradigm for developing phylogenetic methods. A quartet is a set of four sequences, and a quartet topology is an evolutionary tree for a set of four sequences. Quartet puzzling, introduced by Strimmer and von Haeseler (1996), is currently the most widely used quartet method. ... "

Some phylogenetic resources are also mentioned like Phylip, Tree of Life, Green Plant Phylogeny, Ribosomal RNA Database Project, Tree BASE http://herbaria.harvard.edu/treebase and phylogenetic Resources http://www.ucmp.berkeley.edu/subway/phylogen.html

6) Genome Rearrangement

Its building block are genes, the structures of interest are chromosomes, abstracted in terms of the linear order of the genes they contain. A gene is simply labeled by a number of very different rearrangement processes that are nonlocal, the scope of which may involve an arbitrarily large proportion of a chromosome.

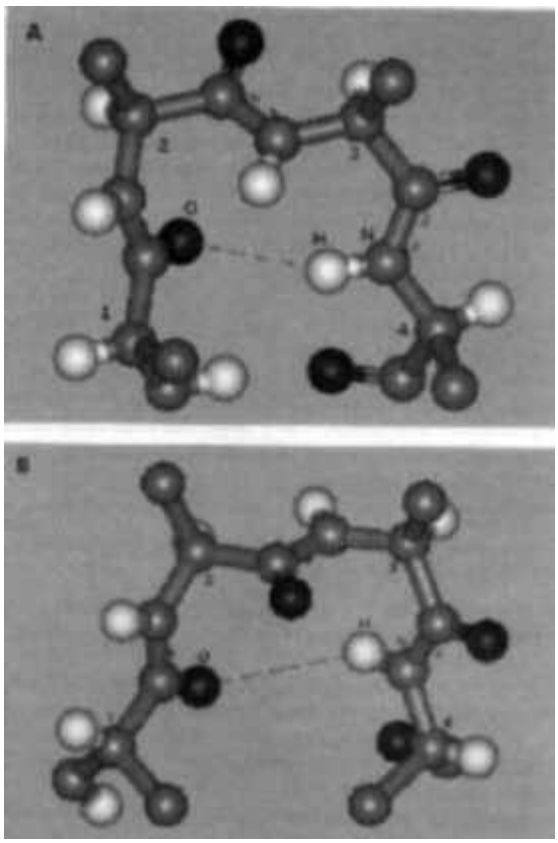
The study of genome rearrangements has focused on inferring the most economical explanation for observed differences in gene orders in two or more species, as represented by their genomes, in terms of a small number of elementary processes.

- Section 6.1 Introduction
- Section 6.2 gives a notion of a genome
- Section 6.3 surveys the problems of genomic distance
- Section 6.4 details the Hannenhalli-Pevzner theory for "signed" distances.
- Section 6.5 is devoted to phylogenetic extensions.

• Section 6.6 is devoted to problems of gene and genome duplication and their implications for genomic distance and genome-based phylogeny.

7) Compressing DNA sequences

Figure 4: This figure appearing on page 375 shows type I and type II of Beta turns. (A) Type 1 beta turn in tenascin (PDB ID 1TEN) at positions 815-818, length of H-bond 2.77 Å. (B) Type II beta turn in mannose permease (PDB ID 1PDO) at positions 84-87, length of H-bond 2.99 Å



Compression is a great tool for genome comparison and for studying various properties of genomes. Different regions on a genome, different genes, different species may have different compression ratios. Such differences may imply, for example, different mutation rates in different genes (Lanctot et al. 2000). DNA Sequence Compression programs can be used to construct whole genome trees e.g. Gen Compress.

III) Data Mining and Pattern Discovery

This section comprises of following chapters (chapters 8 to 13):

8. Linkage Analysis of Quantitative Traits. Shizhong Xu.

9. Finding Genes by Computer: Probabilistic and Discriminative Approaches. Victor V. Solovyev.

10. Computational Methods for Promoter Recognition. Michael Q. Zhang

11. Algorithmic Approaches to Clustering Gene Expression Data. Ron Shamir and Roded Sharan.

12. KEGG for Computational Genomics. Minoru Kanehisa and Susumu Goto.

13. Datamining: Discovering Information from Bio-Data. Limsoon Wong.

8) Linkage Analysis of Quantitative Traits

Quantitative traits/Polygenic traits have a continuous phenotypic distribution (Falconer and Mackey 1996; Lynch and Walsh 1998), such as growth rate of plants. Variances of these traits are often controlled by segregation of many loci. Environmental effects can play a large role in the variation of phenotypic distribution. There are threshold traits (qualitative phenotypically but have a polygenic genetic background) also whose genetic basis is covered by quantitative genetics. Linkage analysis/QTL (quantitative trait loci) mapping is the study of the genetic architecture of quantitative traits using molecular markers (DNA variants). Relative relationship of the markers in the genome (marker map) can be reconstructed using observed recombinant events. Linkage disequilibrium is the foundation for QTL.

