

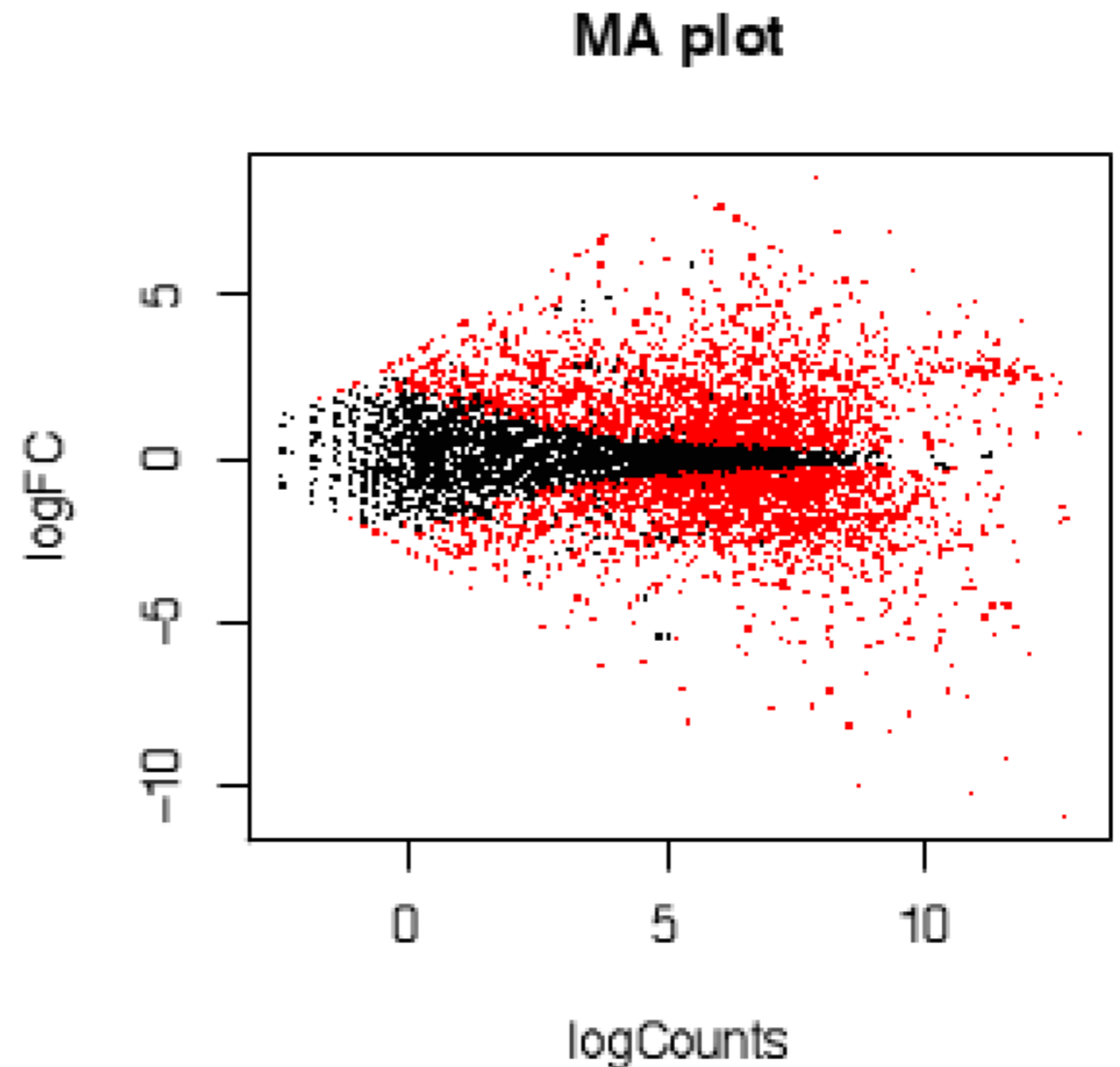
Visualization of RNA-Seq Data

Data Visualizations

- Visualizations are useful for:
 - Identifying patterns/issues in the data
 - Summarizing results
 - Illustrating relationships between variables

