Genet. Res., Camb. (2003), **82**, *pp.* 223–225.
Printed in the United Kingdom

Book Reviews

DOI: 10.1017/S0016672303216608

Bioinformatics for Geneticists. Edited by M. R. Barnes and I. C. Grey. John Wiley and Sons, Chichester, UK. 2003. 350 pages. ISBN 0 470 84394 2. Price £45.00 (paperback). ISBN 0 470 84393 4. Price £90.00 (hardback).

Every practicing biologist now has to understand the jargon of genomics and bioinformatics, and workers in molecular genetics must be able to surf the web of genome data efficiently and effectively. This volume thus promises much: a primer on bioinformatics for, perhaps, the classical geneticist wishing to retrain, or the advanced student entering the field. Unfortunately 'Bioinformatics for Geneticists' is misnamed: the word 'Human' is crucially missing from the title. There is no discussion of other genomes (such as bacterial or model invertebrate) and no lead on how a non-genome-model organism might benefit from bioinformatic/genomic approaches. Many of the tools (and biological-experimental problems) are peculiar to the human genome, and researchers with a focus on other organisms might be better served by one of the other, more general introductory bioinformatics books now available. Bioinformatics for Geneticists reviews the features and landmarks of the human genome, and provides a mixed set of beginners' courses in navigating the bioinformatic tools available (at the end of 2002 in most cases) to the research scientist.

One of the huge strengths of the human genome bioinformatics field is that it is possible, using the 'golden path' human genome sequence, to integrate disparate sources of annotation on one shared framework. Thus one goal of any human geneticist must be to build the skills to allow them to browse and understand these data 'tracks' presented in the genome context. This means not only understanding what the ticks, boxes and colours that appear on the screen mean, but also learning how to navigate between views, ask sensible queries of the system, and how to interpret the output. Interpretation properly also requires an understanding of the 'quality' of the information. For example, if a transcription factor-binding site is shown, how much information does this convey? Are the binding sites on DNA for the transcription factor very carefully and exactly defined, or is it a match to a very 'fuzzy' pattern? How often would such a pattern be expected to appear by chance in a genome segment of this size and base composition? These skills should be being taught at undergraduate level courses now, and are essential catch-up skills for graduate students, postdoctoral fellows and clinical researchers who were educated 'pre-genome'. As the human genome moves from the draft of the 'golden path' to the completed chromosomes being published in the last 2 years, researchers will also have to learn how to keep track of changing base numbering, contig ordering and gene predictions: how to re-find that really interesting gene that was on the last assembly release but appears to have disappeared from today's view.

As the breadth of data and the depth of annotation of the human genome increases, researchers will also have to be able to navigate the variation in the genome: both natural population variation in the form of single nucleotide polymorphisms and more complex variation defined as haplotypes, mutations, rearrangements and 'proper' complex genetic loci.

These annotation sources can be supplied centrally by the large genome centres, but, excitingly, the concept of distributed annotation allows any researcher with a new view, new algorithm, or particular bias, to view their own analyses in the context of all other available ones, and indeed to 'publish' them on the web for others to use. Thus someone with an interest in, say, polynucleotide repeats in relation to gene start sites can generate and display this information for the human genome, and having discovered some utility in the view, publish it for others to comment on. This second aspect of human genome bioinformatics, the generation of new analyses, is a second but distinct need for bioinformatics in human genetics: gaining the skills in data manipulation and analysis to be able to generate new data from the underlying sequence. These skills will be needed by a much smaller community than those of merely consuming data intelligently.

The book is divided into five sections covering introductions to the human genome sequence as

available over the world wide web, the generation, verification and interpretation of sequence annotations, the development of tests based on sequence variation, and the next steps into the realms of regulatory network and protein functional prediction. While the majority focus of the chapters is on navigating the genome and understanding the offered analyses (bioinformatics type one above) some attempt to provide routes to research bioinformatics which, given the space allocated, they fail. The chapters written by Barnes (and co-authors) are coordinated with each other, and aimed at a reader who appears to be new to the human genome, and perhaps to computers and the www. This results in rather simplistic discussions of how to search web-available databases, and a lack of criticism of these resources. Barnes very helpfully uses a single genomic region (the segment around the BACE gene) for many of the analyses and problems presented through the early chapters. This leads the reader logically through both gene-centric and sequence-centric views and emphasises the power of linking many data types in one view of a region. However, other sections are either too general (a list of the biochemical properties of amino acid taking 25 pages) or far too brief, and specific to a particular software solution, to be of general use (for example the chapter on statistical analysis of genetic data, which has little link to the core of the book, the human genomic sequence).

