Explore and visualize your data using Graphpad Prism

Zhengtao Xiao
8.26.2020

zhengtao.xiao@duke.edu

"The simple graph has brought more information to the data analyst's mind than any other device." – John Tukey
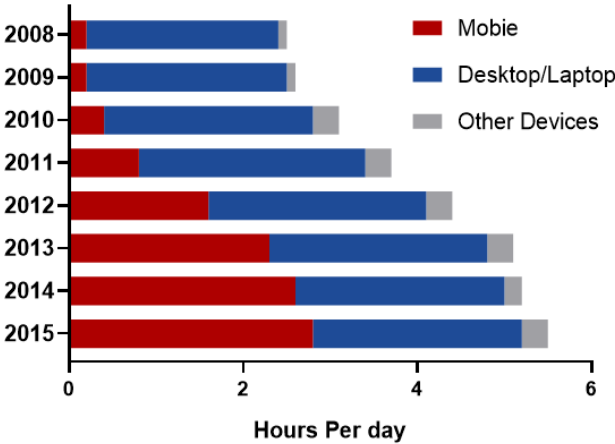
Graph help us interpret scientific data more efficiently
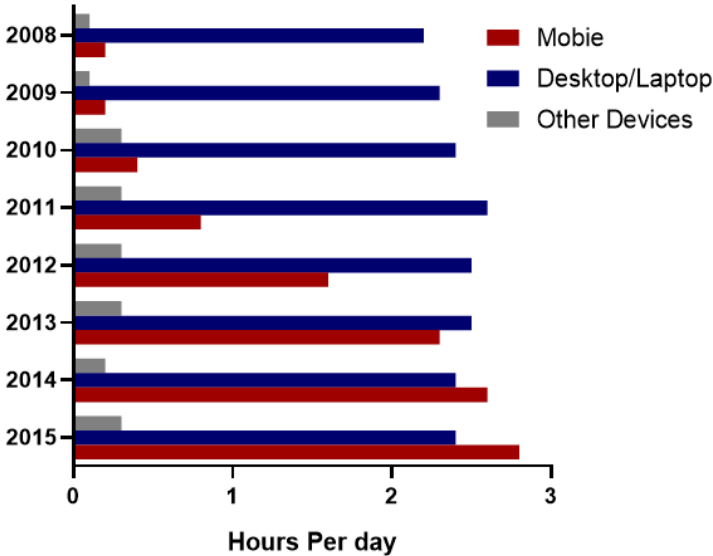
Time spent with digital media

Year	Mobie	Desktop/Laptop	Other Devices
2015	2.8	2.4	0.3
2014	2.6	2.4	0.2
2013	2.3	2.5	0.3
2012	1.6	2.5	0.3
2011	0.8	2.6	0.3
2010	0.4	2.4	0.3
2009	0.2	2.3	0.1
2008	0.2	2.2	0.1

When to use what

Time spent with digital media

Time spent with digital media

Choose the most appropriate graph for your data

https://www.data-to-viz.com/

Why Graphpad Prism?

- Easy and fast
- Commonly used statistical methods
- Graphs and data are automatically updated in real time
- Reusable graph template
- … …

