A statistical approach to determining equivalent functional genomic differences across experiments

Jeffrey Thompson
University of Kansas Medical Center, Kansas City, Kansas

In silico functional genomics have become a driving force in the way we interpret and use gene expression data, enabling researchers to understand which biological pathways or molecular functions are likely to be affected by the treatments or conditions being studied. There are many approaches, but a number of popular methods determine if a set of modified genes has a higher than expected overlap with genes known to function as part of a pathway (functional enrichment testing). Recently, researchers have started to apply such analyses in a new way: to ask if the data they are collecting show similar disruptions to biological functions as some reference data. Examples include studying whether or not similar genes are perturbed in smokers vs. users of e-cigarettes, or whether a new mouse model of schizophrenia is justified, based on its similarity in cytokine expression to a previously published model. However, there is a dearth of robust statistical methods for testing hypotheses related to these questions. This work proposes a novel statistical approach to testing if the observed perturbances in two biological datasets cause equivalent biological functional changes.