Hansen, A.: **Bioinformatik – Ein Leitfaden für Naturwissenschaftler.** – Birkhäuser-Verlag, Basel 2001. ISBN 3-7643-6512-9. SFR 38.00, DM 49.00, öS 358.00.

There are probably many more ways to structure a bioinformatics course than there are people actually teaching the subject, even in this post-genome age. If only a fraction of the teachers decides to turn their lecture-notes into basic introductory books, we still end up with a considerable pile of literature to chose from. Andrea Hansen's thin (112 pages) paperback may be viewed as an example of a course structure that provides useful starting points to students who may later become seriously interested in bioinformatics, while supplying a reasonable amount of factual material to those who only need to pass an exam —and may never touch the subject again.

The book carries unmistakable imprints of its origins in lecture-notes for a practical course taught by the author at the University in Düsseldorf. Great care has been apparently taken to avoid possible overlaps with the contents of other courses. This has substantially contributed to the modest volume (and hopefully also price) of the book, but it also means that the reader needs to be well acquainted with the basic molecular biology and, more importantly, with the questions that will be answered using the bioinformatic tools. Nonetheless, the author admirably avoids any requirements for advanced computer training and, in most cases, also for prior mathematical knowledge, which is consistent with the primary target audience being biology students. To appreciate both advantages and drawbacks of her attitude, one only needs to browse through some contemporary bioinformatics texts of similar size based on a complementary approach, such as Minoru Kanehisa's *Post-genome Informatics* (Oxford University Press, 2000). I, as a biologist, found Kanehisa's book, intended primarily for mathematically literate audience barely touched by molecular biology, far more stimulating. In particular because of the biological passages, which made me think about even more biological questions solvable *in silico*, than I was aware of, although I do admit getting sometimes lost in the mathematics. There is a certain danger that some readers of Hansen's book may end up wondering what is the whole exercise good for, and perhaps shedding any bioinformatics imprints as soon as the exam has been passed. However, if everything were in the books, what would be the purpose of the teacher?

The book covers several selected topics, from a brief account of the current state of sequence databases all over the world, through extensive discussion of sequence alignment, database searching, and sequence comparison, up to construction of phylogenetic trees. All methods discussed can be practically performed using programs that are available for free download or that can be used directly from dedicated free servers, usually from any operating system—a strategy that will be certainly appreciated by a typical less-than-rich student or lecturer. The final chapter and appendices provide a brief survey of specialized databases and web sites, and the author promises maintaining an up-to-date web site to accompany her text.

The strength of the book lies mainly in the thorough, clear and pedagogically exquisite explanation of the principles of methods implemented in programs commonly used for sequence comparison, alignment, and database searching. The author did an admirable work in explaining the basic strategy of the methods without having to include much mathematics, perhaps with a single exception of Hidden Markov Models, where oversimplification might have led to misunderstanding (at least I confess that I did not get it entirely). Importantly, the author manages to convey also the essential point that "programs always produce results, but it is the researcher who has to decide whether these results make any sense".

Nevertheless a contemporary reader might be less than happy about the somewhat arbitrary selection of the topics covered. Although the facts are kept strictly up to date (latest estimates of database size date from October 2000, which is impressive in a book published in February 2001), the overall structure of the text seems to be rooted deep in the pre-genome era (my guess would be mid-nineties). This feeling is supported by the rather curious selection of specialised databases, which devotes a section to enzyme nomenclature and classical enzymological literature databases, while organism-specific genome databases are not considered worthy at least a brief comment.

Especially two topics would have really deserved at least *some* attention. The far from trivial problem of finding potential genes and predicting the location of splice sites and open reading frames in vast chunks of eukaryotic genome data is not mentioned at all. However, nowadays it represents a sizeable proportion of ongoing bioinformatics research, as well as a good entry point into the (likewise neglected) vast area of genome-scale sequence analysis. Second, no account of protein structure and function predictions can be complete without mentioning the resources associated with the EMBL Smart server (smart.embl-heidelberg.de), which provides a simple tool to query all the databases and tools mentioned in the text, and many more, at once. No wonder, if my previous suspicion is correct—this tool has been first introduced only in 1998!

In summary: Hansen's book represents a pedagogically superb basic introduction into some essential methods of bioinformatics. Although the selection of topics is a bit narrow and, above all, strictly pre-genome, the book provides a sound basis for about two thirds of a contemporary introductory bioinformatics course, and as such, it can be recommended to students and teachers alike.

F. CVRČKOVÁ (*Praha*)