Minimum information standards -- all for 1 and 1 for all

August 26 2007

Three papers published by EMBL scientists and their collaborators will make it much easier to share and compare information from large-scale proteomics data. The papers are published in *Nature Biotechnology* on 8th and 26th August.

As the quantity of available biological information and the use of public data repositories increases, consistency in the information held in these databases is vital to allow full integration, exchange and comparison of their contents. As Europe's main provider of biological data, the EMBL-European Bioinformatics Institute (EBI) is involved in setting the precedent for reporting standards by applying these to its own data repositories such as ArrayExpress (microarray and gene expression data), IntAct (molecular interaction data) and PRIDE (protein identification linked to experimental evidence and publications).

The Minimum Information About a Proteomics Experiment (MIAPE) and the Minimum Information required for reporting a Molecular Interaction Experiment (MIMIx) guidelines propose the range of information to be recorded to document proteomics and molecular interaction data, respectively. The standards aim to reduce ambiguity and capture all the necessary information from an experiment to set the experimental results in both a biological and a methodological context, thereby providing a deeper level of understanding to others exploring the data. Henning Hermjakob from the EMBL-EBI, a co-author on both Perspectives papers published on 8th August, said "Through the community-wide uptake of agreed minimum reporting standards, we can

all benefit from easier identification and use of information that is most relevant to our own areas of work. This is the next step in providing freely accessible data repositories of the highest possible quality."

The *Nature Biotechnology Perspectives* papers, published as open-source articles, outline the proposed reporting requirements for proteomics and molecular interaction experiments and discuss their implementation, impact and benefits. The later research paper, published in the same journal on 26th August, shows how implementation of these standards benefits not only the reporting researchers, but also the wider community through the development of more detailed and comprehensive information resources.

Setting the standards

Both sets of reporting requirements, MIAPE and MIMIx, have been shaped by input from the scientific community and developed to minimise the burden on individual researchers. The MIAPE reporting guidelines are being developed as a range of individual modules, which can be combined as necessary to cover an entire experimental workflow – from study design to statistical data analysis. As the first finalised module, the MIMIx guidelines implement the general MIAPE principles for describing molecular interactions and also present recommendations on data deposition prior to publication.

Standards in action

Both sets of reporting requirements were put into practice by Bantscheff et al. in the reporting of a large-scale approach to profile the interaction of protein kinases with small inhibitory molecules. The Cellzome researchers used this method to validate the action of three drugs sharing a particular kinase target and also to identify novel drug targets. The

binding information shed light on the drug binding specificities and downstream effects on signalling pathways. The MIAPE and MIMIx-compliant mass spectrometry and interaction data were entered into the EBI-hosted PRIDE, IntAct, and ChEBI databases, and accession numbers are included in the research publication to provide direct access to the proteomic information, mass spectra and molecular interaction data. The interconnected EBI data resources serve not only to hold the direct experimental results, but also set them within a wider biological context, for example, by linking the identified kinases to known functions in the UniProt database.

The quantitative profiling method used has potential application in drug discovery and in gaining a greater understanding of drug action in patients. The systematic recording of such information in public repositories and adherence to reporting standards ensures that maximum use can be made of this progression in knowledge, offering benefits to the scientific community, and in the case of drug discovery, development and healthcare, there are clearly benefits to be had for society too.

Source: European Molecular Biology Laboratory