Guidelines for reporting the use of mass spectrometry informatics in proteomics

To the editor:

We would like to draw your readers' attention to the Minimum Information about a Proteomics Experiment (MIAPE) Mass Spectrometry Informatics (MIAPE-MSI) guidelines, which are part of the MIAPE documentation system published in last August's issue[1]. MIAPE-MSI specifies the minimum information that should be provided when reporting the use of mass spectrometry in a proteomics experiment (**Box 1**). It was developed through a joint effort between the Proteomics Informatics working group of the Human Proteome Organisation's Proteomics Standards Initiative (HUPO-PSI; http://www.psidev.info/) and the wider proteomics community. It comprises a checklist of information that should be provided about mass spectrometry–based peptide and protein identification and characterization performed in the course of generating a data set that is submitted to a public repository, or when such an experimental step is reported in a scientific publication (for instance, in the materials and methods section). MIAPE-MSI specifies neither the format in which information should be transferred nor the structure of any repository or document. However, HUPO-PSI is not developing the MIAPE modules in isolation; several compatible data exchange standards are now well established and supported both by public databases and by data processing software in proteomics.

The correct analysis of the data produced by mass spectrometry is fundamental to the generation of reliable biological knowledge. Peptide mass fingerprints and peptide fragment fingerprints are two types of data that can be used in peptide and protein identification, quantification, structural characterization and the investigation of protein modifications. The heterogeneity of sample content and complexity on one hand, and the heterogeneity of mass spectrometers (type, sensitivity, accuracy, efficiency of sample introduction, ionization, processing) on the other, strongly affect the type, amount and quality of experimental information to be analyzed by protein and peptide identification and characterization software. This is highlighted by the complex set of input parameters required by such software; the results generated are similarly complex in terms of both data structure and pertinence. These guidelines for the reporting of the use of such software do not prescribe that all of the available information be captured; and given the diversity of tools currently available, the utility of such detail is clearly open to question. However, it is possible to specify (generic) parameters that are representative of the way in which the software was used and that serve to contextualize the data generated, enabling a better-informed process of assessment and interpretation.

The guidelines (**Supplementary Guidelines** and **Supplementary Table 1** online) cover the use of protein and peptide identification and characterization software, and the data generated. They do not cover the mass spectrometry that generated the data or the reduction of 'raw' profile data to peak lists; those areas are addressed in the MIAPE-MS (mass spectrometry) module, the latest version of which can be obtained from the MIAPE home page. Note also that these guidelines do not cover all the available features of a protein and peptide identification and characterization tool (e.g., some of the less frequently used parameters, types of spectra or other experimental data); subsequent versions may have expanded coverage, as will almost certainly be the case for all MIAPE modules.

These guidelines will evolve in step with progress in research. The most recent version of MIAPE-MSI is available at http://www.psidev.info/miape/msi/ and the content is replicated here as supplementary information (**Supplementary Guidelines** and **Supplementary Table 1**). To contribute or to track the process to remain 'MIAPE compliant', browse the website at http://www.psidev.info/miape/.

Note: Supplementary information is available on the Nature Biotechnology website.

Box 1 Content snapshot for MIAPE-MSI

The full MIAPE-MSI document is divided into three parts: an introduction providing background and context; a summary list of the items to be reported; and a glossary with definitions and examples.

The MIAPE-MSI guidelines themselves are subdivided as follows:

1. General features; the software employed.
2. Input data and parameters.
3. The output from the procedure; the list of peptides and proteins identified, characterized or quantified.
4. Interpretation and validation.

Pierre-Alain Binz[1,2], Robert Barkovich[3], Ronald C Beavis[4], David Creasy[5], David M Horn[6], Randall K Julian Jr[7], Sean L Seymour[8], Chris F Taylor[9,10] & Yves Vandenbrouck[11]

[1]Swiss Institute of Bioinformatics, Rue Michel-Servet 1, CH-1211 Geneva 4, Switzerland. [2]GeneBio SA, 25 Av. de Champel, CH-1206 Geneva, Switzerland. [3]Affymetrix, Inc., 3420 Central Expressway, Santa Clara, California 95051, USA. [4]Biomedical Research Centre, University of British Columbia, Vancouver, British Columbia, Canada. V6T 1Z3. [5]Matrix Science Ltd., 64 Baker Street, London W1U 7GB, UK. [6]Agilent Technologies, 5301 Stevens Creek Blvd., Santa Clara, California 95051, USA. [7]Indigo BioSystems, Inc., 111 Congressional Blvd., Suite 160, Carmel, Indiana 46032, USA. [8]Applied Biosystems, 850 Lincoln Centre Drive, Foster City, California 94404, USA. [9]European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SD, UK. [10]NERC Environmental Bioinformatics Centre, Mansfield Road, Oxford, OX1 3SR, UK. [11]CEA, DSV, DRDC, laboratoire de Biologie, Informatique et Mathématiques, 17 rue des Martyrs, Grenoble, F-38054 France.
e-mail: Pierre-Alain.Binz@isb-sib.ch

1. Taylor, C.F. *et al. Nat. Biotechnol.* **25**, 887–893 (2007).