Three steps to generate a graph

❑ Prepare data

❑ Perform analysis

❑ Choose the type of graph

Dataset for practice

Proteomic and Metabolomic Characterization of COVID-19 Patient Sera

Graphical Abstract
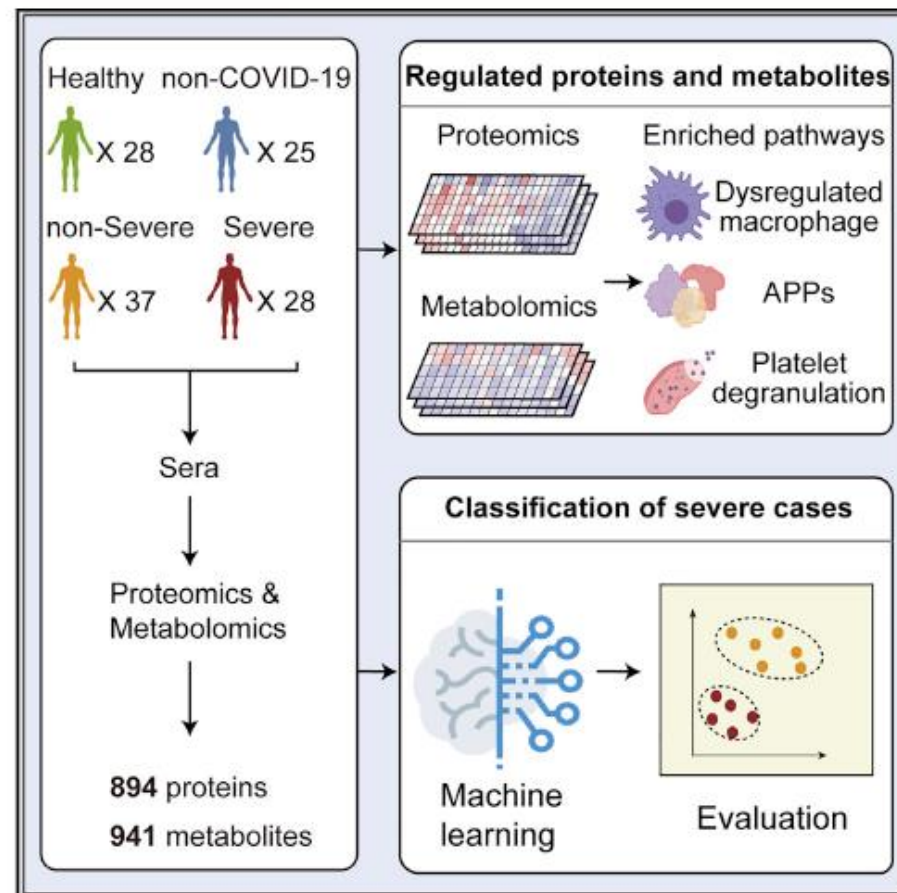
Authors

Bo Shen, Xiao Yi, Yaoting Sun, ...,
Huafen Liu, Haixiao Chen, Tiannan Guo

Correspondence

zhuyi@westlake.edu.cn (Y.Z.),
liuhf1@dazd.cn (H.L.),
chenhx@enzemed.com (H.C.),
guotiannan@westlake.edu.cn (T.G.)

In Brief

Proteomic and metabolomic analysis of COVID-19 sera identifies differentially expressed factors that correlate with disease severity and highlights dysregulation of multiple immune and metabolic components in clinically severe patients.

- ❑ Distribution of metabolites' concentration

- ❑ Correlation of metabolomics across patients

- ❑ T-test (comparison of two groups)

- ❑ One-way anova (comparing more groups)

Practice: Transform the data to log2 scale

1. "log2" changes the distribution to normal-like distribution
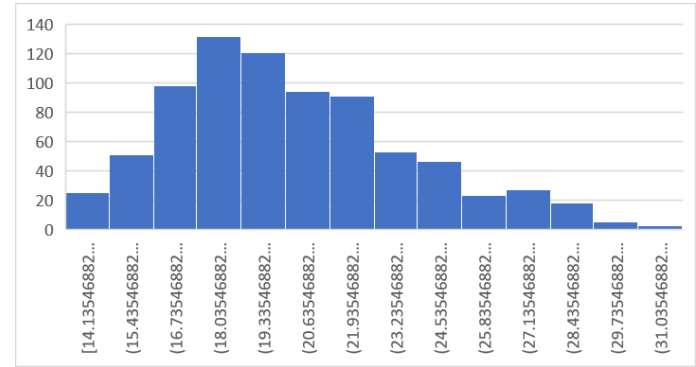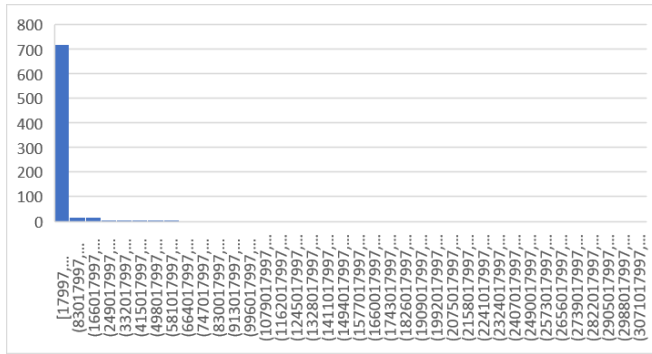
2. "log2" brings the up- and down-regulated genes/metabolites to same scale

