
CIMR: Environmental Analysis Context

Metabolomics Standards Initiative (MSI)

Sponsor: Metabolomics Society (<http://www.metabolomicsociety.org/>)

Reference: <http://msi-workgroups.sourceforge.net/bio-metadata/>

Version: 1.0

Date: 06/12/2006

Authors: Dan Bearden, NOAA, Charleston SC, USA (dan.bearden@noaa.gov)
Jake Bundy, Imperial College, UK (j.bundy@imperial.ac.uk)
Tim Collette EPA, Athens GA, USA (collette.tim@epamail.epa.gov)
Felicity Currie, University of Manchester, UK (felicity.currie@postgrad.manchester.ac.uk)
Matthew P. Davey, University of Sheffield, UK (m.davey@sheffield.ac.uk)
Dawn Field, CEH Oxford, UK (dfield@ceh.ac.uk)
Nathan S. Haigh, University of Sheffield, UK (n.haigh@sheffield.ac.uk)
David Hancock, University of Manchester and EBC, UK (hancockd@cs.man.ac.uk)
Oliver Jones, University of Cambridge, UK (oahj2@mole.bio.cam.ac.uk)
Norman Morrison, University of Manchester, UK (morrison@cs.man.ac.uk)
Simone Rochfort, DPI, Werribee, Australia (simone.rochfort@dpi.vic.gov.au)
Susanna-Assunta Sansone, EBI, Cambridge, UK (sansone@ebi.ac.uk)
Dalibor Štys, University of South Bohemia České Budějovice, Czech Republic (stys@jcu.cz)
Quincy Teng, EPA, Athens GA, USA (Teng.Quincy@epa.gov)
Mark Viant, University of Birmingham, UK (m.viant@bham.ac.uk)

Credits: *Chair:* Norman Morrison

Co-chairs: Mark Viant & Dawn Field

Copyright © 2006 MSI - Metabolomics Society and contributing authors

Table of Contents

1.0. This Document.....	2
2.0. Goals of the Environmental Context Working Sub-Group.....	2
3.0. Recommended Information.....	3
4.0. Minimum Reporting Standards for Environmental Metabolomic Studies.....	4
4.1. Sources.....	4
4.2. Instructions to Authors	4
4.3. Requirement Groups.....	5
Sample.....	5
Environment.....	6
Process.....	7
5.0. Additional Information.....	8
6.0. Bibliography	8

1.0. This Document

This document forms part of the standards for reporting metabolomics experiments developed under the Metabolomics Society (<http://www.metabolomicssociety.org/>) Metabolomics Standards Initiative (MSI). It should be read in the context of top level document for those standards.

The current version of the document is work in progress.

The scope of our effort is to identify, develop and disseminate a core set of reporting requirements necessary for the consistent annotation of environmental metabolomic experiments. This working group considers environmental metabolomics as the application of metabolomics to investigate free-living organisms obtained directly from the environment (whether studied in the environment or returned to a laboratory for further experimentation), as well as studies on laboratory-reared organisms (whether studied in the laboratory or transferred to the environment for further experimentation), where any laboratory experiments specifically serve to mimic scenarios encountered in the environment.

The aim of this document is to provide a minimum set of reporting requirements that describe and support the dissemination and re-use of metabolomic data. It is not our intention to be prescriptive about how one should perform a metabolomic experiment. Furthermore, given the complexity of environmental metabolomics experiments, the requirements outlined below should not be considered as exhaustive. Such reporting standards will specify the data identified as strongly recommended for adequate and correct reporting in a range of identified contexts, such as submission to academic journals and public databases. The group also recommends that data exchange standards be developed in order to provide a transparent technical vehicle which meets or exceeds the reporting requirements.

2.0. Goals of the Environmental Context Working Sub-Group.

The ECWSG seeks to:

- I. Work cooperatively on a consensus document for a minimum set of reporting requirements needed to unambiguously interpret or reproduce the results of a metabolomics based experiment generated in an environmental context.
- II. Represent the views of community members working in the area of environmental metabolomics in an unbiased and open fashion. The group will engage with environmental metabolomics practitioners through a process of “use case” provision and validation of selected requirements.
- III. Assess previous and relevant standardisation efforts in environmental biology, including similar work in genomics, transcriptomics and proteomics studies.
- IV. Interact with the MICheck (<http://micheck.sourceforge.net/>) project as well as with individual technology specific groups. These include the Microarray Gene Expression Data Society (MGED), the Human Proteome Organisation - Proteomics Standards Initiative (HUPO-PSI) and the Genomic Standards Consortium (GSC).

The group recognises that the larger scientific community will best be served if the resulting specification overcomes duplications across omics domains where commonality of the concepts exists.

