

“Variance Stabilization Approach for Mass Spectrometry Proteomic Data”

Xiaoyong Wu, PhD
McGill University

Abstract

Mass spectrometry allows us to analyze the proteins in cells, tissues or species. The large amount of data generated in mass spectrometry-based proteomics experiments involves many steps from sample preparation to data analysis and are subject to considerable sources of noise, which result in a complicated statistical error structure with across-replicate variances changing as a function of replicate means. This dependence of variance on signal intensity violates a key assumption of standard statistical methods such as the t-test. As a consequence, analysis of proteomic data presents a significant challenge to computational and statistical methods. We develop a variance stabilization approach for mass spectrometry data which removes the variance/signal dependence and in turn facilitates valid statistical analysis. Our approach has been effective for several data sets generated by our lab and one data set published in the journal *Cancer Cell*. Our variance stabilizing transformation, which we have labeled “balance transformation”, could in principle also be applied to other data type with replication, e.g., microarray data.