Control = 20	Control = 160
Treatment = 160	Treatment = 20

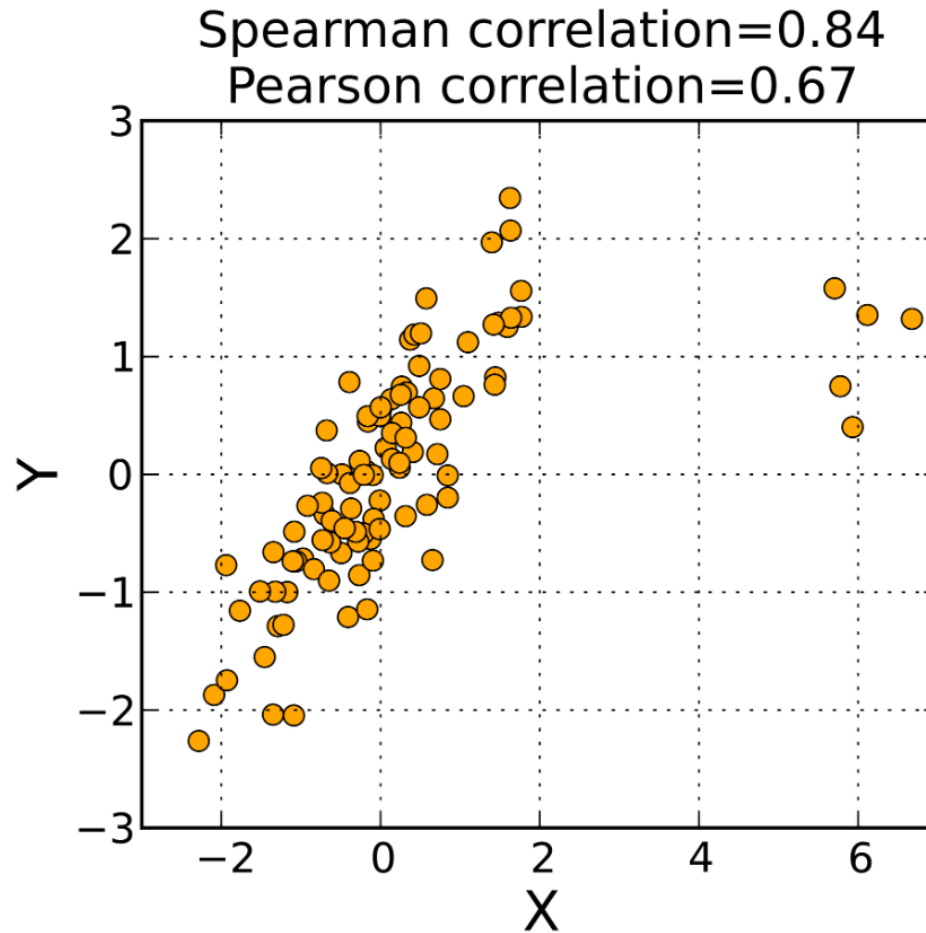Fold change = 8   Fold change = 0.125

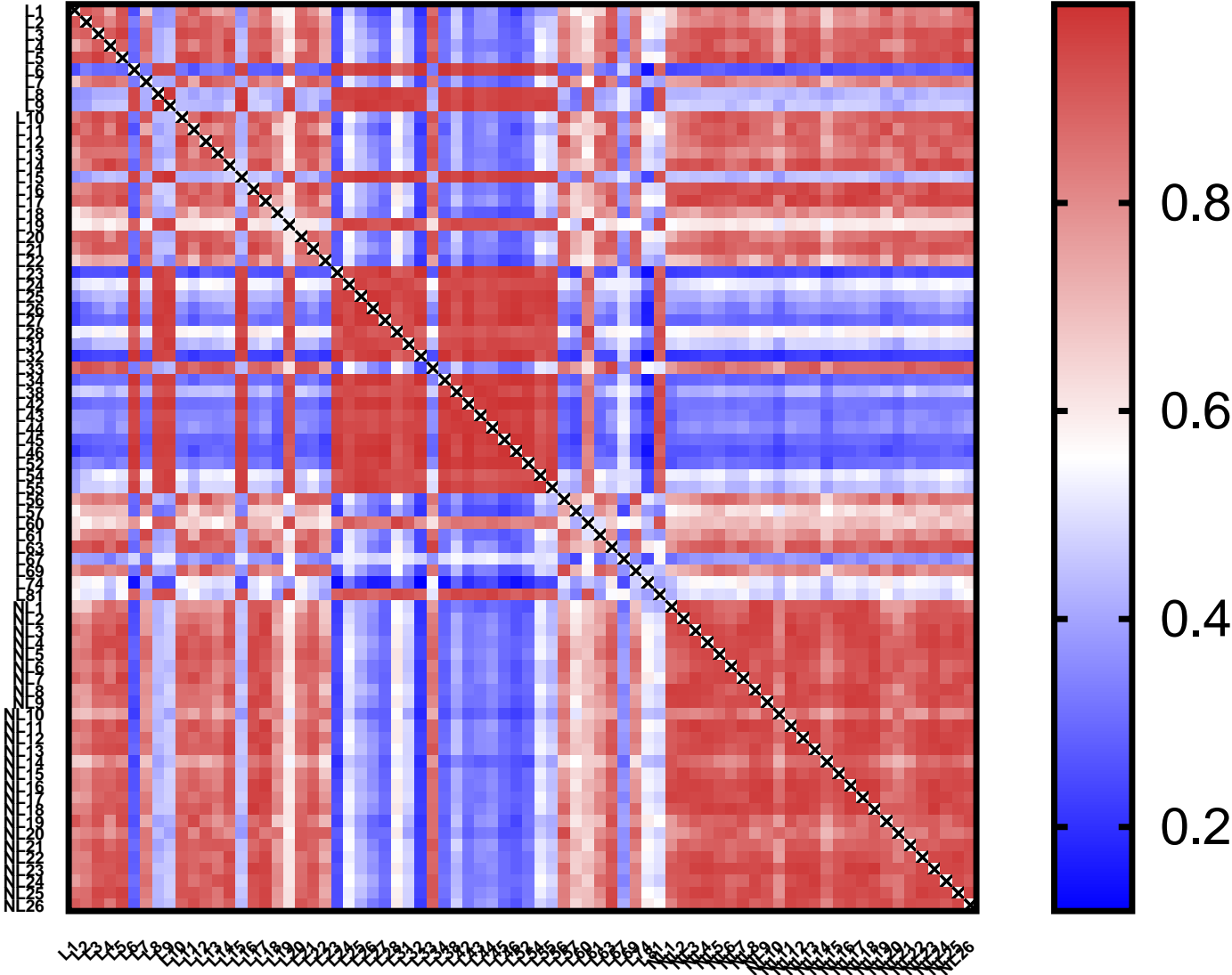Log2 fold change = 3   Log2 fold change = -3