- V. Evaluate our reporting requirements in the context of the other biological context working groups with an aim to identify any overlaps or gaps in our reporting.

The ECWSG aims to provide a reporting structure to cover all possible scenarios occurring in environmental investigations. To do this we propose minimal description sets for – but not limited to – the following four fundamental activities (and combinations thereof).

1. *Field study* - Investigations in which an organism is sampled directly from the field.
2. *Laboratory study* - Investigations in which a sample is laboratory reared or obtained from a standard provider and conditioned in the laboratory under controlled conditions, in order to mimic a scenario encountered in the environment.
3. *Removal of samples from field to laboratory* - Process in which an organism is removed from the field and experimented upon in the laboratory under controlled conditions.
4. *Release of samples from laboratory to field* - Process in which an organism is taken from the laboratory and experimented upon in the field under semi-controlled conditions.

We also recognise that an organism can have a ‘biotic environment’, where the immediate environment of an organism is another organism (for example in a parasitic relationship). This scenario necessitates a nested description of the environment, including a description of the ‘biotic environment’ of the primary organism and that of the host.

3.0. Recommended Information

The minimal information set for reporting the samples in environmental metabolomic experiments builds upon the general biological practice of reporting biological experiments in scientific journals in a way that the ‘materials and methods’ section should include sufficient information to allow the experiments to be repeated (‘Instructions to authors, ASM’ - <http://www.journals.asm.org/>). These include aspects such as:

- Organism
- Source of the organism
- Experimental design

Scientific journals also ask that authors describe those conditions that are critical to the experiments performed in larger detail. Since metabolomics data are extremely sensitive to external factors, as much information as possible should be recorded. Below we have identified those factors that are critical in the description of *environmental* metabolomic experiments.

The challenge is not to provide reporting requirements for everything, but to provide a framework where the environmental features considered relevant to a particular sample, which will be highly context dependant, can be captured in a structured and extensible manner.

All requested information outlined below in standard font should be considered as 'strongly recommended' for submission; any missing information should be justified. Requested information marked in *italics* should be considered as optional further details.

All ontologies or controlled vocabularies (CVs) we suggest in this document are from publicly available resources. Our terminology requirements and recommendations will also be collected by the MSI Ontology WG (<http://msi-ontology.sourceforge.net/>), which is registered under the Open Biomedical Ontologies umbrella (OBO, <http://obo.sourceforge.net>).

4.0. Minimum Reporting Standards for Environmental Metabolomic Studies

4.1. Sources

- The International Council for the Exploration of the Sea (ICES) integrated environmental reporting format document [1]
- Data reporting requirements from the PETRAEA project [2]
- Plant Information Management System (PIMS) Database Schema [3]
- UK Protocols for Standard Measurements at Freshwater Sites [4]
- Sharing Publication-Related Data and Materials: Responsibilities of Authorship in the Life Sciences [5]
- MIAME - Minimum information about a microarray experiment [6]
- MIAME/Env - Annotation of environmental OMICS data: application to the transcriptomics domain [7]
- A proposed framework for the description of plant metabolomics experiments and their results [8]
- Potential of metabolomics as a functional genomics tool [9]
- The concept of a sample in OMICS technology [10]

4.2. Instructions to Authors

We have presented these reporting requirements as a collection of small *groups*, each focussed on one small portion of the reporting process. These groups are designed to be interconnected to enable as wide a variety of possible experimental models to be described as possible.

These groups of requirements are not a linear checklist in the traditional manner, therefore you should not attempt to work through the groups in sequence. Instead, when describing

your study, please consider the requirements from any of the following groups which are relevant. Please note that you may need to use some of the groups more than once in your description and some you may not use at all.

The requirements have been split into 3 major groups:

1. Sample (**S**) - descriptions of the biological sample(s) involved in the study.
2. Environment (**E**) - descriptions of the environment(s) involved in the study.
3. Process (**P**) - descriptions of the processes involved in the study.

For each part of your experimental description, you should identify which group is most suitable, and follow the requirements of that group.

For example:

1. If your study involved collecting shellfish from the shoreline and cultivating them in a laboratory before analysis, then you should examine groups: **S1, E1, E2, E3.2 and P1-6.**
2. If your study involved collecting whole plants from the side of a mountain, dissecting tissue and examining metabolites, then you should examine groups: **S1, E1, E3.1, P1 and P6.**

4.3. Requirement Groups

Sample

Please note: Currently, there is only one group under the general heading of Sample. The group has been categorised 'S1' (as opposed to just 'S') in order to maintain consistency with the naming convention applied to other groups and to allow for potential extension and/or sub-division of groups as the reporting requirements evolve (**Sn**).

