

BOOK REVIEW

The expanding scope of bioinformatics: sequence analysis and beyond

Bioinformatics – From Genomes to Drugs (2 vols)

Thomas Lengauer
Wiley-VCH, Weinheim, Germany. 2001; 648 pp.
£170.00, hardback. ISBN 3-527-29988-2

Reviewed by RB Altman

Heredity (2003) 90, 345. doi:10.1038/sj.hdy.6800225

The precise definition of bioinformatics has always been a matter of some debate. Although some use a narrow definition which limits it to the analysis of genome sequences (and more recently microarray analysis), it is probably wiser to be inclusive, and include a wide variety of issues relating to the analysis of nucleic acid and protein sequences, structure, expression and networked function. This is the approach taken by Lengauer and co-authors in this impressive two-volume set that offers a useful annotated introduction to the primary molecular biological challenges and algorithmic approaches currently facing the field of bioinformatics. An additional advantage of this collection is that there is a bias towards the use of technologies toward drug development, which helps to explain some of the choices in content that are made. The relevance of informatics technologies to drug discovery is often left implicit, and so the authors provide a service by trying to make this connection explicit. The two-volume set is divided logically into the first entitled 'Basic technologies', which reviews the general landscape of bioinformatics, and then algorithms for sequence alignment, gene identification, characterization of regulatory regions, modelling protein structures, predicting structures, and docking structures. The second volume entitled 'Applications' introduces the problems of data-base integration, support for genome sequencing efforts, sequence variation analysis, proteome analysis,

finding drug targets, and screening drugs. Taken together, these two volumes provide a very good set of descriptions and references for the interested reader who wants to enter these fields. The quality of the writing is clear and concise, although the great advances in bioinformatics over the last 10 years makes it difficult to treat all topics fully. Not surprisingly, therefore, some discussions are remarkably short or absent (for example, there is no discussion of RNA secondary structure, RNA three-dimensional structural modelling, and the discussion of Gibbs Sampling and EM for sequence motif detection is very short), while others reflect the biases of the authors (for example, the discussion of data-base integration focuses chiefly on the SRS technology, and the extended discussion of complexes and docking reflect the research interests of the editors). The other challenge facing the authors is the emergence of entirely new areas that dominate the landscape. Thus, for example, there is very little discussion of the analysis of microarray expression data (clustering, classification, learning genetic networks), despite the fact that these data have taken the bioinformatics world by storm, and even dominated recent meetings. There is little the editors can do to anticipate such swings of interest, except start planning for a revised version. To their credit, however, there is a very cogent introduction to the methods of proteomics, which is widely considered to be the next wave in bioinformatics data-analysis challenges. Despite this somewhat uneven treatment (no perfect treatment exists), the focus of the book on broad coverage, and the attempt to relate technologies to drug discovery (not the only reason for bioinformatics, but a major driving force) makes it a valuable reference book, introduction for those entering the field, and text for courses in computational molecular biology and bioinformatics.

RB Altman
Stanford University, 251 Campus Drive,
MSOB X-215
Stanford, CA 94305-5479, USA
E-mail: russ.altman@stanford.edu