

"Probe Level Analysis of Microarray Data Using the S-Score"

Richard Kennedy

PhD Candidate

Virginia Commonwealth University

Abstract

The analysis of Affymetrix GeneChip data is a complex, multistep process. Most often, methods condense the multiple probe level intensities into single probe set level measures (such as RMA, dChip and MAS5), which are then followed by application of statistical tests to determine which genes are differentially expressed. An alternative approach is a probe-level analysis, which tests for differential expression directly using the probe-level data. Probe-level models offer the potential advantage of more accurately capturing sources of variation in microarray experiments. However, this has not been thoroughly investigated, since current research efforts have largely focused on the development of improved expression summary methods. We present recent work by our group, which examines a novel probe-level analysis algorithm for performing two-chip comparisons, the S-Score. As described by Zhang and colleagues, the S-Score is based on an error model in which the detected signal is proportional to the probe pair signal for highly expressed genes, but approaches a background level (rather than 0) for genes with low levels of expression. This model is used to calculate relative change in probe pair intensities that converts probe signals into multiple measurements with equalized errors, which are summed over a probe set to form the S-Score. Under the null hypothesis of no differential expression, the S-Score is approximately normally distributed, allowing direct *P* value calculation. This presentation will review our initial studies of the S-Score, which have shown that its performance compares favorably to other methods that use probeset expression summaries, as an example of probe-level analysis. Limitations remaining in the S-Score algorithm will be discussed, as well as work in progress to address these limitations and improve the flexibility of the S-Score.