Applications of QTL mapping include - improvement of efficiency of selective breeding; application of transgenic technology to quantitative traits; identification of alleles causing predisposition to common multifactorial diseases could lead to improved methods of prevention and knowledge of the number and properties of genes will make the quantitative genetics theory more realistic, which will in turn improve our understanding of evolution (Falconer and Mackey 1996).

Figure 5: This figure appearing on page 236 of this book (in chapter 9 entitled "Finding genes by computer") depicts a Dbscan visualization of results of searching similar regions for a mouse mRNA in a database of known human genes. We can see that all exons shown by red boxes (in the second window) have corresponding conserved regions (in the first window). By this we can detect exon boundaries in the mouse RNA. In third window we can display alignment by marking some conserved region in the first window. [Note by the reviewer: Although the caption of this figure mentions red boxes, the figure appears in black and white in the book. The editors might want to have a look at this discrepancy in the next edition]



---

9) Finding Genes by Computer: Probabilistic and Discriminative Approaches

Computational gene identification is important as a tool of identifying biologically relevant features (Protein coding sequences) that often cannot be found by the traditional sequence database searching technique. This chapter describes statistically based methods for the recognition of eukaryotic genes. The structure and significant characteristics of gene components are reviewed. Recent advances and open problems in gene-finding methodology are discussed along with its application to sequence annotation of long genomic sequences. The application of gene expression

data for large-scale verification of predicted genes is also considered. Details of Web servers for eukaryotic gene and functional signal prediction are also given.

10) Computational Methods for Promoter Recognition

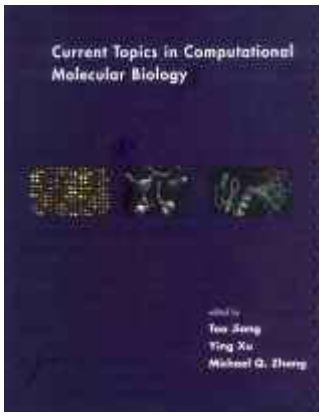
A brief introduction to the biology of promoter structure and function is given in the beginning followed by a review of some of the current computational approaches to the problem, with emphasis on basic concepts and methodologies in real application. Promoter is the most important regulatory DNA region that controls and regulates the very first step of gene expression: mRNA transcription. Delineation of the promoter architecture is fundamental for understanding gene expression patterns, regulation networks, cell specificity, and development. It is also important for designing efficient expression vectors or to target specific delivery systems in gene therapy. In the large-scale genomic sequencing era, promoter prediction is also crucial for gene discovery and annotation.

Computational approaches can be divided into two classes.

1) general promoter recognition method

2) specific promoter recognition methods

Primary goal for the general methods is to identify Transcriptional start sites (TSS) and/or core promoter elements for all genes in a genome; the specific methods focus on identifying specific regulatory elements, Transcription factor binding sites (TF sites) that are shared by a particular set of transcriptionally related genes. Specific methods can have very high specificity when searching against the whole genome and can provide immediate functional clues to the downstream gene. But because of their broad coverage, the general methods are extremely useful for large-scale genome annotation.



" ...Technologies for generating high density arrays of cDNAs and oligonucleotides are developing rapidly, changing the landscape of biological and biomedical research. The information obtained by monitoring gene expression levels in different development stages, tissue types, clinical conditions, and different organisms can help the understanding of gene function and gene networks, and assist in the diagnosis of disease and effects of medical treatments. ... "

11) Algorithmic Approaches to Clustering Gene Expression Data

Technologies for generating high density arrays of cDNAs and oligonucleotides are developing rapidly, changing the landscape of biological and biomedical research. The information obtained by

monitoring gene expression levels in different development stages, tissue types, clinical conditions, and different organisms can help the understanding of gene function and gene networks, and assist in the diagnosis of disease and effects of medical treatments.

