

BY JOHN C. WOOLEY, PhD

Computing Has Changed Biology Forever

In 1991, a prescient editorial in *Nature* by Harvard's **Walter Gilbert, PhD**, ("Towards a paradigm shift in biology") included these observations on the utility and impact of computing:

"The new paradigm now emerging is that all the 'genes' will be known (in the sense of being resident in databases available electronically), and that the starting point of any biological investigator will be theoretical. ... We must hook our individual computers in the worldwide network that gives us access to daily changes in the database. ... The programs that display and analyze the material for us must be improved—and we must learn how to use them more effectively."

Today, Gilbert's paradigm shift at the interface of biology and computation is essentially complete: Genomic data allow researchers to start their investigations through a theoretical approach—an analysis of sequence databases; computer literacy is an essential part of the biologist's toolkit; data resources are available anywhere anytime; and software for biologists is becoming more widely available and useable.

Another emerging idea for biology is third party open access and standards (such as the open development efforts that extended Linux) to establish annotation for the vast amounts of data emerging. No single group will ever be able to annotate the data arising from the ever accelerating pace of genome sequencing, let alone that from metagenomics (Venter et al., 2004). Subsequent automated re-annotation following advances in biological understanding must be a feature of knowledge management. Similarly, the development of sophisticated computational methods for predicting function is needed to refine experiments in functional genomics and make explicit the information flowing from high throughput sequencing (Friedberg et al., 2006).

An exclamation mark for the introduction of computational approaches to biology has recently appeared: the 2005 report on the frontier at the interface, which summarizes a study of the National Research Council (*Catalyzing Inquiry at the Interface of Computing and*

Biology). While the field has grown so much in the past decade that no study could be inclusive, the NRC report is a comprehensive review of the history and current status of the field. The report, pointing also to the future, was conducted independent of the interests of individual agencies and subdisciplines, and should build support for research on the frontier.

Catalyzing Inquiry documents the paradigm shift through reviewing the accomplishments, opportunities and challenges at the interface; provides further sup-

An exclamation mark for the introduction of computational approaches to biology has recently appeared: the 2005 report on the frontier at the interface.

port for the commitment envisioned by Gilbert for advancing software for quantitative biological research, such as the efforts of the National Centers for Biomedical Computation and the computing-based NIH Research Resources; and points to the importance of programs for bioinformatics education for extending the impact of computing.

In 1998, Stanford professor **Michael Levitt, PhD**, reflected that "computing has changed biology forever, even if most biologists don't know it yet." Today, perhaps, they do. □

REFERENCES

Catalyzing Inquiry at the Interface of Computing and Biology, 2005 report of the National Research Council of the National Academies, National Academy Press, Washington, DC. <http://www.doegenomestolife.org/pubs/NRCCComputingandBiology/index.shtml>. Editors: John C. Wooley and Herbert S. Lin.

Friedberg, I., Harder, T., Godzik, A. (2006) "A Protein Function Annotation Metaserver," *Nucleic Acids Research*, in press.

Gilbert, W. (1991) "Towards a paradigm shift in biology." *Nature* 349: 99.

Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rush, D. (2004) "Environmental Genome Shotgun Sequencing of the Sargasso Sea." *Science* 304: 66-74.

DETAILS

John C. Wooley, PhD, is Associate Vice Chancellor of Research at the University of California, San Diego and Senior Fellow of the San Diego Supercomputer Center.