Environmental genomics is a diverse, heterogeneous discipline, often involving multi-factorial experiments that can have an almost infinite number of experimental parameters. Data of this nature is inherently more complex than that generated by other disciplines. Therefore, there is a greater necessity to sufficiently capture the minimal information associated with these parameters.

MIAME (Minimum Information About a Microarray Experiment) outlines a conceptual structure for defining the core information that is common to most microarray experiments. Recognising the importance of being able to disseminate, interpret unambiguously and potentially reproduce and verify array-based gene expression monitoring experiments, the Environmental Genomics (EG) community is keen to conform to, and make use of, the guidelines outlined by the MGED society[1] in the MIAME 1.1 document[2].

A recent effort, MIAME for Toxicogenomics[3] (MIAME/Tox), seeks to expand the core standard to support more closely the field of toxicogenomics. Realising some of the limitations of the core MIAME standard with respect to Environmental Genomics (EG), MIAME/Env aims to define a similar set of guidelines for the capture and exchange of microarray data, common to most environmental genomic experiments. In order to preserve the core structure of MIAME and in keeping with its existing extensions, MIAME/Env is based on the MIAME/Tox draft proposed by the EMBL-EBI, NIHES National Center for Toxicogenomics (NCT) and ILSI-HESI[3]. Defining a core that is common to most environmental genomics experiments has been of particular importance due to the Natural Environment Research Councils (NERCs) commitment to making MIAME compliance a de-facto standard within the Environmental Genomics Science Programme.

Most environmental genomics experiments can be placed into one of three main categories:

1. Field trials – experiments in which an organism/biosource is taken directly from the field, samples and extracts are prepared and the hybridisations then performed.
2. Conditioned field trials – experiments in which an organism/biosource is taken from the field and conditioned in the lab under specified animal husbandry conditions (preconditioning) and where “treatments” (conditioning) may be performed pre or post sample/extract preparation.
3. Lab experiments – experiments in which an organism/biosource is lab reared or obtained from a standard provider and where “treatments” (conditioning) are usually performed pre or post sample/extract preparation.

MIAME/Env envisages that these three experimental types will require different levels of experimental detail to be provided in certain sections of the document. For example - for field trials, the environmental history and sample source location (section 2.1) will be of paramount importance and hence the information provided here will give a high degree of detail, whereas sections such as animal husbandry and sample manipulations (section 2.2) may be of less importance and might be relatively sparsely populated. As such **Guidance notes are provided at the beginning of some sections of the document to aid in their completion.**

**The MIAME/Env Ontologies and Controlled Vocabularies**

What is an ontology?

(From 'The Wiley Dictionary of Bioinformatics’ 2004)

An ontology is a way of capturing how a particular community thinks about its field. An ontology attempts to capture a community’s knowledge or understanding of a domain of interest. Like any representation an ontology is only a partial representation of the communities understanding. Ontologies are used to share a common understanding between both people and computer systems. Most ontologies consist of at least concepts, the terms that name those concepts and relationships between those concepts. Ideally, the concepts have definitions and the collection of terms (lexicon) provides a vocabulary by which a community can talk about its domain. The ontology provides a structure for the domain and constrains the interpretations of the terms the ontology provides [3]. The goal of an ontology is to create an agreed-upon vocabulary and semantic structure for exchanging information about that domain [4].

The MGED Society is developing an ontology to provide standard terms for the annotation of microarray experiments. The MGED Ontology [5,6] is intended for the use of investigators in annotating their microarray experiments and for software and database developers. These terms will enable structured queries of elements of the experiments and allow unambiguous description of how the experiment was performed. A core MGED Ontology will be established by Sept. 2003, and will remain unchanged to facilitate software applications development. However, as new applications of microarray technology arise, requiring new descriptive terms, a second layer of ontology will be built. The MGED Ontology also participates in the global open biological ontologies effort (GOBO) [7] that collects freely available ontologies for use within the genomics and proteomics domains. Therefore, the MGED Ontology will only include reference to available free resources.
Environmental genomics applications require specific, common terminologies that need to be identified and subsequently integrated within the MGED Ontology. MIAME/Env content areas for gene expression experiment descriptions include sections that will be provided in a free text format, along with information that are recommended to be provided by maximum use of controlled vocabularies or external ontologies (such as species taxonomy, cell types, anatomy terms and chemical compound nomenclature). The use of controlled vocabularies is needed to enable database queries and automated data analysis.