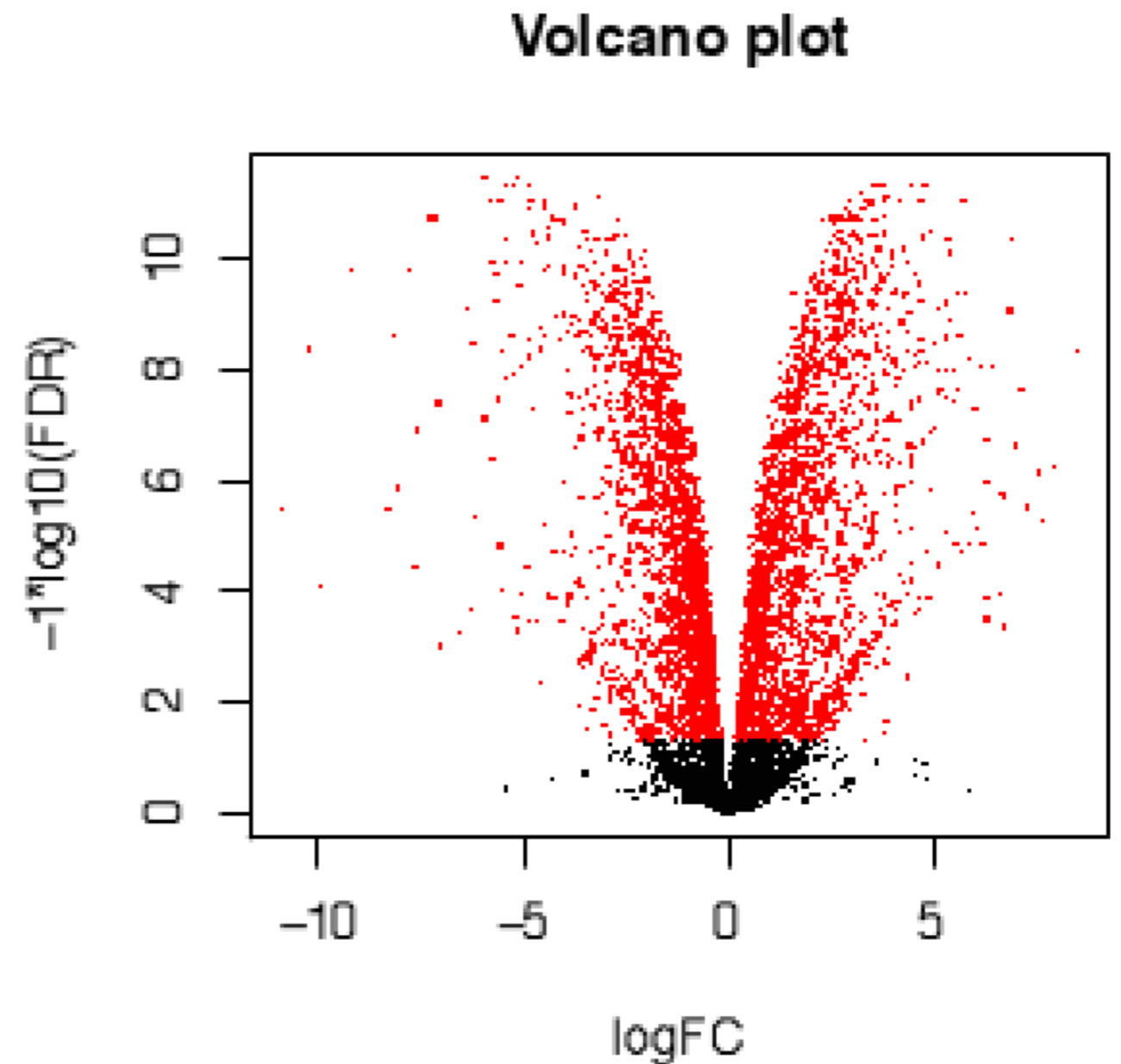
MA Plots

- For visualizing differences in measurements (in this case, gene expression) between two groups.
- M- log fold change (differences between two groups)
- A - mean gene expression (average value across samples)



