



Microarrays Inc

Data Sheet

38.5K Mouse Genomic Array

For the production of our Mouse Genomic Array, MI utilizes oligonucleotide probes from the Illumina, Inc.-MEEBO (Mouse Exonic Evidence Based Oligonucleotide) collaboration. This optimized collection of 70mer probes was largely derived from constitutively expressed exons, and allows for the interrogation of nearly 25,000 mouse genes. An exon-centric probe design was used for the differentiation of constitutively expressed versus alternatively expressed exons. This design also supports comparative genome hybridization (CGH) analysis. In addition to the exon-centric probes, these arrays contain an extensive assortment of controls that facilitate accurate evaluation of expression results. MI's mouse array enables the study of transcription patterns and, as broadly as possible, alternative splicing.

The 38.5K arrays contain 35,302 gene specific 70mer oligonucleotide probes. In addition, each array includes 3,482 controls (see below). All of the oligonucleotide probes have been T_m normalized and are designed to minimize cross-hybridization and secondary structure.

Under licensure from Oxford Gene Technology and in combination with our own proprietary printing technology, MI is able to precisely deposit these 38,467 probes onto a single 1" x 3" standard-sized slide. MI employs a proprietary slide chemistry and buffer combination offering high DNA-binding capacity, deposition uniformity and stability. Also, in an effort to optimize results and reduce background, all arrays are surface inactivated post-printing and packaged in special materials chosen to reduce the out-gassing associated with most plastic containers and packaging. MI's attention to detail and quality control verification of every lot insures dependable product consistency between production lots and experiments. Included with each array are detailed quality control production data, an optimized sample labeling/hybridization protocol, and a CD containing annotation and array design data.

MOUSE PROBES

The array contains a total of 35,302 probes targeting mouse genes. The probes fall into the following categories:

- **Constitutive Exonic Probes (30,125 probes):** A probe that will recognize all known transcripts of a gene.
- **Alternatively Spliced / Skipped Exonic Probes (4,201 probes):** Probes that will recognize exons that are present in some, but not all transcripts of a gene.
- **Non-Coding RNA Probes (196 probes):** Probes recognizing non-protein coding transcripts (ribosomal RNAs, miRNAs).
- **BCR / TCR Genic / Regional Probes (372 probes):** Probes recognizing transcripts from genes that undergo somatic rearrangement.
- **Mitochondrial Probes (13 probes):** Probes recognizing mouse mitochondrion derived DNA sequences.
- **Transgenic / Cassette Probes (37 probes):** Probes recognizing elements commonly used for transgenic constructs (e.g. GFP, YFP).
- **Murine Viral Probes (358 probes):** Probes recognizing mouse viral pathogen sequences.

CONTROLS

The array contains a total of 3,482 controls. The controls fall into the following categories:

- **Negative Controls (317 empty wells and 97 probes):** Empty wells and 97 random sequences are positioned throughout the set to assist in determining background.
- **Positive Controls (1,152 probes):** Probes recognizing a small subset of mouse transcripts.
- **Doped Controls (1,916 probes):** Probes recognizing non-mouse sequences that can be spiked into RNA samples.

DESIGN APPROACH

Mouse Sequence Selection and Probe Design

A systematic methodology was applied to identify the exons, generate all possible 70mer candidate probes, and select the optimal probe from the candidates. The pipeline for sequence selection and probe design included three steps:

1. Collect and curate exon sequences, supplement

as needed with transcript sequences.

2. Design the candidate 70mer probes for exon or transcript sequences.

- ArrayOligoSelector, an open source tool for selecting 70mer oligo probes from a defined set of sequence data, was used to generate a list of candidate probes for each exon or transcript sequence.

- Multiple filters including uniqueness, selfbinding, complexity, GC, content, and user defined parameters were used to narrow and rank the list of candidate probes.

3. Pick the best probe from the list of candidate probes. Several criteria were used to identify the optimal probe from the list of candidate probes:

- **Uniqueness:** Probes that had binding energies of > -35 kcal / mol for other sequences were preferred.
- **3' Proximity:** Probes that were less than 1,000 bases from the 3' end of the transcript were preferred.
- **Constitutive:** Probe should be present in all transcripts.

If the above three criteria could not be met, more than one probe would be selected for the exon or transcript sequence.

