

Mixed effects model for assessing RNA degradation in Affymetrix GeneChip experiments

Kellie J. Archer¹, Suresh E. Joel², Viswanathan Ramakrishnan¹

¹Department of Biostatistics, Virginia Commonwealth University,
Richmond, VA 23298-0032

²Department of Biomedical Engineering, Virginia Commonwealth University,
Richmond, VA 23298-0694

Abstract

Due to the high cost of microarray experiments, investigators typically select designs with biological rather than technical replicates. Therefore, it is essential that the quality of the RNA hybridized to a microarray meets certain standards. The process of transcription begins with reverse transcriptase binding at the 3' end of a gene and continuing toward the 5' end. However, transcription generally does not continue to completion. That is, reverse transcription typically drops off before reaching the 5' end. Affymetrix GeneChips includes probe sets which interrogate both the 3' and 5' ends of selected control genes to assess the quality of transcription. The Microarray Suite software estimates the 3':5' ratio after the perfect match (PM) and mismatch (MM) level data have been summarized into probe set expression measures. Unfortunately, an inherent problem with the application of any probe set expression summary method is that the 3' and 5' probe sets of interest are only represented on the GeneChip once. This leads to the unfortunate consequence of no replicates for variance estimation. This research proposes the use of pixel level intensities to increase the number of observations per probe set in order to obtain an estimate of the 3' to 5' ratio and an associated confidence interval. Since there is an inherent hierarchical structure to GeneChip data, where pixels are nested within probes and probes are nested within probe sets, the proposed method is to assess RNA degradation by fitting mixed effects ANOVA models to estimate the 3' to 5' ratio treating probe set as a fixed effect, treating PM level data as random effects, and treating pixel level data as a subsample nested within PM probe. The proposed method thus enables the construction of confidence intervals about the estimated ratio. This estimated confidence interval will more appropriately indicate whether the RNA was of sufficient quality rather than judging quality based on the ratio being below an arbitrarily selected threshold. Results from both HG-U133A and HG-Focus GeneChips will be presented.