

## Commentary

### MIAME/Plant – adding value to plant microarray experiments

Philip Zimmermann<sup>1\*</sup>, Beatrice Schildknecht<sup>2</sup>, David Craigon<sup>2</sup>, Margarita Garcia-Hernandez<sup>3</sup>, Wilhelm Gruissem<sup>1</sup>, Sean May<sup>2</sup>, Gaurab Mukherjee<sup>4</sup>, Helen Parkinson<sup>4</sup>,  
Seung Rhee, , Ulrich Wagner<sup>1</sup>, Lars Hennig<sup>1</sup>.

<sup>1</sup> Swiss Federal Institute of Technology - ETH Zurich  
ETH Center, CH-8092 Zurich, Switzerland.

<sup>2</sup> The Nottingham Arabidopsis Stock Center (NASC)  
Division of Plant Sciences, University of Nottingham,  
Sutton Bonington LE12 5RD, UK

<sup>3</sup> The Arabidopsis Information Resource (TAIR)  
Carnegie Institution of Washington  
260 Panama St  
Stanford, CA 94305, USA

<sup>4</sup> European Bioinformatics Institute (EBI)  
European Bioinformatics Institute, EMBL outstation  
Wellcome Trust Genome Campus, Hinxton  
Cambridge CB10 1SD, UK

\* To whom correspondence should be addressed. E-mail:  
philip.zimmermann@ipw.biol.ethz.ch

## **Abstract**

Appropriate biological interpretation of microarray data calls for relevant experimental annotation. The widely accepted MIAME guidelines provide a generic, organism-independent standard for minimal information about microarray experiments. In its overall structure, MIAME is very general and specifications cover mostly technical aspects, while relevant organism-specific information useful to understand the underlying experiments is largely missing. If plant biologists want to use results from published microarray experiments, they need detailed information about biological aspects, such as growth conditions, harvesting time or harvested organ(s). Here, we propose MIAME/Plant, a standard describing which biological details to be captured for describing microarray experiments involving plants. We expect that a more detailed and more systematic annotation of microarray experiments will greatly increase the use of transcriptome data sets for the scientific community. The power and value of systematic annotation of microarray data is convincingly demonstrated by data warehouses such as Genevestigator® or NASCArrays, and better experimental annotation will make these applications even more powerful.

High-throughput technologies are rapidly reshaping the plant research landscape. Gene expression microarrays, in particular, have impacted our way of thinking and challenged our concepts about research by providing genome-wide results at an increasing rate. A recent search in PubMed with the terms “microarray” and “plant” yielded more than 600 references, most of which report results obtained using RNA profiling technology. A major portion of the published datasets can be obtained via repositories, as supplements to scientific publications, or from the authors directly. Precise experiment annotation, however, that is easily accessible and suitable for comparisons, reproduction and/or correct interpretation of experiments is not always fully available. In fact, standard terms to describe conditions for plant microarray experiments so far have not been defined. While the technical details of a microarray experiment are usually abundant, the biological details are frequently poorly described, unsystematic between laboratories, or even missing.

MIAME (Minimum Information About a Microarray Experiment) is a standard that provides a conceptual framework for core information to be captured from most microarray experiments [1]. The microarray community has been very favorable to the establishment and implementation of this standard. Manufacturers, software developers and international databases have contributed to the development and adoption of the MIAME guidelines. Many scientists reporting microarray experiments now use this standard, which is required by an increasing number of scientific journals.

The MIAME standard has proven very useful so far. Data from repositories can be grouped according to standardized categories with respect to technology platform, labeling and hybridization procedures, measurement data, and array design. Nevertheless, experiment descriptions remained unsatisfactory for many researchers browsing databases, because the use of a common denominator for all biological sciences makes a full annotation of specialized community-level research difficult.

Specifically, the core MIAME standard is limited in its ability to capture domain-specific information on experimental design and sample preparation. Readers of publications describing plant microarray experiments are not only interested in hybridization or normalization protocols, but also e.g. whether the plants were grown under high or low light conditions, on soil or in vitro, or which organs were harvested at what age etc. Domain-specific extensions of the existing general MIAME standard are needed to collect experimental annotation in a structured way. MIAME/Plant, MIAME/Env and MIAME/Tox are such extensions (see <http://www.mged.org/Workgroups/MIAME/miame.html>) [2,3].

