

*Spine Equation Modeling and Univariate Scaling (SEMUS) for Microarray Data Normalization*

Ni, T.

Vanderbilt University, Nashville, Tennessee

Microarray technology is used to quantify the transcripts that are products of independent or related genetic elements. Large scale data generated is typically normalized based on the assumption that the underlying data sets to be compared are identical and independently distributed (iid), either globally or locally. However, the dependency between the transcripts, the amount of genes that are differentially expressed, and the presence of nonlinear measurement error could skew the distribution or invalidate the iid assumption. Thus it is important for any normalization method to be robust in the presence of variety sources of variations. The new normalization algorithm, SEMUS, a multi-step nonparametric approach for normalizing microarray data, is developed to be also applicable and robust when the two distributions are significantly different. The performance of the method against high variable data is presented.