The MIAME Checklist for Dataset C*

[* Dataset C was generated by having the individual laboratories (n=7) apply a standardized image analysis software and set of feature extraction parameters to the raw microarray images that were produced for Dataset B. Thus, all elements of the MIAME checklist for Dataset C are the same as for Datasets A and B, with the exception of the image analysis. All Supplementary information, including materials, methods, and primary data can be accessed at http://www.niehs.nih.gov/research/atniehs/core/microarrays/trc/.]

Experiment Design:

- **Type of experiment**: The experiment was designed to determine the reproducibility of gene transcript level measurements between multiple laboratories (n=7) for a single microarray platform called the Standard Spotted Array. Each laboratory was provided with large aliquots from two different mouse standard RNA samples: a liver RNA sample and a five-tissue pooled sample (liver, kidney, lung, brain, spleen in equal amounts) that were prepared in one of the consortium laboratories (see URL for Supplementary Methods, Standard Protocols, Tissue Extraction and Standard RNA Sample Preparation. Each laboratory was provided *Standard Spotted Arrays* that were made in Laboratory 1 by spotting 70-mer oligos (Operon) representing 18,000 mouse genes onto poly-l-lysine coated glass slides.
- **Experimental factors**: Using the standard RNA samples allowed us to focus on variation in the technical and analytical approaches to microarray experimentation, such labeling and hybridization protocols and image analysis, rather than the biological variation.
- **The number of hybridizations performed in the experiment.** Each of the 7 laboratories carried out 8 hybridizations; four that co-hybridized liver RNA versus liver RNA (LvsL) and four that co-hybridized liver RNA versus pooled RNA (LvsP); each set of four hybridizations had two dye-swapped samples (Cy3 versus Cy5).
- **The type of reference used for the hybridizations**: See answer to "Type of experiment" (above).
- **Hybridization design**: See answer to "Type of experiment" and "The number of hybridizations" (above).
- **Quality control steps taken**: See answer to "The number of hybridizations" (above).
- **URL of any supplemental websites or database accession numbers**: Supplemental information, including materials, methods and primary data can be accessed at http://www.niehs.nih.gov/research/atniehs/core/microarrays/trc/.

Samples used, extract preparation and labeling:

- **The origin of the biological sample:** For tissue extraction, twenty-five C57 black male mice (10-11 weeks) were purchased from Taconic Farms and housed five in a cage in pathogen-free rooms determined by sentinel animal survey. They acclimated for 7-10 days until they reached 12 weeks of age. They were housed in shoebox cages on sawdust and fed NTP 2000 chow; RNA isolation from tissue was performed by one laboratory (see URL for Supplementary Methods, Standard Protocols, Tissue Extraction and Standard RNA Sample Preparation).

- **Manipulation of biological samples and protocols used**: See URL for Supplementary Methods, Standard Protocols, Tissue Extraction and Standard RNA Sample Preparation.
- **Protocol for preparing the hybridization extract**: RNA isolation from tissue was performed by one laboratory and distributed to the other six participating laboratories (see URL for Supplementary Methods, Standard Protocols, Tissue Extraction and Standard RNA Sample Preparation).
- **Labeling protocol(s)**. The individual laboratories were provided the standard RNA samples and used a variety of resident (in-house) protocols for RNA labeling (see URL for Supplementary Methods, Supplementary Table 4 and Resident Protocols for Laboratories 1-7).
- **External controls (spikes)**. cRNA control *in vitro* Arabidopsis transcripts (70mers) provided by The Institute for Genomics Research (http:pga.tigr.org) were added to each of the standard RNA samples as a means of assessing experimental variability for individual RNA analyses (labeling and hybridizations). A total of 10 Arabidopsis genes were chosen for this spike in design, with fixed concentrations of each transcript in two distinct sets, A3 and A5 (see URL for Supplementary Methods, Standard Protocols, Quality Control Genes).

