# Visualizing Effect of Data Transformations on Errors

Robert M Flight and Hunter NB Moseley

Department of Molecular and Cellular Biochemistry

Markey Cancer Center

Resource Center for Stable Isotope-Resolved Metabolomics

University of Kentucky

In many omics analyses, a primary step is data transformation. A transformation is generally employed to convert proportional error (variance) to additive error, which most statistical methods appropriately handle. However, omics data frequently contain error sources that result in both additive and proportional errors. To our knowledge, there has not been a systematic study on detecting proportional error in omics data, or the effect of transformations on the error structure. In this work we demonstrate that a set of three simple graphs can facilitate the detection of proportional and mixed error in omics data. The graphs plot the 1) absolute range, 2) standard deviation and 3) relative standard deviation across replicates against the mean signal across replicates. In addition to showing the presence of different types of error, they also readily demonstrate the effect of various transformations on the error structure as well. Using these graphical summaries we find that the log-transform is the most effective method of the common methods employed for removing proportional error.