



CONTACT SCIENCE

Letters to the Editor

May be submitted via e-mail (at science_letters@aaas.org), fax (202-789-4669), or regular mail (Science, 1200 New York Avenue, NW, Washington, DC 20005, USA). Letters are not routinely acknowledged. Full addresses, signatures, and daytime phone numbers should be included. Letters should be brief (300 words or less) and may be edited for clarity or space. They may appear in print and/or on the Internet. Letter writers are not consulted before publication.

Subscription Services

For change of address, missing issues, new orders and renewals, and payment questions, please contact AAAS at Danbury, CT: 800-731-4939 or Washington, DC: 202-326-6417, FAX 202-842-1065. Mailing addresses: AAAS, P.O. Box 1811, Danbury, CT 06813 or AAAS Member Services, 1200 New York Avenue, NW, Washington, DC 20005 • Other AAAS Programs: 202-326-6400

Member Benefit Contacts

For Credit Card: MBNA 1-800-847-7378; Car Rentals: Hertz 1-800-654-2200 CDP#343457, Dollar 1-800-800-4000 #AA1115; AAAS Travels: Betchart Expeditions 1-800-252-4910; Life Insurance: Seabury & Smith 1-800-424-9883; Other Benefits: AAAS Member Services 1-202-326-6417.

Reprints

Ordering/Billing/Status 800-407-9190; Corrections 202-326-6501 Permissions 202-326-7074, FAX 202-682-0816

Internet Addresses

science_editors@aaas.org (for general editorial queries); science_news@aaas.org (for news queries); science_letters@aaas.org (for letters to the editor); science_reviews@aaas.org (for returning manuscript reviews); science_bookrevs@aaas.org (for book review queries); science@science-int.co.uk (for the Europe Office); membership@aaas.org (for member services); science_classifieds@aaas.org (for submitting classified advertisements); science_advertising@aaas.org (for product advertising)

Information for Contributors

See pages 108 and 109 of the 2 January 1998 issue or access www.sciencemag.org/misc/con-info.shtml.

Editorial & News Contacts

North America 1200 New York Avenue, NW, Washington, DC 20005 Editorial: 202-326-6501, FAX 202-289-7562 News: 202-326-6500, FAX 202-371-9227 • Bureaus: Berkeley, CA: 510-841-1154, FAX 510-841-6339, San Diego, CA: 760-942-3252, FAX 760-942-4979, Chicago, IL: 312-360-1227, FAX 312-360-0537

Europe Headquarters: Bateman House, 82-88 Hills Road, Cambridge, UK CB2 1LQ; (44) 1223-326500, FAX (44) 1223-326501 Paris Correspondent: (33) 1-49-29-09-01, FAX (33) 1-49-29-09-00

Asia News Bureau: Dennis Normile, (81) 3-3335-9925, FAX (81) 3-3335-4898; dnormile@twics.com • Japan Office: Asca Corporation, Eiko Ishioka, Fusako Tamura, 1-8-13, Hirano-cho, Chuo-ku, Osaka-shi, Osaka, 541 Japan; (81) 6-202-6272, FAX (81) 6-202-6271; asca@os.gulf.or.jp • China Office: Hao Xin, (86) 10-6255-9478; science@public3.bta.net.cn

Staying Afloat on the Seas of Data

Floyd E. Bloom

This issue of *Science* contains a milestone in the life sciences: the first complete genome of a multicellular eukaryote, the nematode *Caenorhabditis elegans*. Those who worked so exhaustively to collect this mass of sequence information deserve the praise of the entire scientific community. Not only did they evolve an effective strategy for large genome analysis with efficiency and accuracy, they also assembled the data in an organized manner that will now permit many other scientists to enjoy and extend the fruits of their labor. And determining the sequence of this genome is only the first stage of a heroic endeavor to understand the meaning of the genes and the interactions of their gene products. Truly, these researchers have generated a massive sea of data.

However, scientists regard data with mixed emotions. They avidly collect, validate, and interpret their own observations to produce the scientific publications on which their careers will be built. Yet for overloaded scientists focused on their own work, genome-sized bursts of data might be skeptically viewed as inaccessible or too massive to study usefully. In fact, this view is false. Providing the data in an organized form suitable for examination by the powerful search, alignment, and similarity-seeking tools that have been made publicly available means that these data can now be examined by any scholar.

The databases that have accrued through the several microbial genome projects as well as the yeast and the *C. elegans* genomes now begin to offer ways to look at other whole genomes from other species that are still quite incomplete in order to look for conservation of the most well-defined gene products. All of these genome databases can synergize constructively with intermediate scale databases on types of molecular function (receptors, transductive mechanisms, DNA transcriptional regulators, and classes of enzymes) or structural motifs representing consensus protein functions. Whole-genome databases are a major advance over the molecule-by-molecule methods that expanded the DNA and protein databases over the past quarter-century. The next decade will bring a host of other complete genome sequences from other organisms and from humans. In short, the scientific community can now not only gain sequence precision from the completion of the *C. elegans* genome but they can also provide data of their own to hasten the task of providing functional annotations that will clarify the role of these gene products in the physiology of the nematode.

Moreover, as important as these whole-genome efforts will be, other technological advances for rapidly analyzing whole-organism gene expression, such as the DNA chip array, will likely provide still more enormous data sets. Comprehensive determination of the genetic responses of specified cells under defined experimental challenges will likely reveal new regulatory systems, especially as tools emerge to refine analysis of their temporal, spatial, and quantitative differences across cells in other organs and in other species.

The growth of databases poses many challenges for scholars and journals. Traditionally, scientific databases have been both public and private. Although many issues of standardization remain to be resolved for the multiple public databases, corrections and annotations require that all databases be constantly updated. Should databases from which observations have been extracted for a scientific report be frozen and made accessible for others to examine for confirmation or alternative insights? Should they be maintained by the scholar or the journal, and for how long and in what form? Would such data sets be an archived laundry list of information or could they provide a new dimension to guide future experimentation?

Science believes that when an experimental database provides a generalized understanding for a broad community and is crucial to the evaluation of insights presented in a paper, it will benefit scholarship for that database to be reviewed, publicly archived, and made accessible in easily imported file structures. Moreover, we look forward to the challenge of reviewing, evaluating, and properly presenting papers based on such large data sets, and to the extended insights that the community may then be able to glean from accessing them. We await your submittals.

The growth of
databases poses
many challenges
for scholars
and journals.