nature biotechnology

Time for leadership

The example set by leading proteomics laboratories will be a major factor in determining the successful implementation of new reporting guidelines in the wider community.

This issue includes two perspectives that propose reporting guidelines for proteomics and molecular-interaction data sets (p. 887 and p. 894). The "minimum information about a proteomics experiment" (MIAPE) and an associated module on molecular interaction experiments (MIMIx) were developed by the Proteomics Standards Initiative of the Human Proteome Organization with the aim of standardizing the reporting of proteomics research.

The common deficiencies of proteomics papers are well known: a lack of important data or information, idiosyncratic data formats, inadequate statistics and so forth. MIAPE is not intended to remedy all these deficiencies. It does not provide instruction on how to carry out experiments or how to assess data quality. It does not require particular data formats or particular controlled vocabulary lists. Simply put, MIAPE is a list of the information to include when reporting the results of a proteomics experiment. At the same time, however, it defines a framework in which data formats, controlled vocabularies, databases and other relevant resources can be established.

The movement to impose some kind of order on the reporting of proteomics experiments began several years ago with an initiative by the editors of *Molecular & Cellular Proteomics* (**3**, 531–533, 2004), which resulted in a set of guidelines for papers submitted to that journal. The MIAPE project has continued this effort, focusing on reporting requirements rather than on issues of data quality. MIAPE is an ongoing enterprise, a work in progress. The modules, each of which covers a specific technology or group of technologies, will continue to evolve as new technologies appear. They will also be revised in an iterative fashion in response to community feedback. We would urge all interested parties—experimentalists, bioinformaticians, instrument and software vendors, funding agencies, journals—to get involved in evaluating and contributing to the modules. Before MIAPE and MIMIx were accepted for publication, *Nature Biotechnology* hosted the manuscripts on our website for public review (http://www.nature.com/nbt/consult/index). Although the number of people who submitted critiques was fewer than we would have wished, those who took the time to comment offered many thoughtful suggestions that were incorporated into the papers. We thank them for their input.

Whether *Nature Biotechnology* ultimately elects to require compliance with the MIAPE guidelines will depend on their reception by the scientific community. This March, we began recommending (not requiring) that proteomics and molecular-interaction data sets be deposited in a public repository before the associated manuscript is submitted to this journal (*Nat. Biotechnol.* **25**, 262, 2007). But we would not consider enforcing the MIAPE guidelines until such time as the proteomics community has reached a consensus that the benefits of compliance outweigh the burden.

Before this can happen, at least two critical pieces of infrastructure must be in place. First and foremost, appropriate software tools must be developed and made freely available to all. Second, databases must improve their capabilities for transferring and storing MIAPE-compliant data sets.

At present, a considerable amount of labor is often required to convert a data set into a form that is MIAPE compliant—a situation that favors large, well-funded laboratories and penalizes smaller ones with fewer resources. Under Framework Program 6, the European Union is providing €1 ($1.4) million over 30 months to fund the ProDaC project, which among other things is developing software tools to import data generated by proteomics laboratories into public repositories. The availability of robust open-source conversion tools will go a long way towards encouraging participation by the wider community of experimentalists.

The major vendors of proteomics instrumentation have an important role to play in making MIAPE compliance a less painful exercise for less expert laboratories. Vendors tend to produce better quality software code than academic laboratories (which focus more on *ad hoc* solutions). In the field of DNA microarrays, vendors were encouraged by big pharma companies to agree on standards, and this contributed to the eventual adoption of the MIAME (minimal information about a microarray experiment) standard. Similarly, for MIAPE, it would be beneficial if vendors were collaborative and open minded with respect to open-source software—at least providing code interfaces so that scientists can 'plug-in' modules of their own software to enable data analyses that are particular to their projects. Efforts by mass spectrometry instrument vendors to accept the mzData format are a step in the right direction.

Databases must also provide easy-to-use submission pipelines, efficient data processing, rapid provision of accession numbers and a simple mechanism for anonymous access by referees. Here, the example of the MIAME guidelines offers a cautionary tale: when the *Nature* journals first recommended deposition of MIAME-compliant data sets, databases such as ArrayExpress were initially overwhelmed by the high volumes of submitted data.

Software development and database management require funding. We support the emerging trend among funding agencies to underwrite such projects, to ensure that data generated through their grants are shared publicly, and to consider enforcing appropriate reporting guidelines.

In the meantime, we encourage large proteomics laboratories, and, in particular, those scientists who are authors on the MIAPE and MIMIx papers, to lead by example. By embracing the new guidelines in their publications, prominent proteomics researchers can jumpstart the process by which MIAPE is widely adopted. They have the vision, the resources and the clout to make this happen; those who contributed to the guidelines as authors have a special obligation to follow through. Without their leadership, implementation of MIAPE may languish. We will follow future developments with great interest.