Since few controlled vocabularies have been fully developed, MIAME/Env encourages the users, if necessary, to provide their own qualifiers and values identifying the source of the terminology. This is achieved through the use of

\[(\text{qualifier, value, source})\]

triplets, for instance,


This is recommended instead or in addition to free text format descriptions wherever possible and will allow the community to build a knowledge base of the most useful controlled vocabularies for describing microarray experiments.

The MIAME/Env structure

Overview

Microarrays are often manufactured independently of particular experiments and their design description can be given separately. Therefore, as with MIAME and MIAME/Tox, MIAME/Env has two major sections.

- Array design description;
- Experiment description.

The two components of MIAME/Env are discussed in further detail below. Whilst the array design description remains almost identical to that in MIAME and MIAME/Tox the second section is extended to fulfill the needs of the Environmental genomics community.
Array design description

The array design specification consists of the description of the common features of the array as the whole, and the description of each array design elements (e.g., each spot). Following terminology used in MAGE, we distinguish between three levels of array design elements: feature – the location on the array, reporter – the nucleotide sequence present in a particular location on the array, and composite sequence – a set of reporters used collectively to measure an expression of a particular gene, exon, or splice-variant. The details that should be given for each of them are described below. For controlled vocabularies that can be used to assist completion see http://www.mged.org/ontology

1) Array related information
   - array design name
   - platform type: in situ synthesized, spotted or other
   - surface and coating specification
   - physical dimensions of array support (e.g. of slide)
   - number of features on the array
   - availability (e.g., for commercial arrays) or production protocol for custom made arrays

2a) For each reporter type
   - the type of the reporter: synthetic oligo-nucleotides, PCR products, plasmids, colonies, other
   - single or double stranded

2b) For each reporter
   - sequence or PCR primer information: (If known, otherwise a unique identifier and an accompanying reference to a database containing additional information on that identifier)
   - sequence or a reference sequence (e.g., for oligonucleotides), if known
   - sequence accession number in DDBJ/EMBL/GenBank, if exists
   - primer pair information, if relevant
   - approximate lengths if exact sequence not known
   - clone information, if relevant (clone ID, clone provider, date, availability)
   - element generation protocol that includes sufficient information to reproduce the element for custom-made arrays that are not generally available

3a) For each feature type
   - dimensions
• attachment (covalent/ionic/other)
• technology used to generate the feature

3b) For each feature
• which reporter and the location on the array

4) For each composite sequence
• which reporters it contains
• the reference sequence
• gene name and links to appropriate databases (e.g., SWISS-PROT, or organism specific databases), if known and relevant

5) Control elements on the array
• position of the feature (the abstract coordinate on the array)
• control type (spiking, normalization, negative, positive)
• control qualifier (endogenous, exogenous)

For each array that is not generally available (e.g., commercially available), the provided information should be sufficient to reproduce the array and all its design features.

**Experiment description**

By *experiment* MIAME refers to a set of one or more hybridisations that are in some way related (e.g., related to the same publication, study or holistic biological question). The minimum information for an environmental genomics experiment includes a description of the following four parts.

1. Experimental design (experimental type, factors and design parameters)
2. Biological materials used, extract preparation and labeling (biosource properties, pre-conditioning, conditioning and sample processing)
3. Hybridization procedures and parameters
4. Gene expression measurement data and specifications of data processing (raw data description, image analysis and normalization)

*Note:* A potentially reusable part of the experiment description is ‘laboratory protocols,’ including data processing methods (e.g., normalization). MIAME encourages the user to assign unique identifiers to all reusable parts of the experiment description and to reference these when the respective parts are reused (indicating any deviations). A standard for the description of protocols, including the data transformation protocols is being developed by the MGED Society [19].
MIAME/Env recommends the following details on each of these sections:

1. Experimental design

This section is common to all hybridizations performed in the environmental genomic experiment, such as the goal, brief description, experimental factors tested. The following information is included in the experimental design.

