

Sharing Publication-Related Data and Materials: National Research Council Publishes Report

Diego Pineda

A new report from the (US) National Research Council addresses issues of sharing data and materials mentioned in scientific papers in the life sciences. The activities leading to the report originated in the debate on access to the sequence of the human genome, but the report is of considerably broader scope.

The publication of the draft sequence of the human genome in *Science* in February 2001 generated not only excitement in the scientific community but many complaints. "After *Science* published the Celera paper on the human genome, allowing the data to stay on the Celera Web site, [the National Academy of Sciences] got calls from scientists expressing the need to talk about the issue of publication by the for-profit sector and the convention of depositing everything in Genbank", said Robin A. Schoen, of the National Research Council.

Many journals require submission of DNA sequences to public databases like Genbank (www.ncbi.nlm.nih.gov/Genbank) before publication of a paper so that an accession number may appear in the paper. Celera Genomics posted its sequence on its own Web site and restricted the amount of data downloadable from the Web site at any one time. "Many scientists thought that move violated the principle that data integral to a publication (and necessary for another scientist to determine its validity) need to be freely available upon publication", said Thomas R. Cech, president of the Howard Hughes Medical Institute.

Cech chaired the Committee on Responsibilities of Authorship in the Biological Sciences, created in October

DIEGO PINEDA, a graduate student in science and technology journalism at Texas A&M University, is an intern at *Science* Editor.

Principles and Recommendations for Sharing Publication-Related Data and Materials*

The Committee on Responsibilities of Authorship in the Biological Sciences identified five principles that represent what the scientific community commonly expects from authors regarding the sharing of data and materials mentioned in their publications. These are the principles as stated in the report:

Data and Software

Principle 1. Authors should include in their publications the data, algorithms, or other information that is central or integral to the publication—that is, whatever is necessary to support the major claims of the paper and would enable one skilled in the art to verify or replicate the claims.

Principle 2. If central or integral information cannot be included in the publication for practical reasons (for example, because a dataset is too large), it should be made freely (without restriction on its use for research purposes and at no cost) and readily accessible through other means (for example, online). Moreover, when necessary to enable further research, integral information should be made available in a form that enables it to be manipulated, analyzed, and combined with other scientific data.

2001 by the National Research Council to create consensus on authors' responsibilities for sharing data and materials referenced in their publications. To help to analyze the issue, the committee organized a workshop, "Community Standards for Sharing Publication-Related Data and Materials", held on 25 February 2002 in Washington, DC. The resulting report, *Sharing Publication-*

Principle 3. If publicly accessible repositories for data have been agreed on by a community of researchers and are in general use, the relevant data should be deposited in one of these repositories by the time of publication.

Materials

Principle 4. Authors of scientific publications should anticipate which materials integral to their publications are likely to be requested and should state in the "Materials and Methods" section or elsewhere how to obtain them.

Principle 5. If a material integral to a publication is patented, the provider of the material should make the material available under a license for research use.

In addition to the five principles, the committee developed 10 recommendations for the life-science community to discuss and consider. The recommendations include attention to contextual factors, such as costs, administrative barriers, and commercial agreements.

Recommendation 1. The scientific community should continue to be involved in crafting appropriate terms of any legislation that provides additional database protection.

Recommendation 2. It is appropriate for

Related Data and Materials: Responsibilities of Authorship in the Life Sciences, published in spring 2003, synthesizes the discussions at the workshop and the deliberations of the committee. An appendix includes the five hypothetical scenarios examined in the workshop.

"It is hoped that the report stands as a reminder to all about the purpose of the

continued

scientific reviewers of a paper submitted for publication to help identify materials that are integral to the publication and likely to be requested by others and to point out cases in which authors need to provide additional instructions on obtaining them.

Recommendation 3. It is not acceptable for the provider of a publication-related material to demand an exclusive license to commercialize a new substance that a recipient makes with the provider's material or to require collaboration or coauthorship of future publications.

Recommendation 4. The merits of adopting a standard MTA [material transfer agreement] should be examined closely by all institutions engaged in technology transfer, and efforts to streamline the process should be championed at the highest levels of universities, private research centers, and commercial enterprises.

Recommendation 5. As a best practice, participants in the publication process should commit to a limit of 60 days to complete the negotiation of publication-related MTAs and transmit the requested materials or data.

Recommendation 6. Scientific journals should clearly and prominently state (in the instructions for authors and on their Web sites) their policies for distribution of publication-related materials, data, and other information. Policies for sharing materials should include requirements for depositing materials in an appropriate repository. Policies for data sharing should include requirements for deposition of complex datasets in appropriate databases and for the sharing of software and algorithms integral to the findings being reported. The policies should also clearly state the consequences for authors who do not adhere to the policies and

the procedure for registering complaints about noncompliance. Recommendation 7. Sponsors of research and research institutions should clearly and prominently state their policies for distribution of publication-related materials and data by their grant or contract recipients or employees.

Recommendation 8. If an author does not comply with a request for data or materials in a reasonable time period (60 days) and the requestor has contacted the author to determine if extenuating circumstances (travel, sabbatical, or other reasons) may have caused the delay, it is acceptable for the requestor to contact the journal in which the paper was published. If that course of action is not successful in due course (another 30 days), the requestor may reasonably contact the author's university or other institution or the funder of the research in question for assistance. Those entities should have a policy and process in place for responding to such requests for assistance in obtaining publication-related data or materials.

Recommendation 9. Funding organizations should provide the recipients of research grants and contracts with the financial resources needed to support dissemination of publication-related data and materials.

Recommendation 10. Authors who have received data or materials from other investigators should acknowledge such contributions appropriately.

*These principles and recommendations are reprinted with permission from *Sharing Publication-Related Data and Materials*, copyright 2003 by the National Academy of Sciences, courtesy of the National Academies Press, Washington, DC.

publication process and how it benefits the whole community—to make it possible to quickly propel science forward”, said Schoen, study director of the project that included the workshop and the report. “Sometimes, authors don't share materials with colleagues who are competing with them. We hope that this report gives a little bit of ammunition to those who are frustrated by a lack of adequate sharing or access. They can dare to complain, and in doing so, they can point to the community principles laid out in the report.”

The report notes that editors can play an

important role in the sharing of scientific data, mediating between authors and fellow scientists requesting materials from them.

“I think the editor will have to decide what the policy of the journal will be with respect to sharing different kinds of materials and data related to publications”, Schoen said, “and then make sure prospective authors know about the policy and the consequences of not following it. That implies that an editor is willing to explore complaints and make tough decisions about enforcement.” Some journals, including the journals of the American Society for

Microbiology and the Nature Publishing Group, already have such policies.

Cech said the benefits of sharing scientific data are easy to see. “Science is facilitated by sharing materials and databases upon publication”, he said. “The obligation to share is, over time, recompensed by the flow of materials and data to those who share.”

The report can be ordered or read online at www.nap.edu/catalog/10613.html. Printed copies of the report are also available from the Board on Life Sciences, 500 Fifth St NW, Washington DC 20001. 