Homework Assignment 2

RMA Normalization

Normalization refers to the process of modifying microarray data for effects caused by technological variations rather than biological differences amongst RNA samples or printed probes. Robust Multiarray Averaging (RMA) is a microarray normalization process that background corrects, normalizes, and reports probe level data without relying on information gained from MM probes. RMA is an expression matrix creation algorithm based on Affymetrix data. Background correction, log2 transformation, and quantile normalization are all applied to the raw intensity values. Multiplying the previous Rolling Average by n-1 periods, adding today's symbol price, and then dividing the result by n periods yields the most recent Rolling Average. It's worth noting that the RMA's initial value is based on a Simple Moving Average. Using the RMA expression measure, this function converts an AffyBatch object to an ExpressionSet object.

Mas5.0 Normalization

Providing an algorithm that integrates the signals from the many Perfect-Match (PM) and Mismatch (MM) probes that target each transcript into a single value that sensitively and precisely depicts its concentration is a key difficulty with Affymetrix expression data. MAS5.0 does this by calculating a robust average of the (logged) PM-MM values; increased variation is observed at low signal strengths and is at least in part due to the extra noise generated by subtracting the MM values from their PM partners. The first value, an estimate of transcript concentration, and the second, a measure of how much the software believes the first, are returned by MAS5.0. The fact that this result is referred to as the "detection p-value" and is then utilized to generate a recognition call that marks the transcript as Present, Marginal, or Absent (P/M/A) may cause some confusion. In fact, the detection p-value is more akin to a reliability score, and labels like Reliable, Marginal, and Unreliable may be more applicable.

References:

Molecular Organisation and Assembly in Cells.” *RMA and GC-RMA Normalisation*, https://warwick.ac.uk/fac/sci/moac/people/students/2003/sam\_robson/usergroups/rmavsmas5/.

Pepper, S.D., Saunders, E.K., Edwards, L.E. *et al.* The utility of MAS5 expression summary and detection call algorithms. *BMC Bioinformatics* **8,**273 (2007). <https://doi.org/10.1186/1471-2105-8-273>