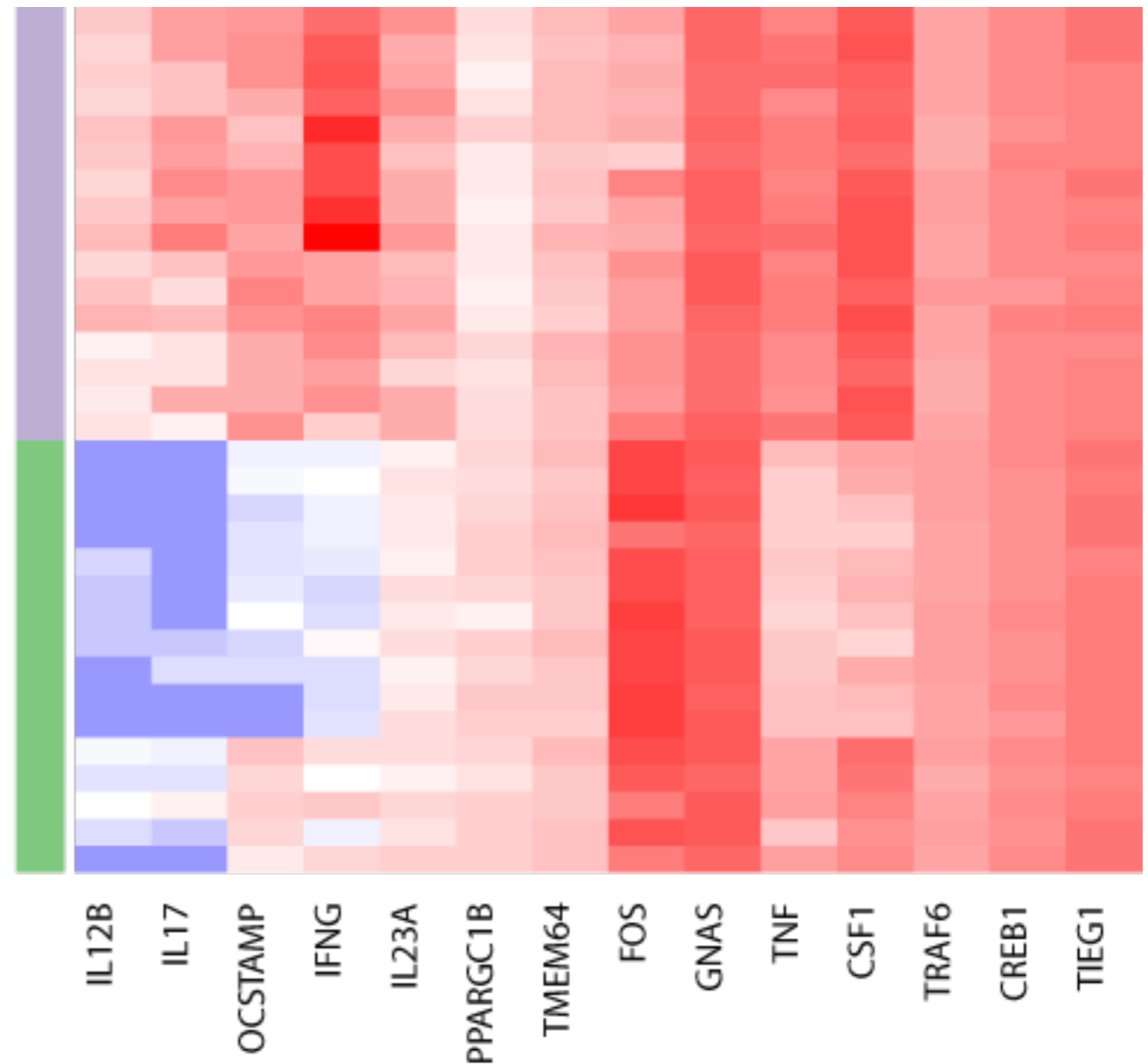
Volcano Plots

- Plots fold change vs significance value for all genes.
- Helps quickly see how many significantly differentially expressed genes are present.