Practice: calculate correlation
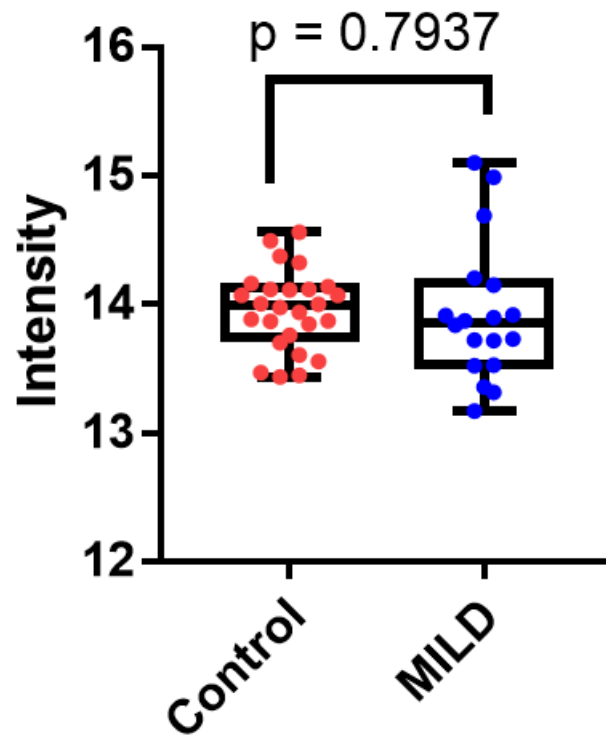
Correlation (pearson or spearman?)

The Spearman correlation is less sensitive than the Pearson correlation to strong outliers

Practice: correlation matrix

Practice: using t-test to compare two groups

Advice: when to plot SEM vs. SD

- **SD**: standard deviation, quantifies how much the values vary from one another

- **SEM**: standard error of the mean, represents the accuracy of the true mean for the population. SEM = SD/(square root of sample size)

- **SD**: If you want to show the variation of your data. For example, if you want to present that one metabolite is much more stable than another one by stimulus.

M1   M2

- **SEM**: If you want to show how precisely you determine the true mean. For example, you want to present that one metabolite is up regulated after infection.
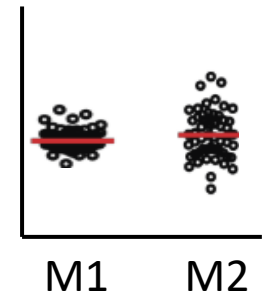
Practice: using anova to compare multiple groups

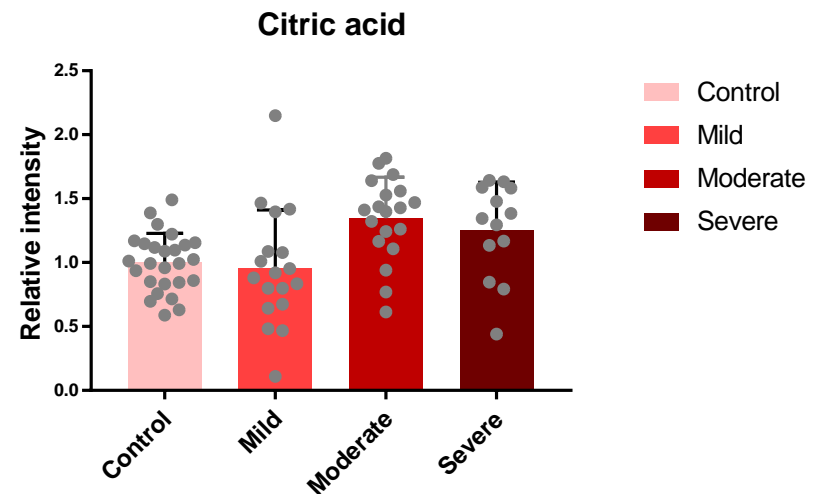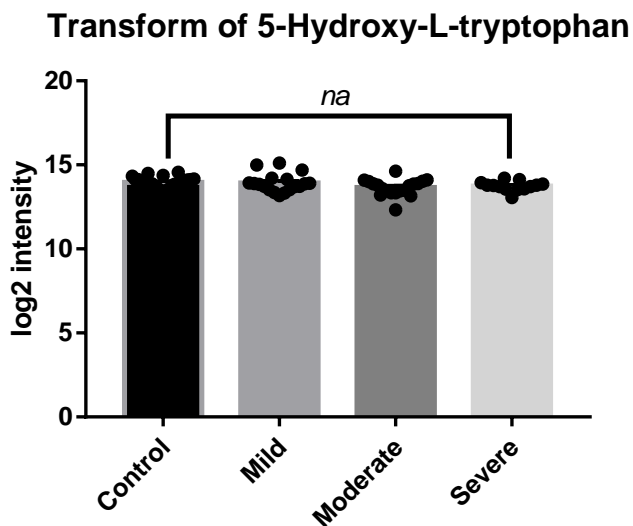**We want to see if selected metabolite significantly changed after infected.**

