

# Editorial

## BIOINFORMATICS - CHALLENGES IN 2001

### Bioinformatics - solving biological problems

When the department of Biocomputing was formed in 1986 at the EMBL Heidelberg, my colleague Arthur Lesk insisted on the spelling BIOinformatics. If we could not answer biological questions, we would fail as computational biologists. The same sort of thinking and the desire to focus on modern biology had an effect on the journal CABIOS (Computer Applications in the Biosciences), this journal's direct predecessor, renamed at the end of 1997.

Five years ago, even Bioinformatics moved to reporting biologically driven research rather than simply descriptions of new software. Reports such as 'PRED: yet another implementation of the Chou-Fasman method on Windows 2.6' were no longer accepted and, more generally, software reports shortened to at most 2 pages as 'Application Notes'. Structured abstracts now encourage putting a paper into biological context. 'Discovery Notes' report biologically interesting discoveries using computational techniques, with topics ranging from sequence motif detection, structural similarities, and gene structure prediction to comparative genomics and metabolic pathways. Review papers are actively solicited from all areas of computational biology, not just algorithmics.

Let's look beyond the journal. With the entry of new researchers into the field of bioinformatics, many trained in computer science, we continually need to widen the roads that lead from algorithm and computation to biologically meaningful results. And as genomic technologies invade most wet labs, we have the opportunity, even the obligation, to lead fields of biological and medical research in the design of experiments and of tools used to generate the data on which computation is based. We assert: computational planning and analysis is an integral part of the biological discovery process.

### Leadership - formulating biological questions

More biological thinking, increasing focus on answering biological questions: what does this mean in concrete terms? Perhaps it's as simple as putting the horse before the cart. Don't just talk about the challenge of analyzing massive data in the age of completed genome sequences and high throughput technologies; instead, talk about risk for disease, about human genetic differences, about how evolutionary changes in genotype lead to changes in function and how to use data to answer these questions. Don't just talk about support vector machines as a cool machine-learning tool; talk about classifying cancer

cells for diagnostic purposes and pick the best toolkit. Don't just talk about data mining; talk about defining gene function using expression profiles or about defining sets of functionally related genes that are involved in an important cellular process using a combination of expression profiling and SNP frequencies. Add your own examples.

Scientific culture evolves with new technologies, with new types of data, but usually too slowly. We need leadership: training eclectic physicists and near-nerdy computer scientists in key aspects of biological systems; training wet lab biologists in anticipating the results of computational data analysis and designing their experiments accordingly. Training everyone in thinking about how to derive systems properties using computational tools on organized datasets produced using miniaturized and parallelized technology.

### New ideas - evolving the publication

The journal needs to evolve to make a larger contribution. Yes, we have seen fascinating new papers. But - are all interesting parts of computational biology covered? No. Is an impact factor of about 2.5 spectacular? No. So, your suggestions are invited, creative submissions of things written are welcome.

Here are some suggestions, not yet peer-bounced, of some new sections and new types of papers. New and views: comments on important developments (who would write these?), more immediate and interesting than review papers. Software reviews: short anecdotal accounts of successful uses of software tools - or warnings of where they might go wrong. Failures: reports on serious research projects that failed (imagine the lessons one could learn!). Bioinfo groups: featured bioinformatics centers (new and rejuvenated ones). Design of experiments: let computational biologists think ahead and specify what kind of data need to be collected using current (genomic and high throughput) technologies to answer important biological questions and which new technological tools (reasonably realistic, please) are needed. Ideas and concepts: proposals for investigations, hypotheses to be tested. This latter proposed section (as well as the one on failures) would need careful attention by one of the editors to filter nonsense and nurture novel scientific communication.

### Editorial board - adding critical gray matter

This journal ([bioinformatics.oupjournals.org](http://bioinformatics.oupjournals.org)) is growing: more submissions, more pages published per year, more readers and more authors. With this comes the need to review more and more papers and to improve overall quality. In addition to the superb efforts of Dr Barbara Cox, who runs the editorial office in rural Cambridgeshire, and the current associate editors (Martin Bishop, Christos Ouzounis, Liisa Holm, Russ Altman, Toshihisa Takagi), more critical gray matter is needed.

So, as 2001 dawns, David States and Charlie Hodgman have joined as associate editors, David in the US (otherwise known as an appendix of Texas), Charlie in the UK (otherwise known as part of the EU). David has broad experience ranging from medicine and protein folding to sequence analysis and genetics. David is a professor in the department of genetics and a primary member of the Center for Computational Biology at Washington University in St.Louis and active in ISCB. His group's current projects focus on the application of quantitative analytic methods and machine learning technology to problems of molecular biology and genome analysis. Charlie Hodgman (at the merged Glaxo SmithKline) has a long history of exploring sequence-function relationships. However, over the last 5 years, he has focused on the bioinformatics of molecular interactions and how they combine into the biological networks that underpin anatomy and physiology. His projects extend the scope of BIOinformatics into chemistry and the physiome project. As newly appointed editors, we expect that David and Charlie will make effective and creative contributions to the future of the journal and of the field.

#### **ISCB and Bioinformatics - defining the goals**

Bioinformatics has become the official journal of the International Society for Computational Biology, ISCB.

This is both an opportunity (growth with an exciting community!) and a challenge (satisfy a demanding and intelligent set of customers and keep up with changing conditions!). Oxford University Press, Editors, Authors are you listening?

Talking about the biological orientation of BIOinformatics. The ISCB has got it right: 'The International Society for Computational Biology is dedicated to advancing the scientific understanding of living systems through computation; our emphasis is on the role of computing and informatics in advancing molecular biology' ([www.iscb.org](http://www.iscb.org)).

One fundamental quibble with these goals, though: what do we mean by 'understanding' and, in what is the implied direction of 'advancing'? In my view, understanding in science relates to the ability to predict future outcomes, at least qualitatively, as well as the ability to intervene in those outcomes. Monkeys have probably honed these skills before the arrival of homo scientiae. For biology, this includes the ability to anticipate what may happen to living beings and to optimize their wellness. So, computational biology is challenged to have a strong bridge to medical and environmental care. Let's make sure this journal helps build that bridge.

Chris Sander, Editor