1.1 Authors, laboratory, contact
1.2 Type of the experiment. See the MGED ontology, specifically:
   http://mged.sourceforge.net/ontologies/MGEDontology.php#ExperimentDesignType
   
   For instance:
   - multiple tissue comparison
   - temperature shock
   - time course
   - dose response
   - effect of gene knock-out
   - effect of gene knock-in (transgenics)
   - challenged vs. unchallenged
   - geographical location comparison
   - fitness
   - conditioning
   - competition
   - lab vs. field comparison
   - strain/line or ecotype comparison
   - other

(Note that multiple types are possible, e.g. a 'time course' experiment can also be a 'dose-response' type if multiple doses of the compound are tested over time)

1.3 Experimental factors: The parameter or parameters considered as being the main factors under test in the experiment. See the MGED ontology, specifically:
   http://mged.sourceforge.net/ontologies/MGEDontology.php#ExperimentalFactorCategory

   For instance:
• species
• strain
• ecotype
• sex type
• age and weight
• competitors
• cell line
• cell type
• developmental stage
• disease state
• behaviour
• morphology
• genotype
• fitness
• housing
• feeding regime
• native location
• route of exposure
• temperature
• time of treatments and observations
• dose(s) in standard units
• genetic variation
• response to a treatment or compound

1.4 How many hybridisations in the experiment?

1.5 If a common (standard) reference material used for all hybridisations

1.6 Quality control steps taken:
• Replicates done (yes/no), type of replicates, description
• biological
• technical
• if pools of extracts (yes/no) were used versus extracts from individual samples, description
• whether dye swap is used (only for two channel platforms)
• other (e.g., polyA tails, low complexity regions, unspecific binding)
• other.

1.7 A brief description of the experiment and its goal and a link to a publication if one exists
2. Biological materials used, extract preparation and labeling

By biological material (sample) MIAME/Env refers to the material used in the investigations and from which the nucleic acids were extracted for subsequent labelling and hybridization. In this section all steps that precede the hybridization with the array are described. We can usually distinguish between:

- Source of the sample (biosource properties);
- Treatments applied to the samples (manipulations);
- Extract preparation;
- Extract labelling; and
- Hybridization controls.

Below we list the most essential items that are usually needed.

2.1 Biosource properties

**Guidance note:** This section is especially important for experiments such as field studies where an organism is taken directly from the “wild” for experimentation. In this case as much detail as possible should be provided in the sample source location and environmental history sections:

- Organism (NCBI taxonomy, if known. Where the identity of the organism is unknown or a mixed population is under study a unique identifier plus description should be provided.)
- Sample source provider
- Sample source location (If the biological material is taken from the “wild”, a GPS coordinate and associated time stamp, to indicate where and when the organism was harvested, should be provided. For some samples a more detailed topography may also be necessary)
- Environmental history, if known. Information on the environmental history of the biosource prior to sample manipulation. Potentially, including growth conditions and any pertinent environmental parameters.
- Descriptors relevant to the particular sample (if known). See the MGED ontology, specifically: http://mged.sourceforge.net/ontologies/MGEDontology.php#BioMaterialCharacteristics

For instance:

- sex
- age
weights
morphology (size, colour etc)
organism behavioural characteristics
fitness
development stage (of the organism or the host organism)
organism part (tissue) of the organism’ anatomy
from which the biological material is derived (if samples are cells)
cell type
animal/plant strain, line or ecotype

genetic variation (e.g., gene knockout, transgenic variation)
individual genetic characteristics (e.g., disease alleles, polymorphisms)
disease state or normal
an individual identifier (for interrelation of the biological materials in the experiment)

2.2 Sample manipulations: laboratory protocols and relevant parameters.
MIAME/Env requires that measurable parameters are provided with at least a name, description and associated unit (For atypical measures or units, it should be made clear whether the measure is relative or absolute and a full description of the unit should be provided). In unusual or ambiguous cases a more detailed description/protocol should be provided. For example:

In an experiment comparing the response of an organism to temperature change, it should be made clear which temperature is of interest (soil, air, water etc.) and details of its measurement should be given (e.g. temperature at 5 mm soil depth along a cline at a radius of 12cm from the base of the plant).

