

A point of entry into genomics

Genomes

by Terry A. Brown

John Wiley & Sons, Inc., \$69.95, ISBN 0-471-31618-0, 1999

Reviewed by Peer Bork and Martijn Huynen

European Molecular Biology Laboratory, D-69012 Heidelberg, Germany

Although the complete genomes of viruses and organelles have been sequenced since the late seventies, the genome era is generally regarded to have started on 28 July 1995, with the publication of the genome of the bacterium *Haemophilus influenzae*. Since then, the number of published genomes has been increasing rapidly—at least 30 should be completed by the beginning of the year 2000, including those of at least two multicellular organisms. But where can one find concise information on technologies and strategies when entering the genomics world? How can researchers be trained to convert the terabytes of new data into understanding and public knowledge? *Genomes*, by Terry A. Brown, the first book to focus specifically on the field of genomics, provides a few answers.

Genomes is a well-written, comprehensive textbook on the field of genomics, containing both an overview of experimental and theoretical techniques used to study genomes and nearly present-day knowledge of how genomes function. Although *Genomes* is dense (which is not surprising, given the subject matter), the book is a pleasure to read and easy to digest. It can serve as a textbook for undergraduates and lecturers, and as a reference book for researchers. Much of this dual value is gained from interspersing the main text with “technical notes” and “research updates” that allow more in-depth coverage of separate issues, and offer historical insights into how the knowledge was derived in the first place. But the wide scope of the book can also lead to oversimplification of theoretical

issues. One prominent example is the definition of orthologous genes—one of the key issues in comparative genomics—as ‘homologous genes in the genomes of different organisms’, a definition which is clearly incomplete, if not incorrect.

Genomes is divided into three parts. The first covers the full scale of genome analysis techniques in detail, from various types of genetic mapping and linkage analyses to shotgun sequencing and even basic sequencing principles. Here the author takes a top-down approach, introducing the large-scale analysis techniques before the actual sequencing itself. Compared with the analysis of genome sequences, the analysis of genome expression is under-represented; for example, we could not find any description of techniques such as differential display, serial analysis of gene expression (SAGE) or representational difference analysis (RDA) in the entire book.

The second part of *Genomes* is a generally up-to-date and brief overview of the molecular biology of the genome and gene expression, although some details appear incorrect or incomplete. For example, in the description of splicing, the role of the SF1 factor is missing. With respect to this subject, there are more extensive contemporary works, such as the frequently updated *Genes* series of textbooks by Benjamin Lewin, the latest edition of which (*Genes VII*) will debut this month.

The third and last part of the book concerns the replication and evolution of genomes, a logical but heretofore unseen

combination of topics within a textbook. Once again, the strength of the book is in the linkage of experimental results with present-day knowledge of genome replication. Relative to this, the theoretical analysis of genomes and their evolution is weak and outdated (perhaps yet another result of the current shortage of bioinformatics experts). Homology searches add tremendous value to sequence data (but within seconds via the World-Wide Web, not hours via e-mail as quoted in the book and as was the practice in the early nineties). These and other computational techniques are essential, given the growing fraction of experimentally uncharacterized genes and proteins. Moreover, computational genomics has already evolved to deal with more complex questions than the homology between sequences, such as the evolution or reconstruction of biochemical pathways and the physical interaction of proteins based on the localization of their genes in the genome.

Genomes introduces numerous important technologies, and the author does a good job of trying to incorporate ‘the state of the art’ into many different areas. But given the rate at which this field is advancing, any book that attempts to cover it will be outdated quickly. The sequence data will be amended with large amounts of heterogeneous biological data from other high-throughput techniques. Even now, genomics research faces significant challenges—not only in generating and analysing large amounts of data, but also in integrating them. Genomics is more than just looking at many genes at once. One possible solution might be a companion volume on the World-Wide Web (several of the pointers to web pages were already dated when the printed version appeared) that is updated regularly. The challenges inherent in describing this field resemble those in genomics research itself: new ways of handling the increasing and changing information are required.

This does not reduce the value of *Genomes*, which is the first textbook to address the step from genes to genomes. It currently has no printed alternative, if a brief but broad and intelligible introduction into the genomics world, and specifically into the techniques that are being used to study genomes at various levels of resolution, is what one desires. □