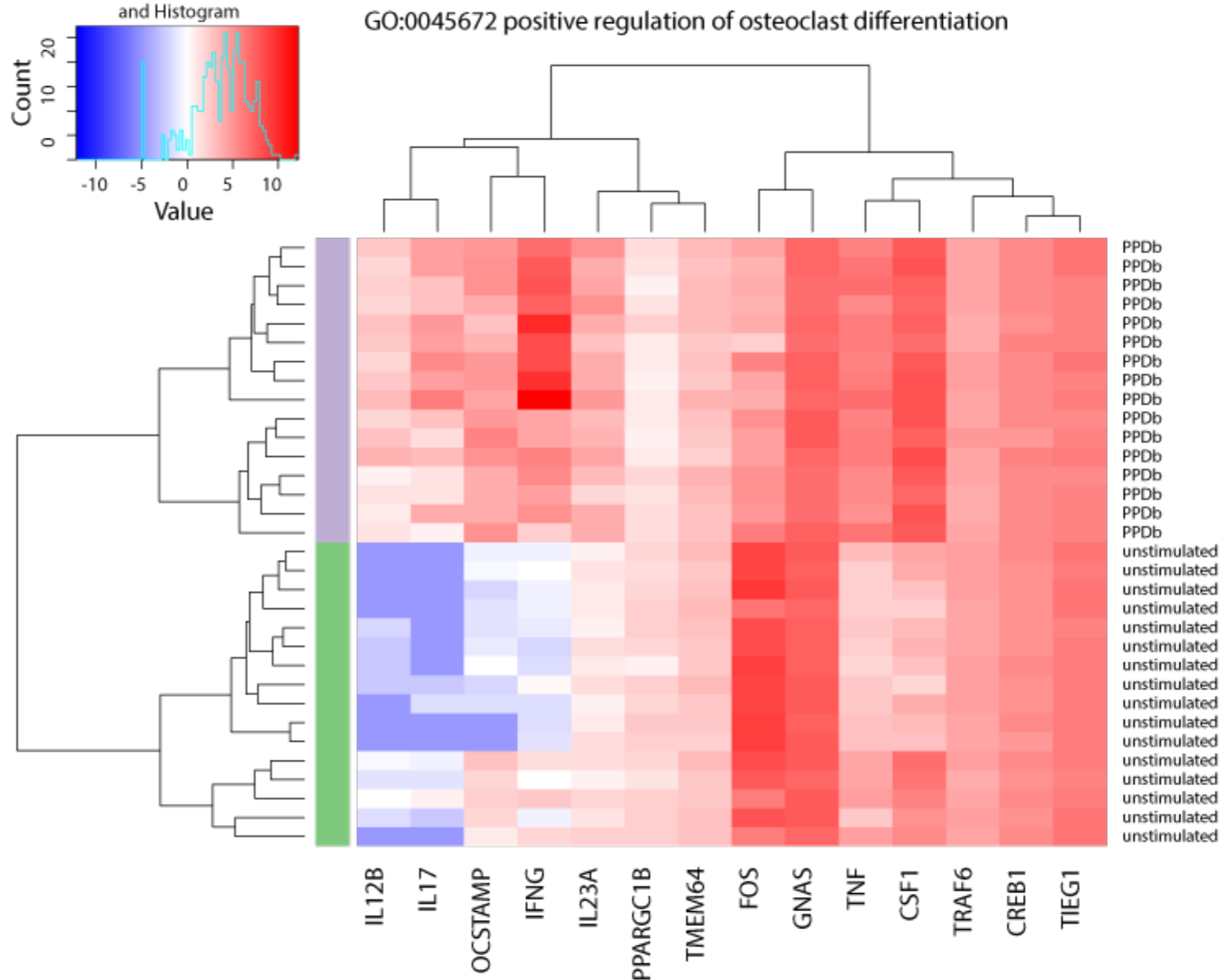
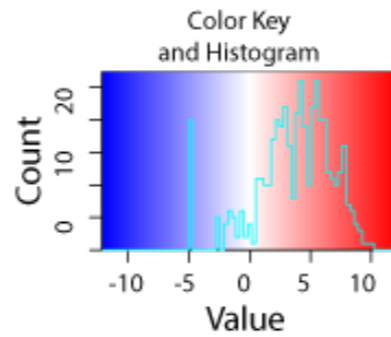
Heatmaps