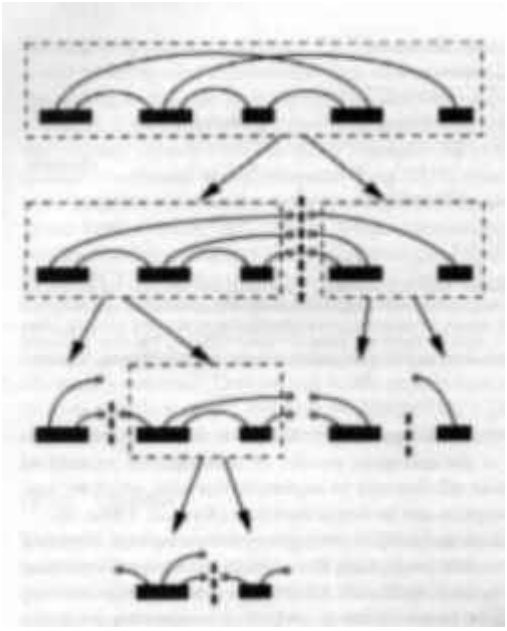
A key step in the analysis of gene expression data is the identification of groups of genes that manifest similar expression patterns. This translates to the algorithmic problem of clustering gene expression data.

A clustering problem consists of elements (usually genes) and a characteristic vector (expression levels of each gene under each of the monitored conditions) and similarity can be measured, for example, by the correlation coefficient between vectors. The goal is to partition the elements into subsets (clusters) on the basis of two criteria - homogeneity and separation. Three technologies that generate large-scale gene expression data are described - cDNA Microarrays, Oligonucleotide Microaarrays and Oligonucleotide Fingerprinting.

12) KEGG for Computational Genomics

Computational molecular biology has been the discipline of choice to analyze sequence and 3D structural information of DNAs, RNAs, and proteins in order to understand molecular functions. Current knowledge on molecular pathways and complexes have been computerized in the PATHWAY database, and analysis of possible relations to the gene catalogs of all the completely sequenced genomes and some partial genomes that are stored in the GENES database in KEGG (Kyoto Encyclopedia of Genes and Genomes) has also been done. KEGG is a computational resource for analyzing networks of symbols at different levels e.g. C for carbon in protein 3D structure at atomic level or to symbols at the molecular level, such as C for cysteine in the amino acid sequence.

Figure 6: This figure appearing on page 480 of this book depicts a partition for a template structure. The first row shows the template with five core secondary structures. The second row shows a partition of the template into two substructures, one with three cores and the other with two cores. A broken arc ended with a circle is called an "open link". The third and fourth rows show further partition of the template until each substructure contains only one core secondary structure. Note that a partition forms a tree structure as indicated by the arrows.



13) Datamining: Discovering Information from Bio-Data

Datamining has attracted increasing attention in the biomedical industry in recent years due to increased availability of a huge amount of biomedical data and the imminent need to turn such data into useful information and knowledge. The knowledge gained can lead to improved drug targets, improved diagnostics, and improved treatment plans.

Datamining is the task of discovering patterns from large amounts of potentially noisy data where the data can be kept in regular relation databases or other forms of information repositories such as the flat text files commonly used by biologists. It is an interdisciplinary subject, relying on ideas and developments in database systems, statistics, machine learning, data visualization, neural networks, pattern recognition, signal processing, and so on.

Section 13.1 Introduction

Section 13.2 presents more background on data mining, where a description of key steps of the knowledge discovery process the diverse functionalities of data mining, and some popular data mining techniques is given.

Datamining has many functionalities, such as association analysis, classification, prediction clustering, and trend analysis. The material in this chapter is presented from the classification perspective, where emphasis is placed on basic techniques for uncovering interesting factors that differentiate one class of samples from a second class of samples. Specifically, the chapter describes data mining techniques for the classification of MHC binding peptides and diabetes clinical study data.

Section 13.3 describes the classification of MHC-binding peptides. It is a target discovery problem in computational immunology. It is an illustration of the application of an artificial neural network to the classification of noisy homogenous biomedical data.

Section 13.4 describes the classification of diabetes clinical study data. It is an illustration of the application of emerging patterns to the classification of heterogeneous biomedical data.

