Microarray Data Analysis

The statistical programming language R (http://cran.r-project.org/) was used. Raw expression measurements for all gene probes for all samples were first log (base=2) transformed then quantile normalized. Quality of data was assured via sample-level inspection by Tukey box plot, covariance-based PCA scatter plot, and correlation-based Heat Map. Raw expression measurements for samples deemed outliers were discarded and quantile normalization repeated. Gene probes not having at least one expression measurement greater then system noise post normalization were deemed "noise-biased" and discarded. System noise was defined as the lowest observed expression measurement at which the LOWESS (locally weighted scatterplot smoothing) fit of the CV (coefficient of variation) by mean for each gene probe for each class of samples (i.e., "ES undiff", "ES EB_ecto", "ES EB_mesend", "iPS undiff", "iPS EB_ecto", "iPS EB_mesend") deviates from linearity. For gene probes not discarded, expression measurements were floored to equal system noise if less than system noise then subject to the one-factor ANOVA (analysis of variance) under BH (Benjamini and Hochberg) FDR (false discovery rate) MCC (multiple comparison correction) condition. Gene probes with a corrected p-value < 0.05 were deemed "potentially informative" and subject to the TukeyHSD (honestly significant difference) post-hoc test. Gene probes having a post-hoc p-value < 0.05 and a fold-change magnitude difference of means >= 1.75 for a specific comparison of classes were deemed to have expression "significantly different" between the two classes. For these gene probes, measurements were subsequently interrogated for association with processing time and/or differences in gender using Spearman correlation and ANOVA respectively under BH FDR MCC condition (alpha < 0.05). Those gene probes having measurements significantly associated with processing time were deemed "processing-biased"; gene probes having measurements significantly associated with differences in gender were deemed "gender-biased". Annotations and associated functions for each gene probe were obtained using IPA (Ingenuity, Inc.).