- Heat Maps represent gene expression by colors.
- For visualizing how gene expression changes in different samples.
- Rows are genes
- Columns are Samples



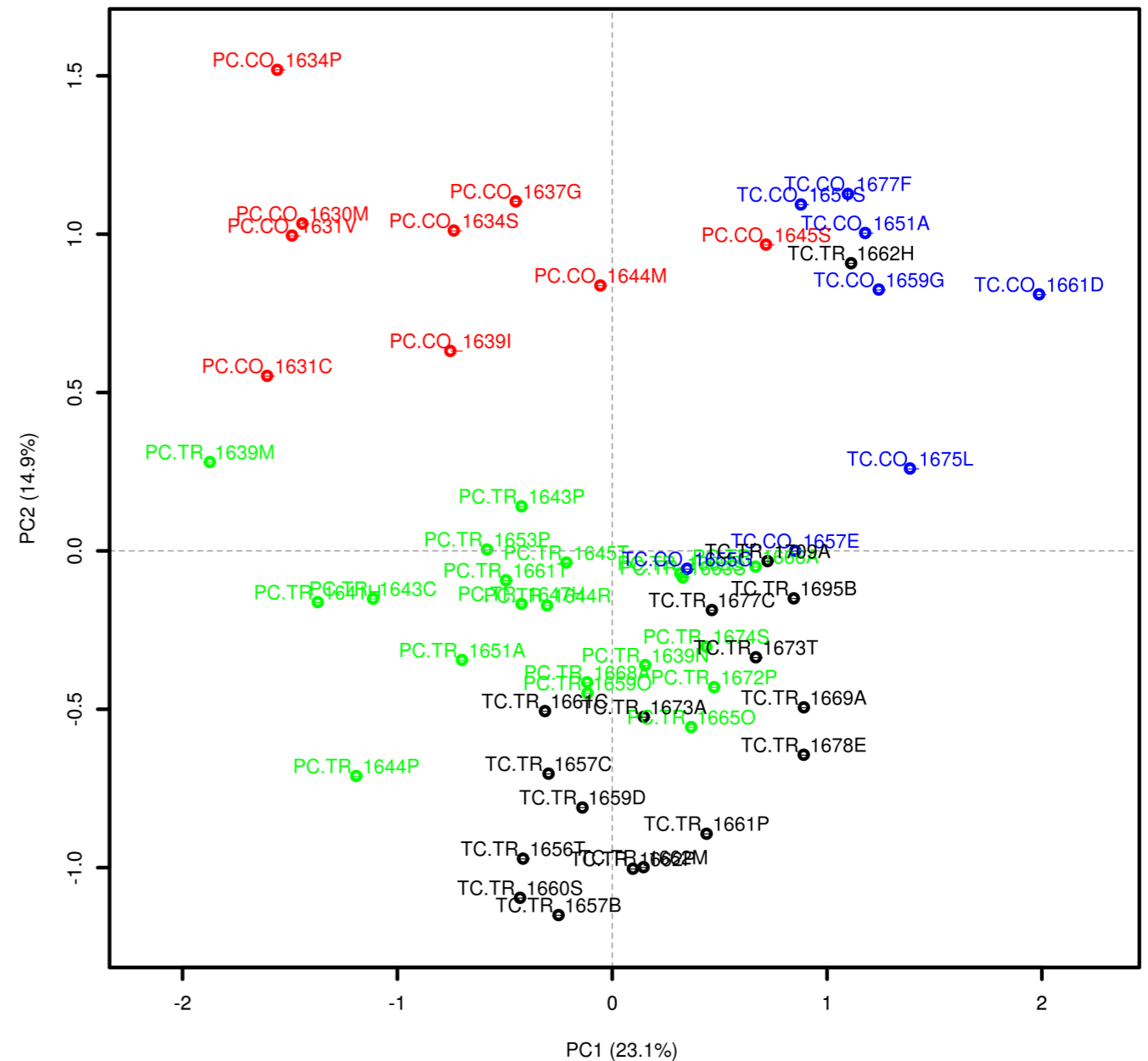
Heatmaps/Clustering

- Dendrograms can be added to heat maps
- Samples can be clustered by gene expression
- Genes can be clustered by gene expression
- time consuming for large number of genes