- **S1** : Description of the biological sample involved in the study:
See also the requirements in P1 – P6
 - Taxonomic classification of organism(s). Please give details of all organisms sampled, as far along the taxonomic scale as possible, ideally to the levels of genus, species and sub-species. Refer to a taxonomic classification, such as <http://www.itis.gov/>
 - Also, include where possible:
 - *Common name(s) (vernacular)*
 - *Genotype(s)*
 - *Ecotype(s)*
 - Sample composition. Please provide details of the organism(s) that constitute the sample. Amounts may be described in 'absolute' (number of individuals) or relative terms (50% Organism X; 25% Organism Y; 25% Unknown)
 - Sample Type. Please give details of the type of sample (For example; Community, Population, Whole Organism, Organ, Biofluid, etc.). Refer to an ontology, such as the Ontology Biomedical Investigations (OBI) - <http://obi.sourceforge.net/>
 - *Condition of specimen(s)*. Please give details of general observations on health, etc.
 - *Phenotypic characteristic(s)*
 - *Weight of specimen(s)*
 - *Age(s) of specimen(s)*
 - *Sex(es) of specimen(s)*

- *Stage(s) of development*
- *Image data.* Please provide photographs of samples taken in the field during collection (or URLs to such images)

Environment

- **E1** : Description of ANY field environment:
See also the requirements in S1 and P1 – P6
 - Geographic location. Please specify latitude and longitude in decimal degrees. If relevant, you can also provide position in a local coordinate system e.g. the UK's OS grid - <http://www.ordnancesurvey.co.uk/>
 - Altitude / Depth. Please specify in meters above/below mean sea level
 - Habitat. Please provide a descriptor of habitat type. Refer to an ontology, such as the environmental habitat subsection of the Ontology of Biomedical Investigations (OBI) - <http://obi.sourceforge.net/>
 - *Meteorological conditions.* For example:
 - *Weather type (for example, sunny, snowing etc)*
 - *Humidity*
 - *Precipitation*
 - *Wind speed and direction*
 - *Lunar/solar phase*
 - *All other measured parameters.* For example:
 - *Pollutant concentration(s)*
- **E2** : Description of ANY laboratory environment:
See also the requirements in S1 and P1 – P6
Please also refer to the MSI – *In vivo* context requirements at <http://msi-workgroups.sourceforge.net/bio-metadata/reporting/invivo/mammalincontext.pdf> and the MSI Plant context requirements at <http://msi-workgroups.sourceforge.net/bio-metadata/reporting/pbc/doc.pdf>
 - Laboratory address and contact details
- **E3.1** : Description of terrestrial environment:
See also the requirements in E1, E2.
 - Inclination and aspect
 - Substrate type
 - Substrate temperature
 - *All other measured parameters.* For example:
 - *Substrate pH*
 - *Substrate organic content*
- **E3.2** : Description of aquatic environment:
See also the requirements in E1, E2.
 - Sample(s) was submerged/emerged (how deep and for how long in this condition)
 - Water temperature
 - *Tidal phase*
 - *All other measured parameters.* For example:
 - *pH*
 - *Salinity*
 - *Dissolved (in)organic content*
- **E3.3** : Description of atmospheric environment:
See also the requirements in E1, E2.

- Atmospheric temperature
- *All other measured parameters*. For example:
 - *Atmospheric pressure*
 - *(In)organic content*
- **E4** : Description of biotic environment:
 Please also refer to the MSI – *In vitro* Biology/Microbiology context requirements: <http://msi-workgroups.sourceforge.net/bio-metadata/reporting/invitro/doc.pdf>
See also the requirements in SI.
 - Description of host organism
 - Relationship of organism(s) to host. For example: Amensalism, commensalism, mutualism, parasitism
 - *All other measured parameters*. For example:
 - *pH*
 - *Temperature*

Process

Any description of a process must be accompanied by information identifying **who performed the action**, and the **time-point (relative or absolute) or interval** over which it occurred. Relative time points can be expressed in terms of a specific interval, since or until, an identifiable event, i.e. 4 hours since dose of environmental toxicant, or 20 minutes after sunrise.