Over one third of the book's 17 chapters are authored by the editors. While this should give it a clarity and conciseness, it remains repetitive and partial. The same software tools are discussed multiple times, sometimes conflictingly, and even some pieces of 'data' are reported differently by different authors. The first chapters' focus on navigating the genome are not effectively carried through in the latter part of the book, and some chapters (for example that on SAGE and microarrays) are too focussed on the authors' particular work to be generally useful.

In closing, the editors describe the promise of the new human genome sequence as a GPS system for genetics compared to the previous 'stars and compass' mapping. The analogy is apposite, but, taking it further, it is instructive to realise that 'public' access to the GPS system is a downgraded version of its actual spatial resolving power: a controlling interest (in this case the (US) military) keeps the best quality data for its own use. Underlying the success of the public human genome project is the vision that all data is available to the community as soon as practical after its generation. It remains a concern that the raw data (sequence, SNPs, marker studies, expression patterns) and the high quality of informatic annotation (gene predictions, functional and structural deductions) remains accessible to all. While Bioinformatics for Geneticists has shortcomings common to multiauthor texts, it does stand as an open way marker to the roads to travel to understanding and exploitation.

MARK BLAXTER
Institute of Cell, Animal and Population Biology
The University of Edinburgh

DOI: 10.1017/S0016672303226604

Plant Biotechnology: The genetic manipulation of plants. A. SLATER, N. SCOTT and M. FOWLER. Oxford University Press. 2003. 346 pages. ISBN 0 19 9254680. Price £19.99 (paperback).

Agricultural biotechnology, especially the generation of transgenic crop plants, is currently one of the most controversial and emotive issues. At least in Europe, the general public are presently overwhelmingly against the introduction of this technology on a commercial scale. This is principally due to perceived food safety issues, the potential environmental impact of these crops and concerns that a handful of agricultural biotechnology companies may establish unparalleled control of the global seed supply.

Into this maelstrom steps 'Plant Biotechnology: The genetic manipulation of plants.' Clearly, there is currently enormous demand from both students and the general public for accurate and unbiased information about this activity. In writing this book, Slater, Scott and Fowler have attempted to provide it. While this has been designed principally as a text book for the teaching of both undergraduate and postgraduate students, the authors have also attempted to make this text accessible to the informed reader. Consequently, the scope of this book is extremely wide and the authors have worked with a particularly broad brush.

The opening four chapters provide basic information on, the organisation and expression of plant genomes, plant tissue culture, procedures for plant transformation and binary vectors. This is generally well written, however, it is rather too complex for a member of the general public, while being, perhaps, too superficial for postgraduate students. For undergraduates though, it is right on the mark. The authors have increased the readability of this book by omitting detailed references from the text. Alternatively, references are presented at the conclusion of each chapter. These are comprised of both original papers and more frequently relevant review articles. The provision of useful web-sites is also a welcome addition.

The remaining chapters of this book address various aspects of plant genetic engineering allied to crop improvement. There are four chapters which provide detailed information on the development of key input traits: herbicide, stress, pest and disease resistance. Crops carrying herbicide or pest resistance have

comprised the first wave of plant biotechnology. However, the development of crops expressing disease or stress resistance is proving to be more problematic. A further chapter covers the manipulation of crop yield, a complex, multigenic trait, which constitutes an even more difficult proposition. All these traits benefit the farmer and not the consumer, which is a point currently not lost on the general public.

In contrast, the second wave of transgenic plants are expected to possess traits that will provide direct benefits to the consumer. For example, the development of plant based vaccines. The development of these so-called 'output traits' is covered during a chapter on molecular 'pharming'. As this represents the future and perhaps the most exciting potential benefits of this technology, I would have liked to have seen more information provided on this area. The final chapter attempts to provide an overview of plant biotechnology, past, present and future, with reference to the legislative framework and economic, social, and ethical issues. Generally, the authors do this rather well and it provides a fitting conclusion for the book. On the debit side, I did notice a number of inaccuracies throughout the book. For example, in the chapter on disease resistance, the NADPH-oxidases thought to be responsible for the production of membrane-impermeable superoxide ($O_2^{\cdot -}$) during the oxidative burst are presumed to be located within the plasma membrane rather than the chloroplasts.

Nevertheless, I would strongly recommend this book for undergraduate students. Moreover, the information covering the cognate social and ethical issues will undoubtedly provide excellent primers for class discussions on these important topics. Finally, students will certainly be delighted to discover such a breadth of information within a single, readable text book.

GARY J. LOAKE
Institute of Cell and Molecular Biology
The University of Edinburgh