ANNOTATION DETAILS

Probe annotation information is provided in the MI_MOUSE_ARRAY.XLS file found on the included CD. The file contains the following data elements:

- **Probe_Name:** Contains the oligo name. The following codes can assist with interpreting the oligo name:

- Rockefeller MouSDB3 constitutive exons / islands (oligo names start with 'scl' followed by a number >0)

- LocusLink constitutive exons / islands (oligo names start with 'scl0' followed by a number >0)

- mRNA derived 70mers which may span intron/exon boundaries (oligo names start with 'scl00' followed by a number >0)

- A collection of alternative spliced / skipped exons generated through extensive curation of 5 published datasets by Max Diehn, Ash Alizadeh, Jean Yang, and Catherine Foo (oligo names start with 'scl000' followed by a number >0)

- Syntenic orthologs of human loci exhibiting cis-antisense transcription based on Yelin et al *Nature Biotech* 2003.

- (oligo names start with 'scl0000' followed by a number >0)

- **LocusLink_ID:** LocusLink ID number.
- **Gid:** GI number.
- **Accession:** Accession number.
- **Symbol:** Gene symbol.
- **Probe Sequence:** 70mer oligo probe sequence.
- **T_m:** Oligo melting temperature.
- **GC:** Percentage of GC content within the probe sequence.

- **Product:** Probe target definition.

- **Description:** The specific annotation for each 70mer is detailed and is caret (^) delimited as follows for MCC plates:

- MCC Details: [LLID (060504) ^ Design type ^ cluster ^ Long Oligoname ^ outputset ^ worstxhyb ^ oligo_3_marg]

- Where LLID refers to LocusLink ID, which can be linked with gene annotation data through a variety of tools including BatchSource.

OLIGONUCLEOTIDE SYNTHESIS AND QUALITY CONTROL

Oligos were synthesized on Illumina's proprietary Oligator® synthesis platform. Illumina utilizes multiple quality control methods to assess oligo quality including:

- **Real Time Digital Trityl Monitoring:** Illumina monitors coupling success of each base addition for every oligo in every plate synthesized. Monitoring is performed real-time during the synthesis process using its proprietary digital trityl monitoring system.

- **Capillary Electrophoresis (CE):** Illumina's uses CE to achieve single base resolution of synthesis success for long oligos; precise abundance measurements and stepwise coupling efficiency values are derived from CE.

- **OD260 Analysis:** Illumina uses OD260 analysis (absorbance measurement) to quantify oligo yield. www.illumina.com

ARRAY MANUFACTURE AND QUALITY CONTROL

All MI arrays are manufactured within HEPA filtered, environmentally controlled clean-rooms. Every production lot is quality control validated:

- **Slide chemistry:** MI utilizes proprietary slide chemistry and an optimized spotting buffer for the manufacture of its arrays. This chemistry/buffer combination provides high DNA binding capacity, deposition uniformity and long-term array stability.

- **Sample deposition and morphology:** Every array is guaranteed to have a minimum of 99.5% of the samples properly deposited and coupled to the substrate. The resultant QC data denotes the exact array location of any suspect depositions, i.e. beyond diameter specifications, irregular morphology, etc.

- **Spatial accuracy:** Every production lot is accessed for spatial accuracy, absolute placement of subgrids, deposition pitch, row and column spacing.

- **Probe/substrate coupling efficiency.** MI uses Veriprobe™, a proprietary quantification technology to monitor attachment of the probe to the substrate. All production lots come with probe attachment data and histogram to insure quality hybridization results.

- **Slide processing:** Following probe deposition, arrays are surface deactivated and processed to remove unbound DNA. The resultant surface has

exceptionally low affinity for non-specific DNA attachment and resists various airborne and buffer-borne contaminants associated with high background levels.

Packaging: Arrays are shipped in a resealable, desiccated storage packs. These materials were specially chosen to reduce out-gassing of the packaging components. Volatiles within packaging components have been associated with reduced shelf-life and high background levels.

CREDITS

MI and Illumina wish to thank the following individuals for their effort in defining and designing the probes on this array:

University of California, San Francisco: Ash Alizadeh, Mike Hagen, Catherine Foo, Jess Leber, David Erle, Jean Yang, Andrea Barczak, Joseph DeRisi, Jing Zhu

Stowers Institute: Chris Seidel

Stanford University: Pat Brown, Max Diehn, Kate Rubins, Stephen Popper, Joseph Marquis, Mike Fero, John Coller, Nicki Chin, Elena Seraia, Peng Zhang, Jon Pollack, Young Kim

Rockefeller University: Terry Gaasterland, Alexey Novoradovsky, Ben Snyder

University of Basel, Switzerland: Mihaela Zavolan

CONTACTING MI

For questions or support, please call 866.627.7297 (toll free) or email us at info@microarrays.com



For Research use only

Microarrays, Inc.
2104 Broadway, Suite 260
Nashville, TN 37203
1-866-627-7297 (toll free)
www.microarrays.com