Pre-conditioning

**Guidance note: This section attempts to capture the animal husbandry or cell culture conditions that are common to all samples within the experiment. For field studies this may be irrelevant or sparsely populated.**

For guidance on terminology see the MGED ontology, specifically: http://mged.sourceforge.net/ontologies/MGEDontology.php#EnvironmentalHistory
Organism husbandry and housing details or cell culture conditions. For example:

- Housing details. For example:
  - Housing type (Make, model etc.)
  - Size
  - Material
  - Shape
- Feeding/cleaning regime
- Growth conditions. For example:
  - Temperature
  - Density
  - Light
- Pre-Conditioning period (i.e. length of time)

**Guidance note: MIAME/Env classifies any change to the animal husbandry regime detailed above as a treatment. Where conditioning is performed in uncontrolled “field” environments, details of the environmental parameters found at the conditioning site should be recorded in a similar fashion to the environmental history captured in biosource properties (section 2.1)**

For examples of controlled vocabularies for use in the following sections see the MGED ontology, specifically:
http://mged.sourceforge.net/ontologies/MGEDontology.php#Action
and
http://mged.sourceforge.net/ontologies/MGEDontology.php#ProtocolPackage

Treatments in vivio, in vitro, measurements, units and detailed experimental protocols on information such as:

- Exposure to chemical stressors. For example:
  - Complex mixtures. For example
    Air/Water samples. Including information on
    - sample source (including relevant environmental parameters)
    - sample analysis (if performed)
  Outflows. Including information on:
  - outflow source (including relevant environmental parameters)
  - outflow analysis (if performed)
Compound / Small molecule. Including information on:
- Treatment compound name, grade, formulation and manufacturer
- Type of compound (e.g. chemical, drug or solvent)
- CASRN, chemical structure/molecular formula
dose (and unit)

Exposure to physical stressors. For example:
- Electromagnetic radiation. Including information on:
  - Type
  - Wavelength
  - Duration
  - Intensity
- Pressure (including depth or height)
- Heat / Cold – shock. Including information on:
  - Duration
  - Temperature

Exposure to biological stressors. For example:
- Parasitic infection
- Introduction of competing organism/s

Perturbation. For example, changes to:
- Housing and or animal husbandry for example
  - Container
  - Diet
  - Cleaning regime
  - Bedding
- Organism density
- Circadian-rhythmicity
- Tidal-cycle

Sample processing
- Separation technique: Of tissues or cells from a heterogeneous sample. Or of a specific organism sub-sample from a mixed population. For example:
  - For tissues and cells
    - none
    - trimming
    - microdissection
• FACS
• other (including description/protocol)

For populations
• size exclusion
• centrifugation
• manual separation (including description/protocol)
  o Method of sacrifice. Including information on:
    date
    time

2.3 Hybridization extract preparation protocol for each extract prepared from the biological material, including
• extraction method
• whether total RNA, mRNA, or genomic DNA is extracted
• amplification (RNA polymerases, PCR)

2.4 Labeling protocol for each labeling prepared from the extract, including
• amount of nucleic acids labeled
• label used (e.g., A-Cy3, G-Cy5, 33P, ...)
• label incorporation method
• Facility details (if this part of the experiments has been carried out in facility different from the sample treatment steps above, e.g. consortium, contracting out.)

2.5 External controls added to hybridization extract(s) (spiking controls)
• element on array expected to hybridize to spiking control
• spike type (e.g., oligonucleotide, plasmid DNA, transcript)
• spike qualifier (e.g., concentration, expected ratio, labeling methods if different than that of the extract)

3. Hybridization procedures and parameters

Each hybridization description should include information about which labelled extract (related to which biological material, which extract) and which array (e.g., array design, batch and serial number) has been used in the experiment; and the hybridization protocol, normally including:
• the solution (erg., concentration of solutes)
• blocking agent
• wash procedure
• quantity of labeled target used
• time, concentration, volume, temperature
• description of the hybridization instruments
4. Measurement data and specifications of data processing

We distinguish between three levels of data processing – raw data (images), image quantitations and gene expression data matrix. Each hybridization has at least one image, each image has a corresponding image quantitation table, where a row represents an array design element and a column to a different quantitation types, such as mean or median pixel intensity. Several quantitation tables can be combined using data processing metrics to obtain the ‘final’ gene expression measurement table associated with the experiment.