One-way anova (you can simply think it as a multiple t-test):

determine whether there are any statistically significant differences between the

means of two or more independent (unrelated)

Null and Alternative hypotheses

H0: $\mu(control) == \mu (mild) == \mu(moderate) == \mu(severe)$

H1: not all μ are equal

Transform of 5-Hydroxy-L-tryptophan

Citric acid
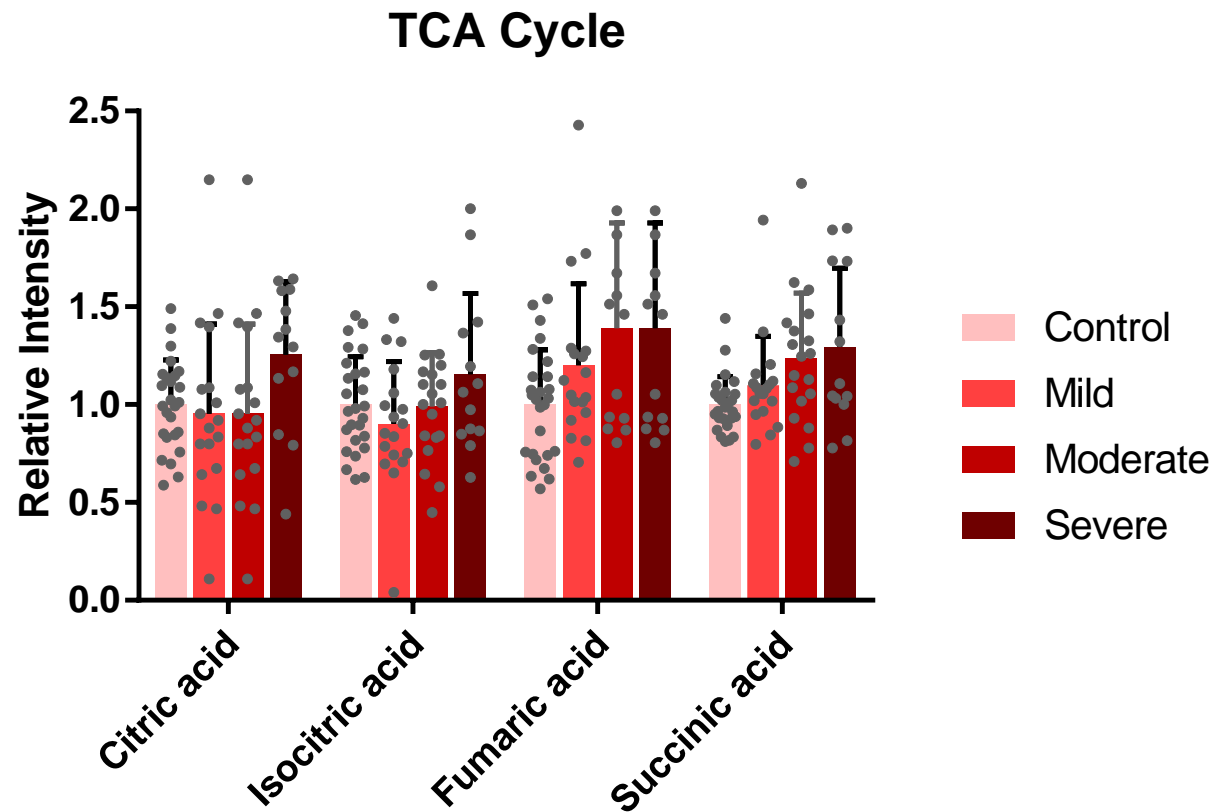
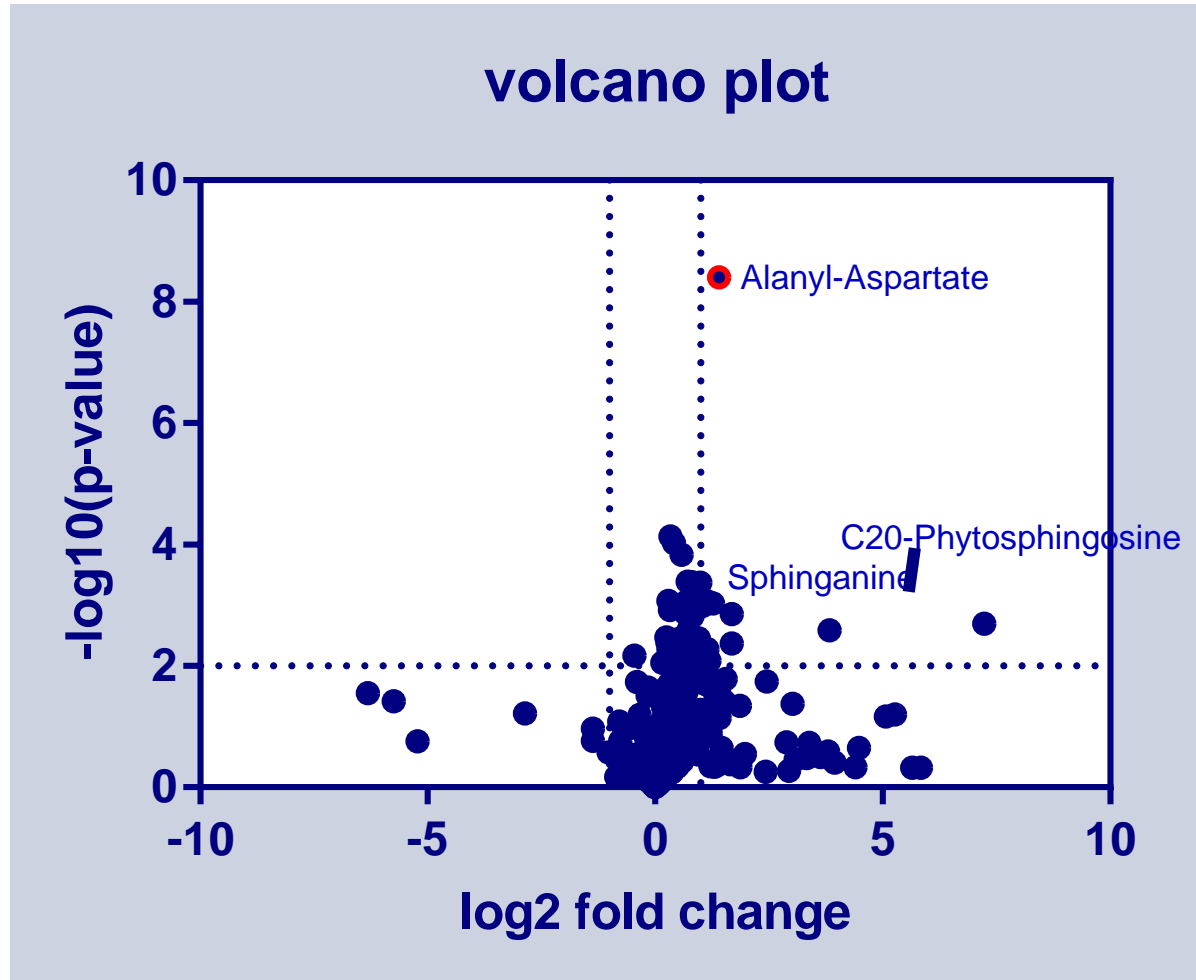We want to see the relative intensities to control group.

Practice: Group Bar Graph

We want to see the relative intensities of all metabolites in TCA cycle

TCA Cycle

Practice: volcano plot

Volcano plot: significance vs magnitude of changes in metabliltes

More information

- Prism user guide
- Prism statistics guide
- Prism Tips

zhengtao.xiao@duke.edu

Next lecture

An Introduction to R

Notes on R: A Programming Environment for Data Analysis and Graphics
Version 3.6.1 (2019-07-05)

https://cran.r-project.org/doc/manuals/r-release/R-intro.pdf

https://github.com/zhengtaoxiao/NCSU_R