Hybridization procedures and parameters:

- **The protocol and conditions used during hybridization, blocking and washing**. The individual laboratories were provided the standard RNA samples and used a variety of resident (in-house) protocols for hybridizations to the Standard Spotted Array (see URL for Supplementary Methods, Supplementary Table 4 and Resident Protocols for Laboratories 1-7).

Measurement data and specifications:

- **The quantitations based on the images**: The individual laboratories used a standardized software package (Axon GenePix Pro software v 4.1.1.28) and set of feature extraction parameters for re-analyzing the raw images from the Standard Spotted Array (Dataset B) to generate Dataset C. Gene expression florescence intensities (of red and green channels) were measured for all spots represented on the Standard Spotted Array (see URL for Supplementary Methods, Standard Protocols, Image Analysis for Standard Spotted Arrays).
- **The set of quantitations from several arrays upon which the authors base their conclusions.**
  - **Type of scanning hardware and software used**: The individual laboratories used a variety of resident (in-house) equipment for scanning the Standard Spotted Arrays (see URL for Supplementary Methods, Supplementary Table 7 and Resident Protocols).
  - **Type of image analysis software used**. The laboratories used a standardized image analysis software package (Axon GenePix Pro software v 4.1.1.28) and set of feature extraction parameters for re-analyzing the raw images for the Standard Spotted Array (Dataset B) to generate Dataset C (see URL for Supplementary Methods, Standard Protocols, Image Analysis for Standard Spotted Arrays).
  - **A description of the measurements produced by the image-analysis software and a description of which measurements were used in the analysis**. The measurements produced using the Axon GenePix Pro software (v 4.1.1.28) and used in the analyses include foreground intensity mean (F532 and F635 mean) and background intensity median (B532 and B635 median).

- **The complete output of the image analysis *before* data selection and transformation**. See URL for Supplementary Information, Download the Primary Data, Raw Data for Dataset C.
- **Data selection and transformation procedures**. Four different versions of the standard liver vs. pooled RNA sample (LvsP) data were used to generate Dataset C. The "raw" data obtained from the GenePix Pro image processing software package were used to generate four normalized datasets by applying: (1) Global Intensity Normalization, (2) Global Intensity Normalization with background adjustment, (3) Lowess Normalization with background adjustment applied to a $\log_2$-ratio versus $\log_2$-geometric-mean-intensity (R-I or M-A) plot, and (4) Lowess Normalization without background subtraction applied to an R-I plot (see Methods).
- **Final gene expression data table(s) used by the authors to make their conclusions *after* data selection and transformation**. See URL for Supplementary Information, Download the Primary Data, Raw Data for Dataset C.

Array Design:

- **General array design.** The Standard Spotted Array was manufactured by Laboratory 1 using sequence-verified mouse oligonucleotides (70mers) representing 18K unique mouse gene elements purchased from Operon and spotted onto poly-l-lysine coated slides using a GeneMachine OmniGrid Arrayer**.**
- **For each feature (spot) on the array, its location on the array and the ID of its respective reporter**. Microarray slides are printed on a Gene Machine OmniGrid Arrayer and Axon GenePix 4000B scanner, and the format dictated from these two files refers to the location of the spots.
- **For each reporter, its type.** Sequence-verified mouse oligonucleotides (70mers) representing 18K unique mouse gene elements were purchased from Operon. Arabidopsis 70mer probes corresponding to the cRNA Arabidopsis control *in vitro* transcripts (spike ins) were spotted onto the slides. Sequence information is available from The Institute for Genomic Research (http:pga.tigr.org).
- **For commercial arrays**. Not applicable.
- **For non-commercial arrays, the following details should be provided**:
- **The source of the reporter molecules**: Sequence-verified mouse oligonucleotides (70mers) representing 18K unique mouse gene elements were purchased from Operon. Arabidopsis 70mer probes corresponding to the cRNA Arabidopsis control in vitro transcripts were provided by TIGR (http:pga.tigr.org). Stratagene Spot report was used.
- **The method of reporter preparation**. To be provided, if accepted.
- **The spotting protocols used**. The Gene Machine OmniGrid Arrayer was used to spot the slides**.**
- **Any additional treatment performed prior to hybridization**.