1) Raw data description should include
• for each scan laboratory protocol for scanning, including scanning hardware and software, scan parameters, including laser power, spatial resolution, pixel space, PMT voltage;
• scanned images;
   It should be noted that MGED does not have consensus whether the provision of images is a part of MIAME.

2) Image analysis and quantitation
• image analysis software specification and version, availability, and the description or identification of the algorithm and all the parameters used
• for each image the complete image analysis output (of the particular image analysis software)

3) Normalized and summarized data – gene expression data matrix data processing protocol, including normalization algorithm (for detailed recommendations, see http://www.mged.org/normalization) gene expression data table(s) derived from the experiment as the whole, derived measurement value summarizing related elements and replicates as used by the author (this may constitute replicates of the element on the same or different arrays or hybridizations, as well as different elements related to the same entity e.g., gene) providing a reliability indicator for each data point (e.g., standard deviation) is encouraged. Note that toxicology data may also be normalized to signal-, fold- change ratios to facilitate comparison to similarly-expressed microarray data (especially when 2-fluor labelling has been used).

This ends the experiment description.

References

MIAME/Env Glossary

MIAME/Env requirements are listed in alphabetical order and definitions are provided.

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amplification method</td>
<td>The method used to amplify the nucleic acid extracted</td>
</tr>
<tr>
<td>Array design</td>
<td>The layout or conceptual description of array that can be implemented as one or more physical arrays. The array design specification consists of the description of the common features of the array as the whole, and the description of each array design elements (e.g., each spot). There are three levels of array design elements: feature (the location on the array), reporter (the nucleotide sequence present in a particular location on the array), and composite sequence (a set of reporters used collectively to measure an expression of a particular gene)</td>
</tr>
<tr>
<td>Array design name</td>
<td>Given name and version for the array design, that helps to identify a design between others</td>
</tr>
<tr>
<td>Array related information</td>
<td>Description of the array as the whole</td>
</tr>
<tr>
<td>Attachment</td>
<td>How the element (reporter) sequences are physically attached to the array (e.g. covalent, ionic)</td>
</tr>
<tr>
<td>Author, laboratory, and contact</td>
<td>Person(s) and organization (s) names and details (address, phone, FAX, email, URL)</td>
</tr>
<tr>
<td>Biological material (sample)</td>
<td>The source material, on which the investigations have been carried out and from which the nucleic acids have been extracted for subsequent labelling and hybridization.</td>
</tr>
<tr>
<td>Biosource properties</td>
<td>Information on the source of the biological material (e.g. organism, sample provider, sex, age, weights or cell type, cell line)</td>
</tr>
<tr>
<td>Clone information</td>
<td>For each reporter, the identity of the clone along with information on the clone provider, the date obtained, and availability</td>
</tr>
<tr>
<td>Common (standard) reference material</td>
<td></td>
</tr>
<tr>
<td>Composite sequence</td>
<td>The set of reporters used collectively to measure an expression of a particular gene, exon, or splice-variant</td>
</tr>
<tr>
<td>Conditioned field trials</td>
<td>Experiments in which an organism is taken from the wild and “conditioned” in a lab environment under specified animal husbandry conditions prior to sample manipulation/treatments</td>
</tr>
<tr>
<td>Conditioning</td>
<td>The period in which the “experimentation” itself is performed. Usually recorded as a list of treatment protocols including changes to animal husbandry/housing.</td>
</tr>
<tr>
<td>Control elements (array)</td>
<td>Array elements that have an expected value and/or are used for normalization</td>
</tr>
<tr>
<td>Control qualifier (array)</td>
<td>Terms used to further define a control element (e.g. endogenous, exogenous)</td>
</tr>
<tr>
<td>Control type (array)</td>
