

3rd International Conference & Exhibition on **Biometrics & Biostatistics**

October 20-21, 2014 DoubleTree by Hilton Baltimore - BWI Airport, USA

Assessing inter-individual variability in gene expression using RNA-seq data from 750 individual *Drosophila*

Lin Y, Golovkina K, Chen Z-X, Smith HE, Sultana H, Serrano Negron YL, Oliver B C and Harbison ST
National Institute of Health, USA

Clear differences in phenotypes such as sleep duration have been observed among individual flies with identical genotypes. One possibility is that gene expression plays a role in producing these differences. However, previous experiments that assessed genetic differences in gene expression were done using pools of flies rather than individuals. The author wanted to determine whether heritable differences in gene expression could be detected among individual flies therefore a multi-factor experiment using RNA extracted from 768 flies was performed. The author harvested RNA from individual flies using 16 inbred lines from the *Drosophila* Genetic Reference Panel. These flies were reared in three biological replicates of the same environmental condition. The RNA was successfully sequenced for 98% of the flies and the output read count data was used in the data analysis. In the data preprocessing, 'bar code' to verify that each sample contained the correct genotype was constructed and Spearman correlations were used to verify that the sex of each sample was correct. Generalized linear models using a negative binomial distribution to account for the characteristics of the read count data were applied. The three main effects of genotype, environment, and sex as well as their two- and three-way interaction terms were included in the models. In addition, a pipeline for RNA-Sequence data analysis consisting of the filtering, normalization, and model fitting steps was applied. The performance of different available methods and tools used in each step of the pipeline was compared. Results showed that the trimmed mean of M-values (TMM) normalization is preferred in our setup and DESeq R package works the best in the model fitting. Preliminary results indicate that differences in gene expression exist in for flies with identical genotypes subjected to identical environments.

Biography

Lin Y got her PhD degree at the Department of Statistics at Purdue University in December 2011. During her graduate study, she focused on the dimension reduction and variable selection problems in high dimensional data. She currently is a Research Fellow in the Laboratory of Systems Genetics at the National Institutes of Health (NIH). She now works on genome wide association study of single fly data and RNA sequencing data analysis of single fly project.

yanzhu.lin@nih.gov