Book reviews

Discovering Genomics, Proteomics and Bioinformatics
Edited by A. Campbell and L. Heyer
Pearson Education, Harlow; 2003;
ISBN 0-8053-4722-4; 352 pp.;
£35.37; Paperback

The advent of rapid DNA sequencing techniques has allowed us to accumulate genomic information at an unprecedented rate; the determination of viral and other small genomes is now commonplace and the human genome sequencing project is nearing completion. Importantly, the field of bioinformatics has emerged to supply us with the tools to store, handle and interpret the enormous quantity of data now available. As a result, biology students are now able to think in more global terms, towards entire genomes and the proteins which they encode.

Discovering Genomics, Proteomics, and Bioinformatics, edited by A. Malcolm Campbell and Laurie J. Heyer, introduces students to this new perspective, by combining a text book with an interactive teaching resource on the web (www.geneticsplace.com). A free CD-ROM accompanies the text and features all figures and tables in colour. The book is composed of four units and takes the reader from DNA sequence acquisition to the analysis of microarray data and the concept of genomic circuits and protein interaction networks. Case studies are used for illustrative purposes throughout, and exemplify the role of the genomic sciences in society. Chapters are punctuated with ‘math minutes’, which provide mathematical background to the biology, and ‘discovery questions’, which encourage students to critique published literature, use online tools and databases and draw their own conclusions from the data presented.

Unit 1 focuses on genome sequences: how they are determined, methods used in their characterisation and the databases which house the information. It describes the concept of homology and how the relatedness of genes can be described in terms of their similarity to each other. The BLAST programs are introduced as a popular similarity search method. Next, it deals with the utility of genome sequences: phylogenetic trees allow us to visualise the ancestral relationships between species; our ability to identify organisms from gene sequence will help us to diagnose existing, and newly emerging, diseases; and the field of biomedical genome research may accelerate the development of new medications. The unit concludes with genomic variation, the existence of single nucleotide polymorphisms (SNPs) and the field of pharmacogenomics, which is concerned with how variations at the genetic level translate to altered efficacy of drugs in the clinic. Some ethical implications are considered, such as the consequences of genetic testing and the use of genetically modified organisms (GMOs).

Unit 2 deals with genome expression, and devotes much of its time to DNA microarrays. The impact of DNA chips on health care, particularly cancer, is discussed and numerous case studies demonstrate the enormous potential of this technology. The remainder of the unit comprises several sections on proteomics — addressing protein structure, function and the interaction networks in which they take part. Some of the quantitative and qualitative methods available for proteome analysis are described, from traditional two-dimensional electrophoresis-based approaches to state-of-the-art techniques that utilise mass spectrometry and the emergence of protein chips.

Unit 3 is concerned with genomic circuits, building from what we know about a well-characterised single gene, sea urchin Endo16, to integrated multi-gene networks. It takes the reader from a single gene perspective towards the idea of a
cellular ‘web’, in which genes interact and regulate each other’s transcription. The concept of genetic toggle switches is introduced; these allow the genome to make ‘decisions’ and regulate its own activity in response to a changing environment. The emergence of systems biology is touched upon, showing how models of interaction can be built which closely mimic experimental systems.

The majority of Unit 4 is presented as case studies, which bring together various elements of the book into a set of problem solving exercises. The book concludes with some of the challenges which lie ahead in our ability to dissect and treat disease at the genetic level. Genomic diversity amongst individuals and the polygenic nature of most diseases are highlighted as major obstacles in creating treatments that are effective in all cases. The extent of this problem is illustrated by the difficulties experienced in altering the specificity of aspirin, a simple and well-characterised small molecule with a long history of use in health care.

The scope of the book is intentionally broad and covers a great deal of ground in a little over 300 pages. Naturally, there is a trade-off in that some topics are under-addressed. In particular, I felt that the introduction to sequence analysis in Chapter 1 could have been expanded upon. In addition to the pairwise alignment methods described, those which utilise multiple sequence alignments are widely used. Alignments of gene families and domain families reveal the existence of conserved motifs, which provide us with signatures that can be used to characterise a novel protein or nucleotide sequence. A description of some of the more common secondary, or pattern, databases would have been useful to supplement those sections dealing with prediction of gene function from sequence. Also, an explanation of the relationship between GenBank, the European Molecular Biology Laboratory (EMBL) and the DNA DataBank of Japan (DDBJ) would have bridged the sections dealing with sequence acquisition, and databases. More importantly, it would have illustrated the collaborative nature of the nucleotide sequence data resources which help to fuel the genomic sciences.

The companion website provides useful background on some of the more commonly-used analytical methods, a collection of three-dimensional structures to accompany the text, some sample amino acid and nucleotide sequences for cut-and-paste submission to online tools, and some links to online databases and bioinformatics resources. The CD-ROM proves particularly useful for the microarray images, where spot patterns can be visualised in full colour. Overall, the general approach of the book sets it apart from related texts; the writing style is relaxed and the emphasis on published literature and current research seems appropriate for a discipline in its infancy. The companion website allows students to inspect data firsthand and equips them to perform basic genomic analysis. Besides biology students, this would be a useful guide for anyone wishing to familiarise themselves with the genomic sciences and the wide range of bioinformatics tools publicly available on the internet.

Paul Bradley,
European Bioinformatics Institute,
Wellcome Trust Genome Campus,
Cambridge, UK

Analyzing Medical Data Using S-plus (Statistics for Biology and Health Series)
Brian Everitt and Sophia Rabe-Hesketh
Springer Verlag GmBH & Co. KG,
Tiergartenstrasse 17, D-69121 Heidelberg, Germany
December 2001
ISBN 0-387 98862 9; pp. 485; £56.00; US$79.95; € 79.95

The analysis of massive amounts of data that has become possible through the new methodologies is still a debated and highly