<td>The type of control used for the normalization (e.g. spiking, normalization, negative, positive)</td>
</tr>
<tr>
<td>Dimensions (feature)</td>
<td>The physical dimensions of each features</td>
</tr>
<tr>
<td>Element (array)</td>
<td>There are three levels of array design elements: feature (the location on the array), reporter (the nucleotide sequence present in a particular location on the array), and composite sequence (a set of reporters used collectively to measure an expression of a particular gene)</td>
</tr>
<tr>
<td>Element generation protocol</td>
<td>A description of how the reporters were generated</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Experiment</td>
<td>A set of one or more hybridizations that are in some way related (e.g., related to the same publication or study).</td>
</tr>
<tr>
<td>Experimental factor</td>
<td>The factors in the study that are experimental parameters or regarded as influencing the experimental results.</td>
</tr>
<tr>
<td>Extraction method</td>
<td>Documentation of the set of steps taken to extract nucleic acids from the biological material.</td>
</tr>
<tr>
<td>Feature (array)</td>
<td>A specific instance of a reporter located on an array, commonly referred to as a spot.</td>
</tr>
<tr>
<td>Field Trial</td>
<td>Experiments in which an organism is taken from the “wild” for use in the experiment without preconditioning or conditioning.</td>
</tr>
<tr>
<td>Final gene expression table(s)</td>
<td>Derived measurement value summarizing related elements and replicates, providing the type of reliability indicator used.</td>
</tr>
<tr>
<td>Hybridization</td>
<td>The process of treating an array with one or more labelled extracts under a specified set of conditions.</td>
</tr>
<tr>
<td>Hybridization extract preparation</td>
<td>Information on the extract preparation for each extract prepared from the sample, the type of extract (e.g. total RNA, mRNA) and the amplification method used (e.g. RNA polymerases, PCR)</td>
</tr>
<tr>
<td>Hybridization protocol</td>
<td>Documentation of the set of steps taken in the hybridization, including: solution (e.g. concentration of solutes); blocking agent and concentration used; wash procedure; quantity of labelled target used; time; concentration; volume, temperature, and description of the hybridization instruments</td>
</tr>
<tr>
<td>Image analysis and quantitation</td>
<td>Each image has a corresponding image analysis output from the particular image analysis software. In an image quantitation table a row represents a array design element and a column correspond to a different quantitation types (e.g. mean or median pixel intensity). This has an associated image analysis protocol, the set of steps taken to quantify the image including: the image analysis software, the algorithm and all the parameters used</td>
</tr>
<tr>
<td>Label incorporation method</td>
<td>The method used to incorporate the label into the extracts</td>
</tr>
<tr>
<td>Label used</td>
<td>The type of the label used (e.g. A-Cy3, G-Cy5, 33P)</td>
</tr>
<tr>
<td>Lab experiment</td>
<td>Experiment in which the biosource is lab reared or obtained from a standard provider, may include preconditioning and or conditioning</td>
</tr>
<tr>
<td>Location (feature)</td>
<td>The abstract coordinate of each features on the array</td>
</tr>
<tr>
<td>Measurements</td>
<td>MIAME distinguishes between three levels of data processing: image (raw data), image analysis and quantitation, gene expression data matrix (normalized and summarized data)</td>
</tr>
<tr>
<td>Normalized and summarized data</td>
<td>Several quantitation tables are combined using data processing metrics to obtain the ‘final’ gene expression measurement table (gene expression data matrix) associated with the experiment. This has an associated data processing protocol, including normalization algorithm and a reliability indicator for each data point (e.g., standard deviation). Toxicology data may also be normalized to signal-, fold- change ratios to facilitate comparison to similarly-expressed microarray data especially for two channel experiments</td>
</tr>
<tr>
<td>Numerical biological endpoint data</td>
<td>Biological endpoint with quantitative measurements (e.g. clinical pathology)</td>
</tr>
<tr>
<td>Preconditioning</td>
<td>The period immediately after the organism is obtained (either from the wild or a provider), including the animal husbandry and housing details common to all samples in the experiment.</td>
