

# PubMed Central: The GenBank of the published literature

Richard J. Roberts

New England Biolabs, Beverly, MA 01915

In 1999, Harold Varmus, then Director of the National Institutes of Health, proposed a bold new initiative called PubMed Central (PMC) designed to provide a central repository for literature in the life sciences [see Marshall, E. (1999) *Science* **284**, 718]. Following an initial period of confusion, PMC now exists. It has a clear mission, a stable home, and a nucleus of papers. Its mission is to provide a comprehensive electronic archive of the peer-reviewed literature relevant to the biological sciences. Its home is the National Center for Biotechnology Information (NCBI), whose director is David Lipman. NCBI is also home to GenBank, the public archive of DNA sequences. The publications already present in PMC and freely accessible to the world's scientific community include all articles published in PNAS that are more than 1 month old and that were in a suitable electronic format, as well as articles from a number of other journals such as *Molecular Biology of the Cell*, *Arthritis Research*, and *Breast Cancer Research*. Several other journals, including *The British Medical Journal* (BMJ) and *Nucleic Acids Research* (NAR), are committed to join. A full list is available at [www.pubmedcentral.nih.gov](http://www.pubmedcentral.nih.gov).

PMC will contain only articles from the peer-reviewed literature and is not intended to be the sole repository or distributor of the publications that it hosts. In fact, journals are encouraged to distribute their material as widely as possible, through their own web sites or online distributors. Furthermore, publishers do not need to relinquish their normal copyright provisions for the further commercial use of the material. The great value that PMC brings to the scientific community is the opportunity to search not just titles and abstracts but entire papers for interesting content. Just as GenBank has proved invaluable to molecular biologists, PMC could serve an equally important role within the broader biological community. Once a central repository and archive for the world's biological literature becomes populated it will have a far-reaching impact on the conduct of scientific research. It will improve productivity

and will allow new approaches to searching the literature. No longer will we need to spend hours searching among the stacks of the local, or not so local, library to find articles essential for our research. Scientists, physicians, teachers, and lay people who are currently disenfranchised from the world's literature because of minimal research budgets will have access, perhaps not to the very latest research, but at least to reasonably current research. Our colleagues in the developing world and many of the smaller research institutions will have unprecedented access to the scientific literature.

To populate PMC, all life science journals are being asked to provide their contents free of charge following a suitable delay beyond the date of print publication. In the case of PNAS, the delay is 1 month; for other journals it may be longer. This delay is to mitigate any deleterious effect on subscriptions and the financial health of the journals that might result from free access. For instance, if a journal were to make its content immediately available to PMC, there would be a real danger that subscriptions to the print or online copy of the journal would drop precipitously as libraries become increasingly pressed to find funds for journals. What is a reasonable delay? I would argue that 6 months seems a reasonable time for a journal to monopolize the content. Most of us would not dream of scanning the contents of a journal published 6 months ago unless we were searching for a specific article. Thus it seems unlikely that a large number of subscriptions would be lost if 6-month-old issues were made freely available. I think rather few worthwhile journals would be adversely affected if they were to institute such a policy. I thus welcome, and have

signed on to, the initiative proposed by Pat Brown of Stanford University. He was one of the chief proponents of PMC and is now circulating an open letter from scientists urging journals to participate. The letter is currently posted at [www.publiclibraryof-science.org](http://www.publiclibraryof-science.org). Signatories show their support for open access and pledge to publish in, edit or review for, and personally subscribe to only those journals that grant unrestricted distribution rights within 6 months of publication to PMC and similar entities. As word of this initiative spreads, many of us hope that thousands of scientists, both senior and junior, will sign on. Even more important, we hope that many journals, especially the more prestigious ones, will join PNAS, NAR, BMJ, and others in agreeing to make their content freely available no later than 6 months after publication.

This initiative is very much a grassroots affair. All scientists from students to professors are being asked to join. It is an initiative that, if successful now, will provide a vital resource to students and their professors alike during the coming years. Why might a journal not join something that is so obviously good for science? Some publishers argue that they will lose revenues from subscriptions. This is hard to take seriously, when many journals make their dated content freely available on their own web site and some even offer prepublication copy. I suspect that many publishers and their senior editorial staff are fearful of losing control and jeopardizing favorite programs that they view as benefiting science and that are presently supported from journal profits. However, when I ask students, they seem over-

This is an invited editorial by Richard J. Roberts. Dr. Roberts is a distinguished scientist who received the Nobel Prize for Medicine in 1993 for his part in the discovery of split genes. He also has been a leader in developing new ways to disseminate scientific information, the subject of this editorial. Dr. Roberts participated in the establishment of GenBank, has been involved with many journals, is now an executive editor of *Nucleic Acids Research*, and is a member of the PubMed Central Advisory Board. We present this editorial as a reasoned statement on a topic of great current interest. Any conclusions or opinions expressed in this editorial are not necessarily the views of the National Academy of Sciences.

—Nicholas R. Cozzarelli, *Editor-in-Chief*

Article published online before print: *Proc. Natl. Acad. Sci. USA*, 10.1073/pnas.041601398.  
Article and publication date are at [www.pnas.org/cgi/doi/10.1073/pnas.041601398](http://www.pnas.org/cgi/doi/10.1073/pnas.041601398)

whelmingly in favor of PMC. Indeed, as a practicing scientist how can one reasonably be against it? It will save much time and make invaluable resources uniformly available. It is good for everybody. Both GenBank and PubMed, also run from NCBI, have been immensely successful and have driven science forward. PMC is the next step.

One might have thought that the scientific societies would have been at the forefront to promote the interests of their members and to promulgate science by all means possible. So why have the major life science societies, such as the American-

Society for Microbiology (ASM), the American Society for Biochemistry and Molecular Biology (ASBMB), the American Association for the Advancement of Science (AAAS), etc., not followed the lead of the National Academy of Sciences and rushed to join PMC? At the very least, the societies should poll their members to gauge their enthusiasm for PMC. Could it be that the societies have become seduced by the cash that their journals produce, and the professional interests of the scientists they represent are taking second place? I would urge all scientific societies and academic publishers such as the uni-

versity and institutional presses to take a hard look at their priorities and ask whether they support science or Mammon. I also urge the large commercial publishers to join PMC. They cannot claim to be serving the best interests of their customers by trying to balkanize the published literature. Imagine how stymied we would all be without GenBank. Most of all, though, I urge our young scientists to think hard and carefully about this issue. Your future is at stake. Here is your chance to make your voice heard and indicate your priorities in the scientific enterprise. Join me and sign on!