In plant science, the use of benchmark terminology for growth stages [4] and plant organs (The Plant Ontology Consortium; <http://www.plantontology.org>) so far has allowed the pioneering development of powerful tools for data storage and analysis [5,6,7]. The rapidly increasing number of microarray studies on plants, together with the prospect of effective data querying by annotation of experimental parameters using controlled vocabularies, now requires a plant-specific MIAME standard. The objective of MIAME/Plant is to facilitate and normalize experiment and sample annotations (Figure 1). This will make annotation browsing and data analysis much easier for end-users, including experimental biologists. More specifically, MIAME/Plant incorporates the core MIAME standard together with a supplementary guideline for experimental description and design and sample preparation protocols (Figures 1 and 2 for more details). The current complete MIAME/Plant guidelines can be downloaded from the Microarray Gene Expression Data Society (<http://www.mged.org>) and from the Nottingham Arabidopsis Stock Center (<http://www.arabidopsis.info>) websites. Currently, MIAME/Plant is being implemented in international repositories [8], and user-friendly tools will further encourage the establishment of these guidelines in the plant community.

The implementation of MIAME/Plant should not create an extra burden on the experimenter, but will enable a targeted workflow through the annotation process that is facilitated by structured data entry forms. While all reported experiments should adhere to MIAME, we do not propose to enforce that every experiment should be annotated to the full detail of MIAME/Plant. Instead, laboratories that decide to collect additional information are encouraged to follow the MIAME/Plant scheme. Systematic and controlled annotation will generate breakthroughs in data mining capabilities, and therefore we encourage the plant community to adopt the MIAME/Plant standard. We expect that the benefits of domain-specific MIAME extensions will also motivate researchers in other fields to develop better community standards.

## References

1. Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, Aach J, Ansorge W, Ball CA, Causton HC, Gaasterland T, Glenisson P, Holstege FC, Kim IF, Markowitz V, Matese JC, Parkinson H, Robinson A, Sarkans U, Schulze-Kremer S, Stewart J, Taylor R, Vilo J, Vingron M (2001) Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat Genet* 29: 365-371.
2. Bao W, Schmid JE, Goetz AK, Ren H, Dix DJ (2005) A database for tracking toxicogenomic samples and procedures. *Reprod Toxicol*. 19:411-419.

3. Field D, Tiwari B, Snape J (2005) Bioinformatics and data management support for environmental genomics. *PLoS Biol.* 3: e297.
4. Boyes DC, Zayed AM, Ascenzi R, McCaskill AJ, Hoffman NE, Davis KR, Gorchach J (2001) Growth stage-based phenotypic analysis of Arabidopsis: a model for high throughput functional genomics in plants. *Plant Cell* 13: 1499-1510.
5. Craigon DJ, James N, Okyere J, Higgins J, Jotham J, May S (2004) NASCArrays: a repository for microarray data generated by NASC's transcriptomics service. *Nucleic Acids Res* 32 Database issue: D575-577.
6. Rhee SY, Beavis W, Berardini TZ, Chen G, Dixon D, Doyle A, Garcia-Hernandez M, Huala E, Lander G, Montoya M, Miller N, Mueller LA, Mundodi S, Reiser L, Tackland J, Weems DC, Wu Y, Xu I, Yoo D, Yoon J, Zhang P (2003) The Arabidopsis Information Resource (TAIR): a model organism database providing a centralized, curated gateway to Arabidopsis biology, research materials and community. *Nucleic Acids Res.* 31:224-228.
7. Zimmermann P, Hirsch-Hoffmann M, Hennig L, Gruissem W (2004) GENEVESTIGATOR: Arabidopsis Microarray Database and Analysis Toolbox. *Plant Physiol* 136: 2621-2632.
8. Mukherjee G, Abeygunawardena N, Parkinson H, Contrino S, Durinck S, Farne A, Holloway E, Lilja P, Moreau Y, Oezcimen A, Rayner T, Sharma A, Brazma A, Sarkans U, Shojatalab M (2005) Plant-based Microarray Data at the EBI. Introducing

AtMIAMExpress a Submission Tool for Arabidopsis Gene Expression Data to ArrayExpress. *Plant Physiol* 139: 632-636.

## **Figure Legends**

**Figure 1.** A schematic representation of the six components of a microarray experiment as defined by MIAME (Brazma et al., 2001). The MIAME/Plant parameters and ontologies extend the basic experiment and sample annotations.

**Figure 2.** Overview of the main classes of ontologies currently represented in MIAME/Plant.