</tr>
<tr>
<td>Physical dimensions (array)</td>
<td>The physical dimension of the array support (e.g. of slide)</td>
</tr>
<tr>
<td>Platform type (array)</td>
<td>The technology type used to place the biological sequence on the array</td>
</tr>
<tr>
<td>Term</td>
<td>Description</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Position (control elements)</td>
<td>The abstract coordinate of each the control features on the array</td>
</tr>
<tr>
<td>Primer information</td>
<td>For each reporter, PCR primers pair sequence and length</td>
</tr>
<tr>
<td>Production protocol (array)</td>
<td>A description of how the custom made array was manufactured</td>
</tr>
<tr>
<td>Qualifier, value, source</td>
<td>Describe any further information about the array or an experiment in a structured manner</td>
</tr>
<tr>
<td>Quality control steps</td>
<td>Measures taken to ensure or measure quality, e.g., dye swap (for two channel platforms), biological replicate (samples taken from more than one organism), technical replicate (same biological material hybridized to more than one array).</td>
</tr>
<tr>
<td>Raw data</td>
<td>Each hybridization has at least one image and associated laboratory protocol for scanning, hardware and software, scan parameters (including laser power, spatial resolution, pixel space and PMT voltage)</td>
</tr>
<tr>
<td>Reference sequence (reporter)</td>
<td>The sequence from which the reporter (e.g. oligo-nucleotides) have been derived</td>
</tr>
<tr>
<td>Reporter</td>
<td>The nucleotide sequence present in a feature. The same reporter can be also used at multiple features to generate replicate measurements.</td>
</tr>
<tr>
<td>Reporter type</td>
<td>Physical nature of the reporter (e.g. PCR product, synthesized oligo-nucleotide, plasmids, colonies)</td>
</tr>
<tr>
<td>Scanner image file</td>
<td>The TIFF file including header</td>
</tr>
<tr>
<td>Sequence information (reporter)</td>
<td>The nucleotide sequence information for reporter: sequence accession number (from DDBJ/EMBL/GenBank), the sequence itself (if known) or a reference sequences (e.g. for oligo-nucleotides)</td>
</tr>
<tr>
<td>Single or double stranded (reporter)</td>
<td>Whether the reporter sequences are single (oligo-nucleotide) or double stranded (e.g. PCR product)</td>
</tr>
<tr>
<td>Spike type and qualifier</td>
<td>The type of spike used (e.g. oligonucleotide, plasmid DNA, transcript) and its qualifier (e.g. concentration, expected ratio, labelling methods)</td>
</tr>
<tr>
<td>Spiking control (hybridization)</td>
<td>External controls added to the hybridization extract (s)</td>
</tr>
<tr>
<td>Spiking control element (array)</td>
<td>Position of the feature (s) on the array expected to hybridize to the spiking control</td>
</tr>
<tr>
<td>Surface and coating specification (array)</td>
<td>Type of surface and name for the type of coating used</td>
</tr>
<tr>
<td>Textual biological endpoint data</td>
<td>Biological endpoint with nominal measurements (e.g. clinical observations and gross necropsy examination)</td>
</tr>
<tr>
<td>Unique identifier</td>
<td>An alphanumeric code or unique name given to a particular item that whose precise identity is difficult to determine (provides provenance if analysis is performed at a later date). Unique identifiers are often used for samples containing mixed species populations or where the identity of a biosequence printed onto an array is unknown</td>
</tr>
<tr>
<td>Environmental genomics experiment description</td>
<td>MIAME/Env distinguishes between: the environmental genomic experiment design (the design, purpose common to all hybridizations performed in the experiment and the parameters), the biological material used (biological material characteristics, the extract preparation and the labelling), the hybridization (procedures and parameters) and the gene expression data (measurements and specifications)</td>
</tr>
<tr>
<td>Environmental genomics experimental design</td>
<td>Design and purpose common to all hybridizations performed in the experiment, including a brief description, person(s), organization (s) names and details, type of the experiment, experimental factors, number of hybridizations, quality controls and link to a publication.</td>
</tr>
<tr>
<td>Type of experiment</td>
<td>A controlled vocabulary that classify an experiment (e.g. Geographical comparison, competition)</td>
</tr>
</tbody>
</table>