

PublisherInfo		
PublisherName	:	BioMed Central
PublisherLocation	:	London
PublisherImprintName	:	BioMed Central

Revisiting Bermuda

ArticleInfo		
ArticleID	:	4719
ArticleDOI	:	10.1186/gb-spotlight-20030311-01
ArticleCitationID	:	spotlight-20030311-01
ArticleSequenceNumber	:	71
ArticleCategory	:	Research news
ArticleFirstPage	:	1
ArticleLastPage	:	4
ArticleHistory	:	RegistrationDate : 2003-3-11 OnlineDate : 2003-3-11
ArticleCopyright	:	BioMed Central Ltd2003
ArticleGrants	:	
ArticleContext	:	130594411

Tabitha M Powledge

Email: tam@nasw.org

Top genome scientists are emphatically reaffirming the requirement that large-scale DNA sequence producers release their data to all other researchers immediately after they generate it, and urging vast expansion of this obligatory pre-publication data-sharing.

The principle, they say, should include all sequence data and should be extended to other international collaborative projects such as the [Mammalian Gene Collection](#), the [SNP Consortium](#), the [International HapMap Project](#) - and perhaps eventually to other "community resource projects," such as large-scale protein structure determination or gene expression analysis. The policy might also apply to microbial sequencing, even when a sequence is generated in a single small lab.

At present, the US [National Human Genome Research Institute](#) (NHGRI) and other genome funders call upon the large-scale sequencers they support to release sequence assemblies larger than 2 kb into the public databases within 24 hours; raw shotgun sequences must be released within a week. Other researchers are permitted to use these publicly available data for all purposes before the sequence producer has published, with the exception of "publication of the results of a complete genome sequence assembly or other large-scale analyses."

NHGRI recently posted a [draft update](#) to the policy on its Web site and is asking for comments by April 25 on a proposed revision of the rules, known as the [Bermuda Principles](#), first formulated in 1996. The draft "reaffirms and extends" NHGRI's "commitment to the Bermuda Principles for all types of large-scale DNA sequence data sets, including those that were not considered when the Bermuda Principles were originally devised."

The policy requiring assemblies of 2 kb to be deposited within 24 hours and raw sequence within a week would remain in force, though the publication exception has been dropped. In its place, the document encourages producers to recognize that even if the sequence data are occasionally used in ways that violate normal standards of scientific etiquette, unconditional release is a necessary risk because the benefits of immediate release are "considerable."

Sequence users are in turn reminded that they "are expected to acknowledge the source of the sequence data through the use of appropriate citations" and urged to recognize that producers have a legitimate interest in publishing their own data.

The immediate need to revise the rules arose from what NHGRI director Francis Collins calls breaches of scientific etiquette involving pre-publication data. "A paper would appear describing the sequence of an organism where the authors had not produced a single base pair of that sequence, but had basically analyzed what they could find for free on the web. In a few instances they didn't even acknowledge where they got the information. That's obviously a fairly egregious breach of politeness," Collins told us.

Collins declined to describe specific examples, except to note that some of the questionable papers had not been published because they were spotted by reviewers or editors. But one very public brawl, over publication of data on the protozoan parasite *Giardia lamblia*, was reported in the pages of [Science](#) last year (**295**:5558, 1206-1207).

The proposed new version of the Bermuda Principles grew out of a January meeting in Florida convened by the [Wellcome Trust](#), which drew together some 40 large-scale sequence producers and sequence users - a group that included computational biologists, representatives of the major public databases, journal editors, and scientists interested in other large-scale data sets.

NHGRI is hoping in particular for comments about the pros and cons of extending the principle of speedy pre-publication data sharing to areas that might come under the heading of "community resources," Collins said. "There was a sense that there had not been an adequate discussion yet among those communities because they weren't very well represented at the Florida meeting, which was primarily to talk about DNA sequencing."

One unresolved topic is how to classify microbial genome projects, especially when they are done in a single small lab. Since microbe projects are funded by several US agencies, the question should probably be decided by the existing interagency committee on microbial genomes, Collins said. "My hope would be that they would strongly endorse it. Most of those sequences are primarily done to benefit the broader science community, which means they ought to fit this definition of a community resource project pretty well."

"The key term is what constitutes a community resource project," Ari Patrinos told us. Patrinos, who participated in the Florida meeting, heads the [Office of Biological and Environmental Research](#) at the Department of Energy, which supports many microbe projects. "If an activity is truly a community resource project, whether it is sequence data or proteomics data or microarray data, I think there's a good argument that Bermuda-like principles need to be adhered to."

That key term is "monopoly," according to Sean Eddy, bioinformaticist at [Washington University School of Medicine](#) in St. Louis, who also attended the Florida meeting. The guidelines requiring pre-publication data release, he said, are intended to apply only to projects where the results will be of general interest and widely used - and where, because of economies of scale, funding agencies have channeled money to a few supercenters rather than into competitive grants. "This is an unusual case. It only applies to the big expensive monopolistic projects. It does not apply across the board to genomics data in general."

The social altruism of the pre-publication data-sharing rules is "courageous and bold," said another meeting participant, Geoffrey Duyk, chief scientific officer of [Exelixis, Inc.](#), of South San Francisco. But he is not so sure they will apply only to big expensive monopolistic projects. While the Florida meeting was mostly about current genome sequencing projects that easily fit that definition, "I do think there was a sense in the room that we need to think about what to do as these large scale technologies become more efficient and less costly," he added.

Duyk forecasts that genome projects will be miniaturized and highly distributed in the future. "Individual labs will be generating what are by historical standards tremendous amounts of data that will be of general and broad interest," he said. "Today it's not possible to do that in an individual lab, but I think there was an embedded wish [at the Florida meeting] that we should think about how that information gets out."

References

1. The Mammalian Gene Collection, [<http://mgc.nci.nih.gov/>]
2. The SNP Consortium, [<http://snp.cshl.org/>]
3. The Haplotype Map Project, [<http://www.genome.gov/page.cfm?pageID=10001688>]
4. National Human Genome Research Institute, [<http://www.genome.gov/>]
5. "Reaffirmation and Extension of NHGRI Rapid Data Release Policies: Large-Scale Sequencing and Other Community Resource Projects," February 2003, [<http://www.genome.gov/page.cfm?pageID=10506537>]
6. Bermuda Principles, [<http://www.gene.ucl.ac.uk/hugo/bermuda.htm>]
7. *Science*, [<http://www.sciencemag.org>]
8. Wellcome Trust, [<http://www.wellcome.ac.uk/>]
9. US Department of Energy, Office of Biological and Environmental Research, [http://www.er.doe.gov/production/ober/ober_top.html]
10. Washington University School of Medicine, [<http://medicine.wustl.edu/>]
11. Exelixis, Inc., [<http://www.exelixis.com/>]