Microarray hybridizations and analyses

A human cDNA microarray (ToxChip version 1.0), developed in-house at NIEHS, was used for gene expression profiling experiments (1). A complete listing of the genes on this chip is available at: http://dir.niehs.nih.gov/microarray/chips.htm. cDNA microarray chips were prepared according to DeRisi et.al (2). The spotted cDNAs were derived from a collection of sequence verified IMAGE clones that covered the 5' end of the gene and ranged in size from 500 to 2000 base pairs (Incyte Genomics, Palo Alto, CA). M13 primers were used to amplify insert cDNAs from purified plasmid DNA in a 100µl PCR reaction mixture. A sample of the PCR products (10 µl) was separated on 2% agarose gels to ensure quality of the amplifications. The remaining PCR products were purified by ethanol precipitation, resuspended in ArrayIt buffer (Telechem, San Jose CA) and spotted onto poly-L-lysine coated glass slides using a modified, robotic DNA arrayer (Beecher Instruments, Bethesda MD). Each total RNA sample (30-75 µg) was labeled with Cyanine 3 (Cy3) or Cyanine 5 (Cy5)-conjugated dUTP (Amersham, Piscataway, NJ) by a reverse transcription reaction using the reverse transcriptase, SuperScript (Invitrogen, Carlsbad, California), and an Oligo (dT) primer (Amersham, Piscataway, NJ). The fluorescently labeled cDNAs were mixed and hybridized simultaneously to the cDNA microarray chip. Each RNA pair (control vs. estrogen- or tamoxifen-treated) was hybridized to a total of 4 arrays employing a fluor reversal accomplished by labeling the control sample with Cy3 in 2 hybridizations and with Cy5 in the other 2 hybridizations. The cDNA chips were scanned with an Axon Scanner (Axon Instruments, Foster City CA) using independent laser excitation of the two fluors at 532 and 635 nm wavelengths for the Cy3 and Cy5 labels, respectively. A custom script has been implemented in the Axon software to allow autobalancing of the two channels (Tucker *et al.*, unpublished).

The raw pixel intensity images were analyzed using the ArraySuite v1.3 extensions of the IPLab image processing software package (Scanalytics, Fairfax, VA). This program uses methods that were developed and previously described by Chen *et al.* to locate targets on the array, measure local background for each target and subtract it from the target intensity value, and to identify differentially expressed genes using a probability-based method (3). We measured the pixel intensity level of "blank" spots comprised of spotting solution. The data was then filtered to provide a cut off at the intensity level just above the blank measurement values to remove from further analyses those genes having one or more intensity values in the background range. After pixel intensity determination and background subtraction, the ratio of the intensity of the treated cells to the intensity of the control was calculated. The ratio intensity data from the 84 housekeeping genes (or, occasionally, all genes) printed on the ToxChip was used to fit a probability distribution to the ratio intensity values and estimate the normalization constants (*m* and *c*) that this distribution provides. The constant *m*, which provides a measure of the intensity gain between the two channels, indicated that the channels were approximately balanced near a value of 1.0. For each array, the ratio intensity values were normalized to account for the imbalance between the two fluorescent dyes by dividing the ratio intensity value by *m*. The other constant, *c*, estimates the coefficient of variation for the intensity values of the two samples. All arrays in this analysis had a *c* value of 0.191 or less. The probability distribution that is fit to the data was used to calculate a 95% confidence interval for the ratio intensity values. Genes having normalized ratio intensity values outside of this interval were considered significantly differentially expressed.

For each of the 4 replicate arrays for each sample, lists of differentially expressed genes at the 95% and 99% confidence levels were created and deposited into the NIEHS MAPS

database (4). For each time point, a query of the database yielded a list of genes that were differentially expressed in at least 3 of the 4 replicate experiments. Any of these genes that indicated fluor bias or high variation were not considered for further analysis. Assuming that the replicate hybridizations are independent, a calculation using the binomial probability distribution indicated that the probability of a single gene appearing on this list when there was no real differential expression is approximately 0.00048.

1. **Nuwaysir EF, Bittner M, Trent J, Barrett JC, Afshari CA** 1999 Microarrays and toxicology: the advent of toxicogenomics. Mol Carcinog 24:153-9.
2. **DeRisi J, Penland L, Brown PO, et al.** 1996 Use of a cDNA microarray to analyse gene expression patterns in human cancer. Nat Genet 14:457-60.
3. **Chen Y, Dougherty E, Bittner M** 1997 Ratio-based decisions and the quantitative analysis of cDNA microarray images. J Biomedical Optics 2:364-374
4. **Bushel PR, Hamadeh H, Bennett L, et al.** 2001 MAPS: a microarray project system for gene expression experiment information and data validation. Bioinformatics 17:564-5.