# Heatmap -- Dendrogram Summary:

Heatmaps are an effective tool for displaying feature variation among groups of samples. The basic concept of a heatmap is to represent relationships among variables as a color image. Rows and columns typically are reordered according to the dendrograms so that variables and/or samples with similar profiles are closer to one another, making these profiles more visible. Each value in the data matrix is displayed as a color, making it possible to view the patterns graphically.

Heatmaps uses an agglomerative hierarchical clustering algorithm to order and display the data as a dendrogram. Two important factors to consider when constructing a heatmap are the type of distance measure and the agglomeration method used. For details on the various methods available see [https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-13-S16-S10].

### Input Summary:

Input Dataset: -> e.csv

Output Datasets and Files: -> Heatmap -> heatmap\_plot.svg -- compound\_order.csv -- sample\_order.csv

**- Scaling Method:** Auto Scaling. Scaling methods are data pretreatment approaches that divide each variable by a factor, the scaling factor, which is different for each variable. They aim to adjust for the differences in fold differences between the different metabolites by converting the data into differences in concentration relative to the scaling factor. It is highly recommended for multivariate statistical analyses. More information is available at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1534033/.

**- Agglomeration Method:** average. Agglomeration is the process by which clusters are merged into larger clusters.

**- Distance Function:** NA. A distance metric is a non-negative number which measures the difference between two objects (e.g. samples/compounds.)

### Result Summary:

The Heatmap and Dendrograms on the dataset with Auto Scaling. See Figure1 for more details.

Figure 1

shows the heatmap and dendrogram analysis on the dataset with Auto Scaling. The row displays compounds and the column represents the samples. The color represent the value scale for the corresponding compound and sample (see colorbar).

The dendrogram on the right is conducted using distance metric NA with agglomeration method of average on the compounds while the top is on the samples. The annotation of the samples (Genotype, Gender) were added on top of the plot. No annotation of the compounds were added to the plot.

Figure 1

shows the heatmap and dendrogram analysis on the dataset with Auto Scaling. The row displays compounds and the column represents the samples. The color represent the value scale for the corresponding compound and sample (see colorbar).

The dendrogram on the right is conducted using distance metric NA with agglomeration method of average on the compounds while the top is on the samples.

The order of samples is determined by the Genotype, and the compounds by the dendrogram. The annotation of the samples (Genotype, Gender) were added on top of the plot. No annotation of the compounds were added to the plot.

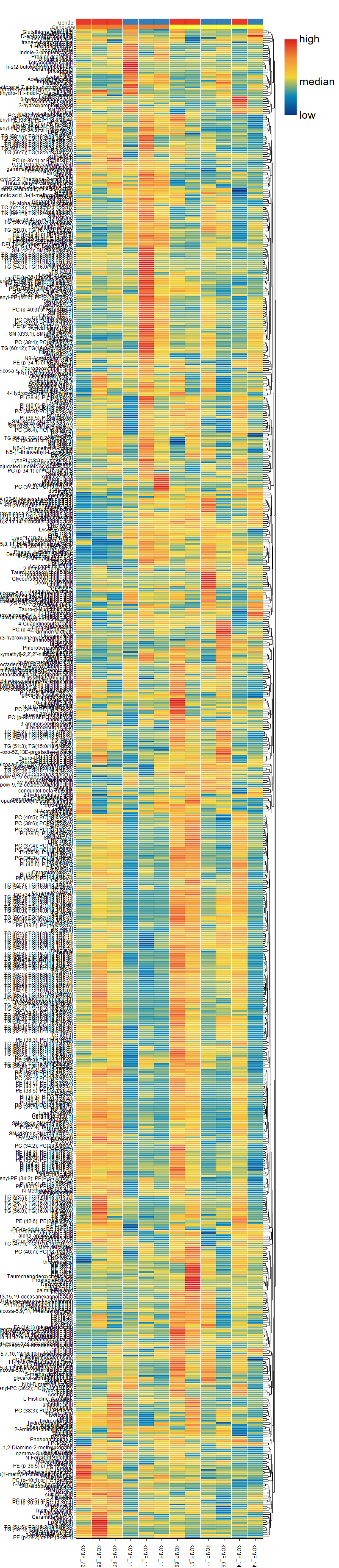


Figure 1: the heatmap of the dataset with Auto Scaling.