Principal Component Analysis

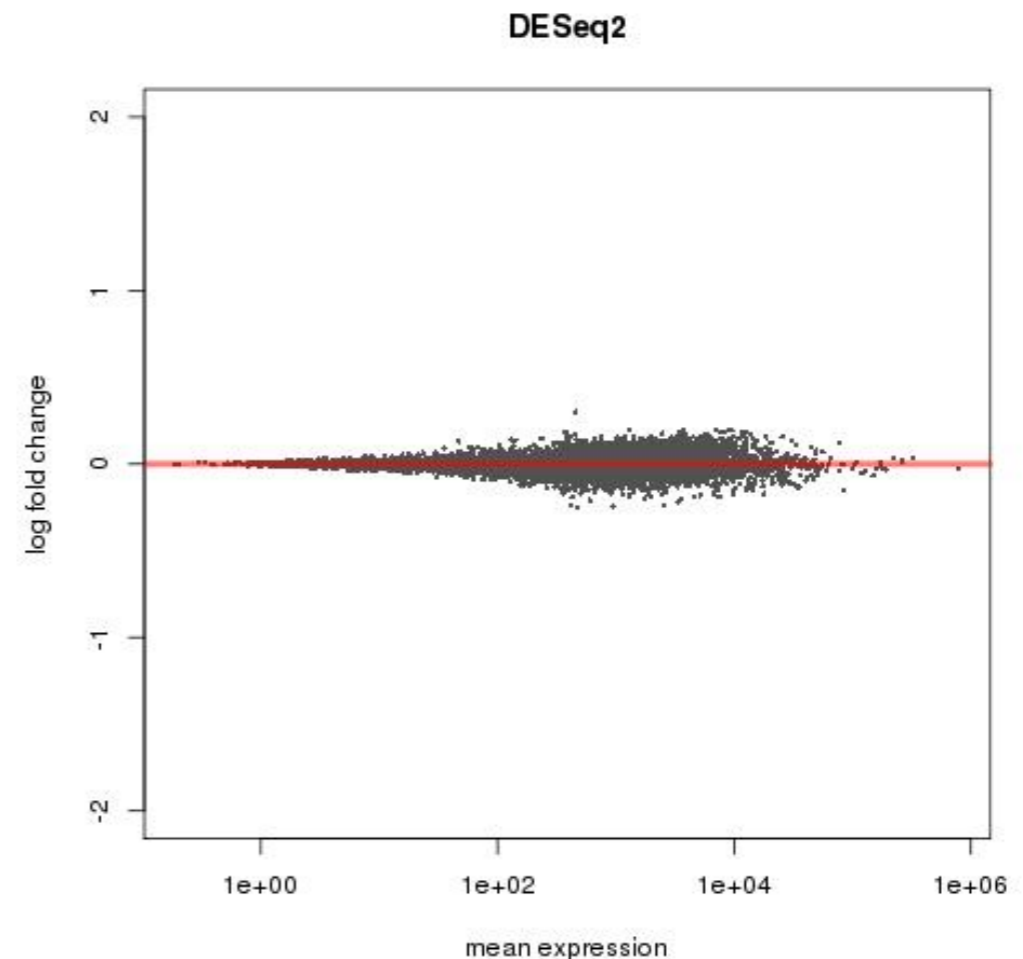
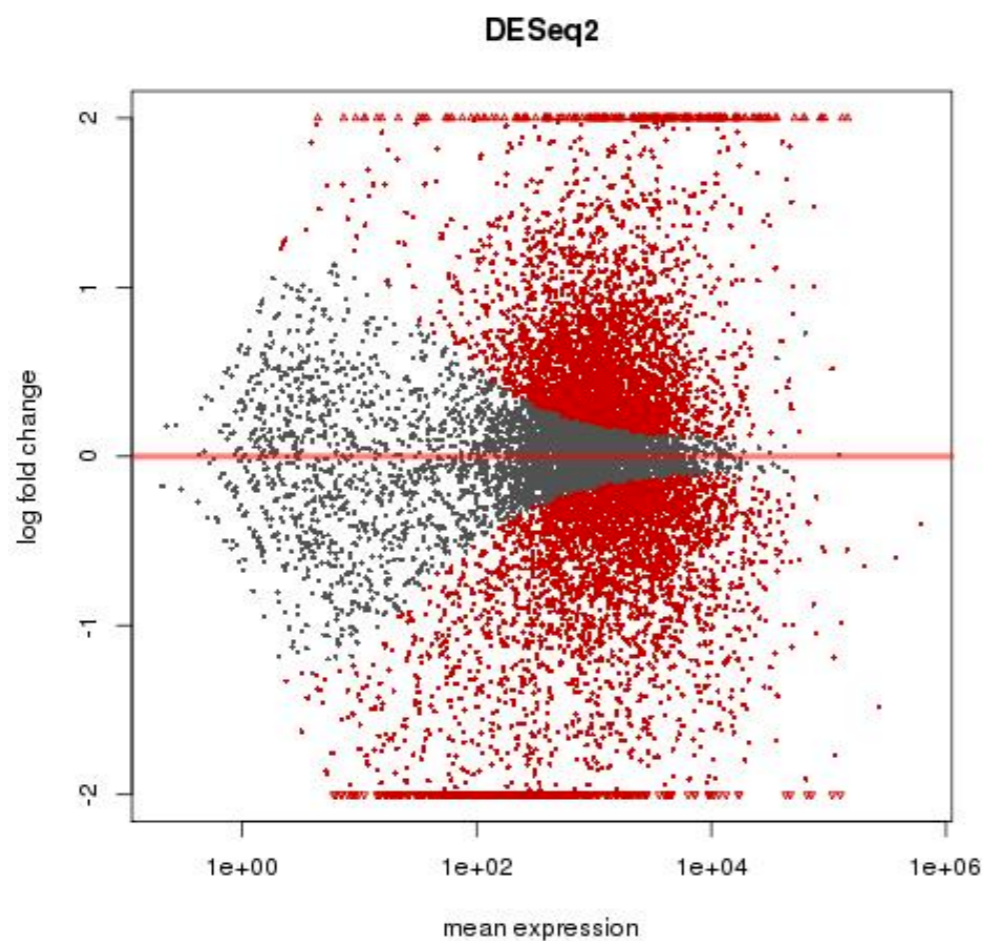
- Each principal component is one dimension in the data.
- Illustrates how the data groups based on the dimensions that represent the highest variability.



Looking at Some Real Data

- **Mysterious results for an experiment with 6 samples across:**

- 2 different time points, 2 different conditions: control vs treated. 3 replicates each.



Looking at Some Real Data

- Can these plots inform us about what might be going on?

