

Bioinformatics and molecular medicine - Introduction and call for papers

Molecular medicine is a rather new, rapidly growing field with numerous emerging methodologies and technologies allowing the elucidation of the molecular basis of many diseases. This is facilitated by genetic approaches which pinpoint disease genes and by the genomics of microbial organisms, opening up new horizons for drug development to treat infectious diseases. With the increasing amount of data on the pathophysiology of diseases and the necessity to access various databases worldwide, the demand for informatics support is continuously growing. Bioinformatics is another young discipline which is booming with the influx of information from genome projects and the need to analyze, interpret and integrate heterogeneous data.

Generally speaking, bioinformatics is the application of computational and informatics methods to biological problems. However, due to the fundamentals of molecular biology and molecular medicine, it is becoming its own independent discipline at the interface of computer science, mathematics and biology (for details see *Trends in Biotechnology*, Bioinformatics issue, August 1996). As it is very hard to find general rules without exceptions in biology, bioinformatics requires a deep understanding of the area to which the tools are

applied, because most approaches are knowledge-based.

Thus, most of the successful methods can be called expert systems (or at least the assumptions required to apply a method are made by experts). Whether the methods are purely statistical or are neural networks which are used to overcome statistics, the adaptation of the tools to the problem to be solved is one of the most important issues. Although the complexity of storage, retrieval and analysis of biological data is orders of magnitudes below the intricacies of astrophysics, the limitations imposed by computational aspects must be considered. Furthermore, due to the Internet and the cross-linkage of heterogeneous data resources, modern technologies are required to handle the dynamics of the data and their error margins. Having the newest information available is often an enormous advantage when applying methods in competitive fields, such as molecular medicine, and this is only possible with complicated updating procedures. In addition, the interface to the users, scientists in molecular biology and medicine, has to be customized according to their requirements, which also change rapidly.

Until a few years ago, molecular bioinformatics was a discipline oriented towards basic research. Now it has become an essential part of molecular medicine. Furthermore, it is evolving from an important "service science" to having a role in guiding experimental work. In recognition of this fact and due to its immediate pharmaceutical relevance, molecular bioinformatics is one of the hottest fields in the current life sciences job market [1-3]. Although most research groups already use bioinformatics tools in their daily research, this fast-moving field has a

Peer Bork
EMBL
Meyerhofstrasse 1
D-69117 Heidelberg, Germany
E-mail: bork@embl-heidelberg.de
<http://www.embl-heidelberg.de/~bork/>
and:
Max-Delbrück-Center for Molecular
Medicine (MDC), Berlin-Buch
Robert-Rössle-Strasse 10
D-13122 Berlin, Germany

Bioinformatics: Bits and Bytes



by Peer Bork

much larger potential, and the public awareness of its powers and pitfalls is far from satisfactory. New trends, strategies and practical aspects of bioinformatics develop so fast that we decided to create a column on bioinformatics in molecular medicine.

This column will review the state of the art in several fields of bioinformatics which are relevant to medical research and also provide information on methods and new developments in the World Wide Web. Last but not least, several conclusions relevant for research in molecular medicine can be drawn just by analyzing the data accumulated worldwide. For example, sequence databases can be perused to make functional predictions based on sequence similarities. Therefore, we will include reviews and research contributions regarding the use of databases from experts in the field in an appealing, concise and easy to digest format. Furthermore, the column will become a forum for discussion which mirrors the demands of experimental research groups.

Bioinformatics covers numerous areas from image analysis to aspects of robotics, and we plan to focus on topics which should interest most readers: analysis of genomic data, disease genes and traits, three-dimensional modeling of macromolecules and aspects of computational drug design.

We will start with contributions from experts in various areas of molecular bioinformatics which also provide entry points to the World Wide Web. They include topics such

as the history of sequence analysis, the accessibility of information on disease genes in the SWISS-PROT sequence database, the value of ESTs in disease research and the identification of genes in genomic data. New ideas and discussion points are welcome and we invite the submission of short contributions (2-3 printed pages, and up to 3 figures). Reports of results achieved with bioinformatics methods and with relevance to molecular medicine are also encouraged if their presentation fits the desired format.

References

1. Marshall E (1996) Hot property: biologists who compute. *Science* 272: 1730-32
2. Marshall E (1996) Federal funding. Appropriators bullish on biomedicine. *Science* 272: 1733
3. Timpane J (1996) *Science* 272: 1897-1913

Upcoming contributions:

- ◆ Sequence analysis of GADD45:
Eugene Koonin (NCBI, USA)
- ◆ History of biocomputing:
R. F. Doolittle (UCSD, USA)
- ◆ Use of ESTs for disease gene analysis:
Greg Schuler (NCBI, USA)
- ◆ Gene identification:
Roderic Guigo (IMAS, Spain)
- ◆ SWISS-PROT and disease genes:
Amos Bairoch (University of Geneva, Switzerland)
- ◆ Molecular databases:
Sandor Suhai (DKFZ, Germany)