- **P1** : Description of capture/sampling of sample or organism(s):
 - Description of capture/sampling procedure
 - Means of capture/sampling. For example: Netted, electrically stunned, anaesthetised, razor cut etc. Refer to an ontology, such as the Ontology Biomedical Investigations (OBI) - <http://obi.sourceforge.net/>
 - *Reason for capture*
 - *Other capture parameters*
- **P2** : Description of storage/preservation of sample(s):
 - Description of storage/preservation procedure
 - Means of storage/preservation. For example: Liquid N, dry ice, formaldehyde, etc. Refer to an ontology, such as the Ontology Biomedical Investigations (OBI) - <http://obi.sourceforge.net/>
 - *Reason for storage/preservation*
 - *Other storage/preservation parameters*. For example:
 - *Temperature*

P3 : Description of maintenance of organism(s): Please also refer to the MSI – *In vivo* context requirements at <http://msi-workgroups.sourceforge.net/bio-metadata/reporting/invivo/mammalincontext.pdf> and the MSI Plant context requirements at <http://msi-workgroups.sourceforge.net/bio-metadata/reporting/pbc/doc.pdf>

- Description of maintenance procedure.
- Means of maintenance. For example: Cage, aquaria, continuous culture, etc. Refer to an ontology, such as the Ontology Biomedical Investigations (OBI) - <http://obi.sourceforge.net/>
- *Reason for maintenance*
- *Other maintenance parameters*. For example:

- *Feeding regime*
 - *Cage dimensions*
- **P4** : Description of transportation of samples or organism(s):
Transportation involves storage and/or maintenance; see also the requirements in either P2 or P3 as appropriate.
 - Description of transportation procedure
 - Means of transportation. For example: Refrigerated container. Refer to an ontology, such as the Ontology Biomedical Investigations (OBI) - <http://obi.sourceforge.net/>
 - *Reason for transportation*
 - *Other transportation parameters.*
 - **P5** : Description of acclimation of organism(s):
Acclimation involves maintenance; see also the requirements in P3 as appropriate.
 - Description of acclimation procedure
 - Means of acclimation. For example: Cage, aquaria, continuous culture, etc. Refer to an ontology, such as the Ontology Biomedical Investigations (OBI) - <http://obi.sourceforge.net/>
 - *Reason for acclimation*
 - *Other acclimation parameters*
 - **P6** : Description of general manipulation of sample or organism(s):
Record details of any controlled manipulation as part of the study
 - Manipulation type. For example: Perturbation such as exposure to a toxicant, dissection, sacrifice etc.
 - Description of manipulation procedure. For example: Perturbation of specific environmental parameter; dissection of specific tissue
 - *Reason for manipulation*
 - *Other manipulation parameters*

5.0. Additional Information

Please be aware that there are additional requirements necessary for *in-vivo*, microbial and plant samples. In addition the processing of samples (e.g. extraction methods and analytical technologies used) is also subject to separate minimum reporting requirements. These are all detailed on the MSI webpage at <http://msi-workgroups.sourceforge.net/>

6.0. Bibliography

1. The International Council for the Exploration of the Sea (ICES) Integrated Environmental Reporting Format document. (Available at <http://www.ices.dk/env/repfor/ERF322.doc>).
2. Data reporting requirements from the PETRAEA project (Available at <http://www.petraea.shef.ac.uk/>).
3. Plant Information Management System (PIMS) Database Schema (Available at http://www.bioinf.shef.ac.uk/pims/images/FieldDB_new.png).

4. Sykes JM, Lane AMJ, George, DG. (Editors). *Protocols for Standard Measurements at Freshwater Sites*. The United Kingdom Environmental Change Network. Institute for Terrestrial Ecology, Grange-over-Sands, UK. 1999: 134.
5. Cech TR, *Sharing Publication-Related Data and Materials: Responsibilities of Authorship in the Life Sciences*. (Available at www.nap.edu/books/0309088593/html).
6. Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, Aach J, Ansorge W, Ball CA, Causton HC *et al*: Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat Genet* 2001, 29(4):365-371.
7. Morrison N, Wood AJ, Hancock D, Shah S, Hakes L, Gray T, Tiwari B, Kille P, Cossins A, Hegarty M *et al*: Annotation of environmental OMICS data: application to the transcriptomics domain. *Omics* 2006, 10(2):172-178.
8. Jenkins H, Hardy N, Beckmann M, Draper J, Smith AR, Taylor J, Fiehn O, Goodacre R, Bino RJ, Hall R *et al*: A proposed framework for the description of plant metabolomics experiments and their results. *Nat Biotechnol* 2004, 22(12):1601-1606.
9. Bino RJ, Hall RD, Fiehn O, Kopka J, Saito K, Draper J, Nikolau BJ, Mendes P, Roessner-Tunali U, Beale MH *et al*: Potential of metabolomics as a functional genomics tool. *Trends Plant Sci* 2004, 9(9):418-425.
10. Morrison N, Cochrane G, Faruque N, Tatusova T, Tateno Y, Hancock D, Field D: Concept of sample in OMICS technology. *Omics* 2006, 10(2):127-137.