IV. Computational Structural Biology

This section comprises of the remaining chapters (chapters 14 to 19):

14. RNA Secondary Structure Prediction. Zuozhi Wang and Kaizhong Zhang.

15. Properties and Prediction of Protein Secondary Structure. Victor V. Solovyev and Ilya N. Shindyalov.

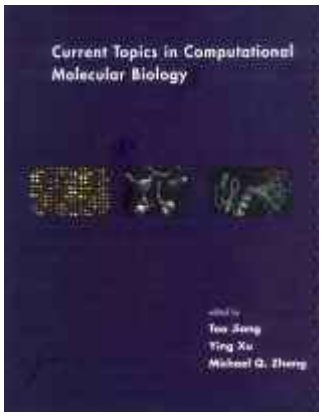
16. Computational Methods for Protein Folding: Scaling a Hierarchy of Complexities. Hue Sun Chan, Huseyin Kaya, and Seishi Shimizu.

17. Protein Structure Prediction by Comparison: Homology-Based Modeling. Manuel C. Peitsch, Torsten Schwede, Alexander Diemand, and Nicolas Guex.

18. Protein Structure Prediction by Protein Threading and Partial Experimental Data. Ying Xu and Dong Xu.

19. Computational Methods for Docking and Applications to Drug Design: Functional Epitopes and Combinatorial Libraries. Ruth Nussinov, Buyong Ma, and Haim J. Wolfson.

14) RNA Secondary Structure Prediction



" ...RNA has recently become the center of much attention because of its catalytic properties, leading to an increased interest in obtaining structural information. Computational methods facilitate the future study of RNA structures, although sometimes they only provide an approximate RNA situational model. Chapter 14 gives a discussion of algorithms/methods (phylogenetic comparative methods, thermodynamic energy minimization method, stochastic context-free grammas method, thermodynamic energy minimization method, stochastic context-free grammar method, equilibrium partition function method, genetic algorithms, etc) to predict RNA secondary structure. ... "

RNA has recently become the center of much attention because of its catalytic properties (Cech and Bass 1988), leading to an increased interest in obtaining structural information. Computational methods facilitate the future study of RNA structures, although sometimes they only provide an approximate RNA situational model. This chapter gives a discussion of algorithms/methods (phylogenetic comparative methods, thermodynamic energy minimization method, stochastic context-free grammas method, thermodynamic energy minimization method, stochastic context-free grammar method, equilibrium partition function method, genetic algorithms, etc) to predict RNA secondary structure.

15) Properties and Prediction of Protein Secondary Structure

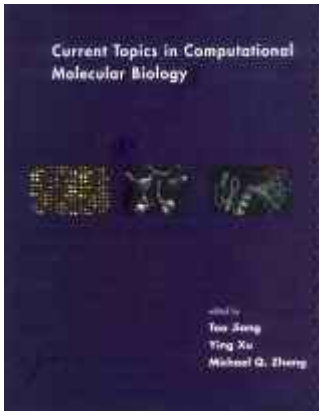
Secondary structure describes regular features of the main chain of protein molecules. Computational prediction of the secondary structure from the amino acid sequence alone is an important step toward our understanding of protein structure and function. It may provide a starting point for tertiary structure modeling, especially in the absence of a suitable homologues template structure, reducing the search space in simulation of protein folding. The predictions can also be used in various aspects of molecular biology research to provide clues about the functional properties of proteins under analysis.

The goal in secondary structure prediction approaches is to extract the maximum information from the primary sequence in the absence of a tertiary structure (King and Sternberg 1996). This chapter deals with the description of secondary structure characteristics, assignment of secondary structure using known 3D coordinates, and prediction of secondary structure based on primary sequence.

16) Computational Methods for Protein Folding: Scaling a Hierarchy of Complexities

Protein are a diverse class of biomolecules performing vital functions in all living things. Therefore, developing a fundamental understanding of how proteins fold is immense intellectual and technological significance. Details of recent protein structure prediction techniques like comparative or homology modeling, fold recognition, and "ab initio" approaches, have been given in this chapter. With the advent of experimental structural genomics, comparative modeling has become increasingly important for providing structural and functional insight into the world's rapidly expanding sequence databases.

17) Protein structure prediction by comparison: Homology - Based Modeling



" ...Understanding the function and the physiological role of proteins is a basic requirement for the discovery of novel medicines (small molecules) and "biological" (protein-based products) with medical, industrial or commodity applications. Functional analysis, the major step after genome sequencing and gene identification, must rely on a combination of technologies. Consequently, new experimental approaches, and their automation for large-scale applications, and their automation for large-scale applications, will need development. Concurrently, and in order to maximize the value of large data sets, one will witness the development of new data mining methods and mathematical models for biological processes simulation ... "

Understanding the function and the physiological role of proteins is a basic requirement for the discovery of novel medicines (small molecules) and "biological" (protein-based products) with medical, industrial or commodity applications. Functional analysis, the major step after genome sequencing and gene identification, must rely on a combination of technologies. Consequently, new experimental approaches, and their automation for large-scale applications, and their automation for large-scale applications, will need development. Concurrently, and in order to maximize the value of large data sets, one will witness the development of new data mining methods and mathematical models for biological processes simulation.

