The MIAME Checklist – update January 2005

The purpose of this checklist is to guide authors, journal editors and referees in helping them to ensure that the data supporting published results for array based chromatin immunoprecipitation (so-called ChIP-on-Chip) experiments are made publicly available in a format that enables unambiguous interpretation of the data and potential verification of the conclusions (see [1]). For more detail regarding the rationale of MIAME see [2]. MGED strongly recommends that the data is made publicly available through one of the public repositories for microarray data (see [3]).

Experiment Design:

- The goal of the experiment – one line maximum (e.g., the title from the related publication)
- A brief description of the experiment (e.g., the abstract from the related publication)
- Keywords, for example, *time course, cell type comparison, array CGH* (the use of MGED ontology terms is recommended).
- Experimental factors - the parameters or conditions tested, such as *growth media (for yeast), or genetic variation ????* (the use of MGED ontology terms is recommended).
- Experimental design - relationships between samples, treatments, extracts, labeling, and arrays (e.g., a diagram or table).
- Quality control steps taken (e.g., replicates or dye swaps).
- Links to the publication, any supplemental websites or database accession numbers.

Samples used, extract preparation and labelling:

- The origin of each biological sample (e.g., name of the organism, the provider of the sample) and its characteristics (e.g., gender, age, developmental stage, strain, or disease state).
- Manipulation of biological samples and protocols used (e.g., growth conditions, treatments, separation techniques).
- Experimental factor value for each experimental factor, for each sample (e.g., '*growth media = glucose rich???*' for a sample in a time course experiment).
- Technical protocols for preparing the hybridization extract (e.g., the RNA or DNA extraction and purification protocol), labeling, *????, crosslinkage?????*
- External controls (spikes), if used.

Hybridization procedures and parameters:

- The protocol and conditions used for hybridization, blocking and washing, including any post-processing steps such as staining

Measurement data and specifications:

- Data

    - The raw data, i.e. scanner or imager and feature extraction output (providing the images is optional). The data should be related to the respective array designs (typically each row of the imager output should be related to a feature on the array – see Array Designs).
    - The processed data, i.e., set of quantifications from several arrays upon which the authors base their conclusions (*for instance, p-values for protein binding???*). The data should be related to the respective array designs (typically each row of the summarized data will be related to one biological annotation, such as a gene name).
- Data extraction and processing protocols,
    - Image scanning hardware and software, and processing procedures and parameters.
    - *Data processing procedures and parameters, such as p-value selection????.*

Array Design:

- General array design, including the platform type (whether the array is a spotted glass array, an *in situ* synthesized array, etc.); surface and coating specifications and spotting protocols used (for custom made arrays), or product identifiers (the name or make, catalogue reference numbers) for commercially available arrays.
- Array feature and reporter annotation, normally represented as a table (for instance see Tables 1, 2 below), including
    - For each feature (spot) on the array, its location on the array (e.g., metacolumn, metarow, column, row) and the reporter present in the location (note that the same reporter may be present on several features).
    - For each reporter unambiguous characteristics of the reporter molecule, including
        - Reporter role – control or measurement
        - The sequence for oligonucleotide based reporters
        - The source, preparation and database accession number for long (e.g., cDNA or PCR product based) reporters
        - Primers for PCR product based reporters
    - Appropriate biological annotation for each reporter, for instance a gene *genomic region, downstream gene???* (note that different reporters can have the same biological annotation)
- Principle array organism(s)

Table 1. Oligonucleotide array description file example:

Feature				Reporter						Biological annotation			
Coordinates on Array				Reporter ID (user defined) Oligo ID	Biosequence Type	Sequence	DDBJ/ EMBL/ Genbank	Reporter Usage	Control Type	ID	Designation	Related Gene Symbol, if appropriate	Database Entry
Meta Col	Meta Row	Col	Row										
1	1	1	1	Cy3Cy5	Oligo	AAAAAAAAAAAAAAAAAA	_	Control	Positive	C001_01	Labeled oligo	_	_
1	1	2	1	M00868_01	Oligo	ACCAGCAGATACCTCCTTG	D83002	Experimental	_	C002_01	Gene	ALK	LocusID 11682
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
4	6	10	8	M00264_01	Oligo	ATGTCCGTTGAATTGG	D83002	Experimental	_	C002_01	Gene	ALK	LocusID 11682
…	…	…	…	…	…	…	…	…	…	…	…	…	…
4	6	11	8	M02404_01	Oligo	AGTGGCGAGGAGGAGGAC	L11065	Experimental	_	C449_01	Gene	OPRK1	LocusID 18387
4	6	12	8	M03172_01	Oligo	CCACCACCAAGACCTACTCC	U34891	Experimental	_	C450_01	Gene	KLRA9	LocusID 16640

Table 2. cDNA array description file example:

Feature				Reporter						Biological annotation			
Coordinates on Array				Reporter ID (user defined) HGMP Ref	Biosequence Type	Clone ID	DDBJ/ EMBL/ Genbank	Reporter Usage	Control Type	ID	Designation	Related Gene Symbol	Database Entry
Meta Col	Meta Row	Col	Row										
1	1	1	1	370503	cDNA clone	IMAGE 32017	R17905	Experimental	_	C1	Gene	FNTA	LocusID2339
1	1	2	1	370504	cDNA clone	IMAGE 2962831	BC005866	Experimental	_	C2	Gene	MLH1	LocusID 4292
1	1	3	1	370505	Genomic clone	Cosmid 9H11	L40416	Control	Positive	_	_	_	_
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮			⋮	⋮
4	8	24	12	380696	cDNA clone	IMAGE 5214483	BC028215	Experimental	_	C285	Gene	PTEN	LocusID 5728

References

[1] Cech, T.R., Sharing Publication-Related Data and Materials: Responsibilities of Authorship in the Life Sciences. Available at www.nap.edu/books/0309088593/html

[2] A. Brazma, P Hingamp, J Quackenbush, G Sherlock, P Spellman, C Stoeckert, J Aach, W Ansorge, C A Ball, H C Causton, T Gaasterland, P Glenisson, F C P Holstege, I F Kim, V Markowitz, J C Matese, H Parkinson, A Robinson, U Sarkans, S Schulze-Kremer, J Stewart, R Taylor, J Vilo & M Vingron. Minimum information about a microarray experiment (MIAME)—toward standards for microarray data, Nature Genetics, vol 29 (December 2001), pp 365 - 371.

[3] C.A. Ball, A. Brazma, H. Causton, S. Chervitz, R. Edgar, P. Hingamp, J.C. Matese, H. Parkinson, J. Quackenbush, M. Ringwald, S. Sansone, G. Sherlock, P. Spellman, C. Stoeckert, Y. Tateno, R. Taylor, J. White and N. Winegarden. Submission of Microarray Data to Public Repositories. PLoS Biology, 2, e317: 1276-1277.