

**IDENTIFICATION OF DIFFERENTIALLY EXPRESSED GENE CATEGORIES  
IN MICROARRAY STUDIES USING MULTIVARIATE NONPARAMETRIC  
ANALYSIS**

by

**Dan Nettleton  
Department of Statistics  
Iowa State University**

**Justin Recknor  
Eli Lilly**

**James M. Reecy  
Department of Animal Science  
Iowa State University**

**December 22, 2006**

**ABSTRACT**

We present a new method for identifying gene categories whose joint expression distribution differs across two or more experimental conditions. A multivariate nonparametric test statistic based on within-condition sums of distances between expression vectors is used to identify categories of interest. A permutation testing approach is used to obtain approximate control of the false discovery rate when testing multiple categories. The method provides an alternative to using Fisher's exact test on gene lists, gene set enrichment analysis (GSEA), significance analysis of function and expression (SAFE), and other approaches for finding gene categories of interest. We discuss some advantages of our approach relative to other methods and illustrate its use on an example experiment aimed at identifying gene categories whose joint expression distribution differs between wild type and myostatin-null mice.