Comparative protein modeling (also often called modeling by homology), consists of the extrapolation of the structure for a new (target) sequence from the known 3D structure of related family members (templates), (Bajorath et al. 1993). Over the next years the focus will be on two main aspects of comparative protein modeling:

1) Improving the sensitivity of the template identification and selection procedure in the sequence similarity "twilight zone", and

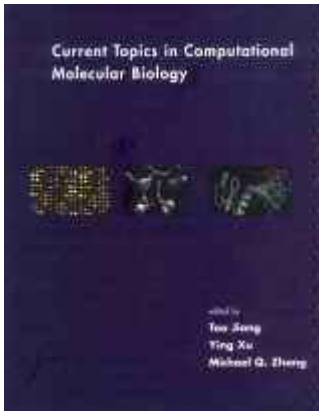
2) Improving the model accuracy

18) Protein structure Prediction by Protein Threading and Partial Experimental Data

As we more toward the post-genome era, it is expected that the demand for rapid protein structure determination will grow drastically Traditional experimental methods for protein structure determination alone, (X-ray crystallography and NMR, will probably not be able to keep up with the pace at which protein sequences are being generated. Computational methods would play a significant role, in conjunction with experimental methods, in protein structure determination on a genome-scale. This chapter gives details of two main classes of tertiary structure prediction, ab initio structure prediction (based on physi-chemical principles directly) and template based structure prediction (using known 3D structures in Protein Data Bank (PDB). Threading uses information related to sequence and structure information such as residue-residue contact patterns in a structure for identifying a homolog or analog through aligning the query protein sequence onto template structures.

A new trend in structure prediction is to incorporate partial experimental data as constraints in the computation process, blurring the boundary between structure prediction and determination. The program PROSPECT is used as an example to illustrate the basic ideas of threading methods. The last part of this chapter, discusses challenging issues and future outlook of the threading method. The cited literature and the Web pages listed in the appendix give more information to the reader.

19) Computational Methods for Docking and Applications to Drug Design Functional Epitopes and Combinational Libraries



" ...The topics covered in this book bring the reader close to the understanding of the function and the physiological role of the gene products in this age, when the draft sequence of the complete genome is ready. I am sure that an interested reader will enjoy his/her journey through bioinformatics like me and through this book the recent developments would also be more accessible to the reader. This book provides a broad-stroke panoramic view of bioinformatics. ... "

Several ingredients are needed in order to efficiently and successfully search a library of inhibitors, or drugs, with the goal of optimally docking them onto a specific target receptor: first, an adequate molecular surface representation; second, efficient docking techniques; third, a practical way of accounting for molecular surface variability; and fourth, providing for molecular flexibility. These four ingredients yield the candidate molecular. The fifth critical component is a fast empirical way of scoring the large number of obtained solutions and ranking them. Currently, although there exist

a variety of computational docking approaches, the scoring step has proven to be the most difficult hurdle.

A description of two computational techniques - rigid body docking technique and hinge-bending motions is given.

The chapter is divided into two parts. In the first, a description of computer vision based rigid and hinge-bending docking algorithms, and the generation of potential solutions is given. The second part focuses on the generation of functional epitopes and some attributes of binding epitopes that have been recently obtained.

The topics covered in this book bring the reader close to the understanding of the function and the physiological role of the gene products in this age, when the draft sequence of the complete genome is ready. I am sure that an interested reader will enjoy his/her journey through bioinformatics like me and through this book the recent developments would also be more accessible to the reader. This book provides a broad-stroke panoramic view of bioinformatics.



-Sherin S. Thomas

Sherin S. Thomas is currently working as a Senior Resident in the Department of Biochemistry at the Maulana Azad Medical College, New Delhi. She completed her MBBS (graduation in medicine and surgery) in 1994 and completed her specialization (MD in Biochemistry) in 2001.

Dr. Thomas is an avid reader of books, journals etc. on a wide variety of topics, especially on Brain theory and neural networks and Computational Molecular Biology. She is a passionate book lover. Her other interests include listening to music and spending time with her five cats and three dogs.

Order this Book by clicking here

-Editor-in-Chief

of Vol. 3, No. 1 (January - June 2004)