Ocean Ridge Biosciences
10475 Riverside Dr. Suite 1
Palm Beach Gardens, FL 33410
Tel: 561-223-3152
Cell: 561-427-5548
Fax: 561-740-8710

ORB Materials and Methods

Gene Expression Profiling

Microarray

Tissue samples in RNAlater / Total RNA / Cultured cells were sent to Ocean Ridge Biosciences (ORB, Palm Beach Gardens, FL) for analysis using Human/Mouse exonic evidence-based oligonucleotide (HEEBO/MEEBO/MI-Rat) microarrays (Lot XXXXX). HEEBO/MEEBO/MI-Rat microarrays were printed by Microarrays Inc. (Nashville, Tennessee) and contained XX,XXX 70-mer oligonucleotides probes complementary to constitutive exons of most human/mouse/rat genes, as well as alternatively spliced exons, and control sequences. For more information on the HEEBO/MEEBO/MI-Rat oligonucleotide set please refer to http://alizadehlab.stanford.edu/.

Sample Processing

Tissue samples were homogenized in Trizol (Invitrogen, Carlsbad, CA) and total RNA was isolated. Biotin-labeled complementary RNA (cRNA) was prepared from total RNA by the method of Van Gelder et al., **(1)**. Briefly, an oligonucleotide containing a 5'-T7-promoter sequence and a 3' T24VN sequence was used to prime reverse transcription of RNA catalyzed by Superscript II (Invitrogen, Carlsbad, CA). Double-stranded cDNA was prepared from the 1st strand product by the method of Gubler & Hoffman **(2)**, and purified on a PCR purification column (Qiagen, Valencia, CA). The double-stranded cDNA was then used as a template for in vitro transcription with T7 RNA polymerase using a high yield transcription kit (Ambion) and including biotin-16-UTP (Ambion) in the reaction mixture. Biotinylated cRNA samples were fragmented, diluted in a formamide-containing hybridization buffer, and loaded on to the surface of HEEBO/MEEBO/MI-Rat microarray slides enclosed in custom hybridization chambers. The slides were hybridized for 16-18 hours under constant rotation in a Model 400 hybridization oven (Scigene, Sunnyvale, CA). After hybridization, the microarray slides were washed under stringent conditions, stained with Streptavidin-Alexa-647 (Invitrogen), and scanned using an Axon GenePix 4000B scanner (Molecular Devices, Sunnyvale, CA).

Data Pre-Processing

Spot intensities for each probe were calculated by subtracting median local background from median local foreground for each spot. Spot intensities were transformed by taking the base 2 logarithm of each value. The spot intensities were then normalized by subtracting the 70th percentile of the spot intensities of probes against mouse constitutive exons and adding back a scaling factor (grand mean of 70th percentile). After removing data for low quality spots, control sequences, and non-human/mouse/rat probes, XX,XXX human/mouse/rat probe intensities remained. The human/mouse/rat probes intensities were filtered to identify all probes with

Ocean Ridge Biosciences
10475 Riverside Dr. Suite 1
Palm Beach Gardens, FL 33410
Tel: 561-223-3152
Cell: 561-427-5548
Fax: 561-740-8710

intensity above a normalized threshold (log2 (3* standard deviation of raw local background) + mean of log2-transformed negative controls), to arrive at XX,XXX probes above threshold in all samples from at least one treatment group

Differential Expression Analysis

For statistical analysis, samples were binned in X treatment groups (X,X,X,etc…). The log2-transformed and normalized spot intensities for the XX,XX detectable probes were examined for differences between the treatment groups by 1-way ANOVA using National Institute of Ageing (NIA) Array Analysis software **(3)**. This ANOVA was conducted using the Bayesian Error Model and X degrees of freedom. The statistical significance was determined using the False Discovery Rate (FDR) method which was proposed by Benjamini and Hochberg **(4)**. It is the proportion of false positives among all probes with P values lower or equal to the P value of the probes that we consider significant. It can also be viewed as an equivalent of a P-value in experiments with multiple hypotheses testing. FDR is an intermediate method between the P-value and Bonferroni correction (multiplying P-value by the total number of probes). The equation is:

$$FDR_r = \min_{i \geq r} \left[\frac{p_i N}{i} \right]$$

where r is the rank of a probe ordered by increasing P values, p_i is the P value for probe with rank i, and N is the total number of probes tested. FDR value increases monotonously with increasing P value.

Gene Ontology Analysis

Gene ontology categories showing significant over-representation of differentially expressed genes were determined using GenMAPP software (Gladstone Institute, San Francisco, CA) for XX,XXX detectable probes with current Entrez Gene IDs. Specifically, the MAPPfinder module of GenMAPP was first used to map all detectable probes, based on their gene targets, to GO and Local MAPP categories. Then MAPPfinder compared the relative representation in each functional group of genes associated with probes meeting one of X differential expression criteria to the relative representation of genes associated with the full set of XX,XXX detectable probes. Significance was determined by permutation of Z scores with correction for multiple comparisons as described in the GenMAPP software manual. Genes meeting the different criteria were further analyzed using online Gene set Analysis Toolkit (Dr. Bing Zhang's group,

Ocean Ridge Biosciences
10475 Riverside Dr. Suite 1
Palm Beach Gardens, FL 33410
Tel: 561-223-3152
Cell: 561-427-5548
Fax: 561-740-8710

ORB Materials and Methods

Vanderbilt University) in order to identify the Wiki, KEGG, Pathway Commons and GO functional pathways showing an over-representation of differentially expressed genes.

Hierarchical Clustering Of Gene Expression Data

Data for the detectable probes were clustered using Cluster 3.0 software **(5)**. The data was pre-processed by three consecutive rounds of gene median centering and then hierarchically clustered using centered correlation as the similarity metric and average linkage as clustering method. Intensity scale shown is arbitrary.

References

(1) Van Gelder Multi-gene expression profile - US Patent 7049102

(2) Gubler, U and Hoffman, B. J. (1983) A simple and very efficient method for generating cDNA libraries. Gene 25 : 263-9.

(3) Sharov, A.A., Dudekula, D.B., Ko, M.S.H. (2005) Principal component and significance analysis of microarrays with NIA Array Analysis tool. Bioinformatics. 21(10): 2548-9.

(4) Benjamini, Y. & Hochberg, Y., (1995). J Roy Stat Soc B 57: 289-300

(5) De Hoon, M. J. L., Imoto, S., Nolan, J. Nolan, and Miyano, S. (2004) Open Source Clustering Software. Bioinformatics, 20 (9): 1453-1454.