

A Framework for Integrating DNA sequence, mRNA and Protein Expression Profiles

Kelvin H. Lee
Cornell University

Biological systems manipulate gene expression in response to different biochemical, environmental and genetic factors. Technology now permits the genome-wide measurement of DNA sequence, mRNA expression levels and protein expression levels. However, the relationship between measured changes in gene expression at the mRNA level to the corresponding changes at the protein level is not well understood. We have developed a genome-wide mathematical framework to integrate DNA sequence information together with mRNA expression and protein expression profiles from *Escherichia coli*. Such a model could be used to obviate the need to measure protein expression profiles in certain systems when chip data is available. More importantly, a model can be used in combination with mRNA and protein data to better understand gene expression regulation. We use Affymetrix Genechip Probe Arrays to monitor mRNA changes and we use two-dimensional protein electrophoresis – mass spectrometry as well as iTRAQ LC MALDI approaches to quantify protein expression and validate the model. The iTRAQ technology permits the simultaneous quantification of protein expression from up to four different samples in parallel using isobaric tags for relative and absolute quantitation. We employ this technology on a TOF/TOF mass spectrometer.