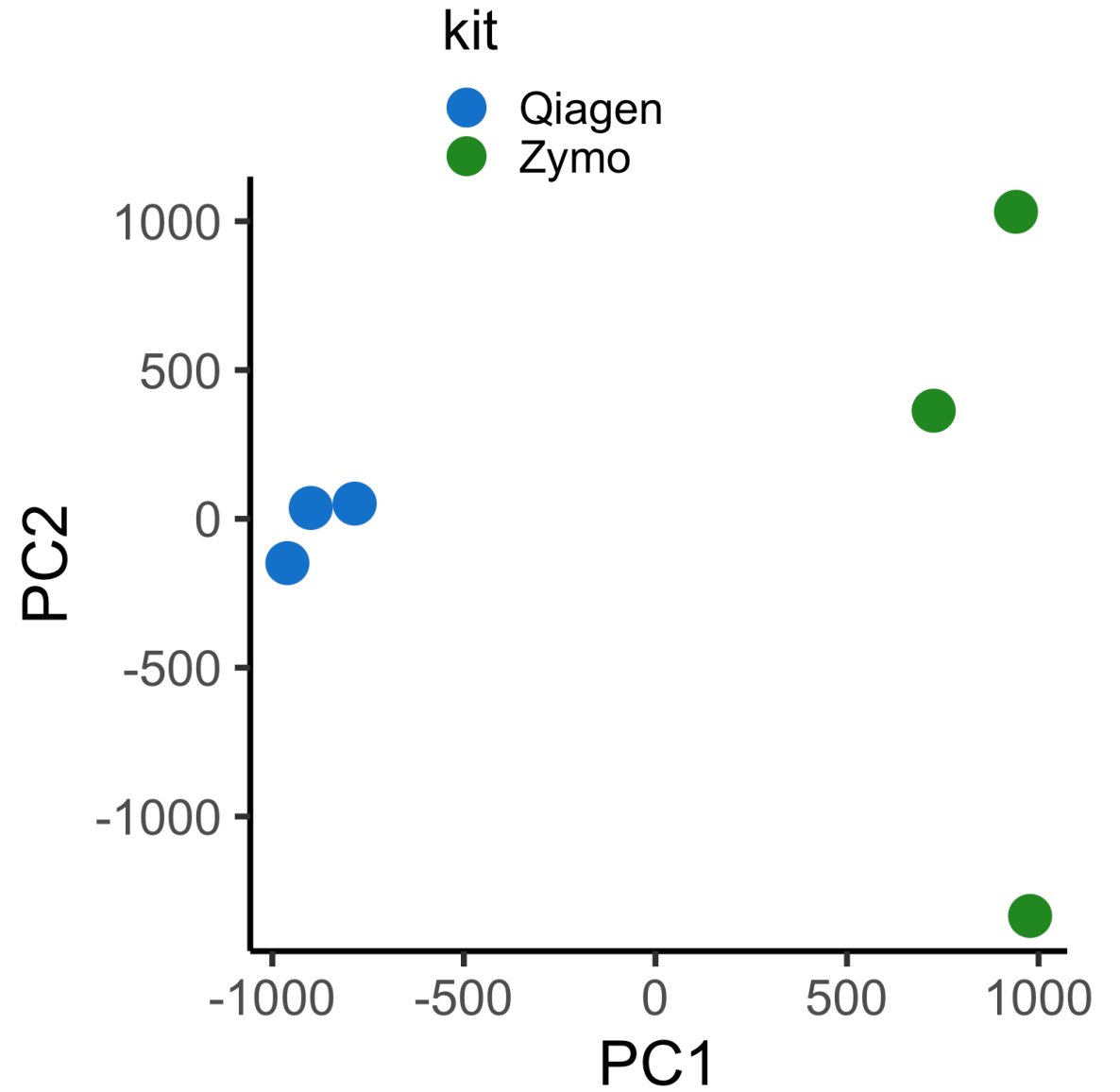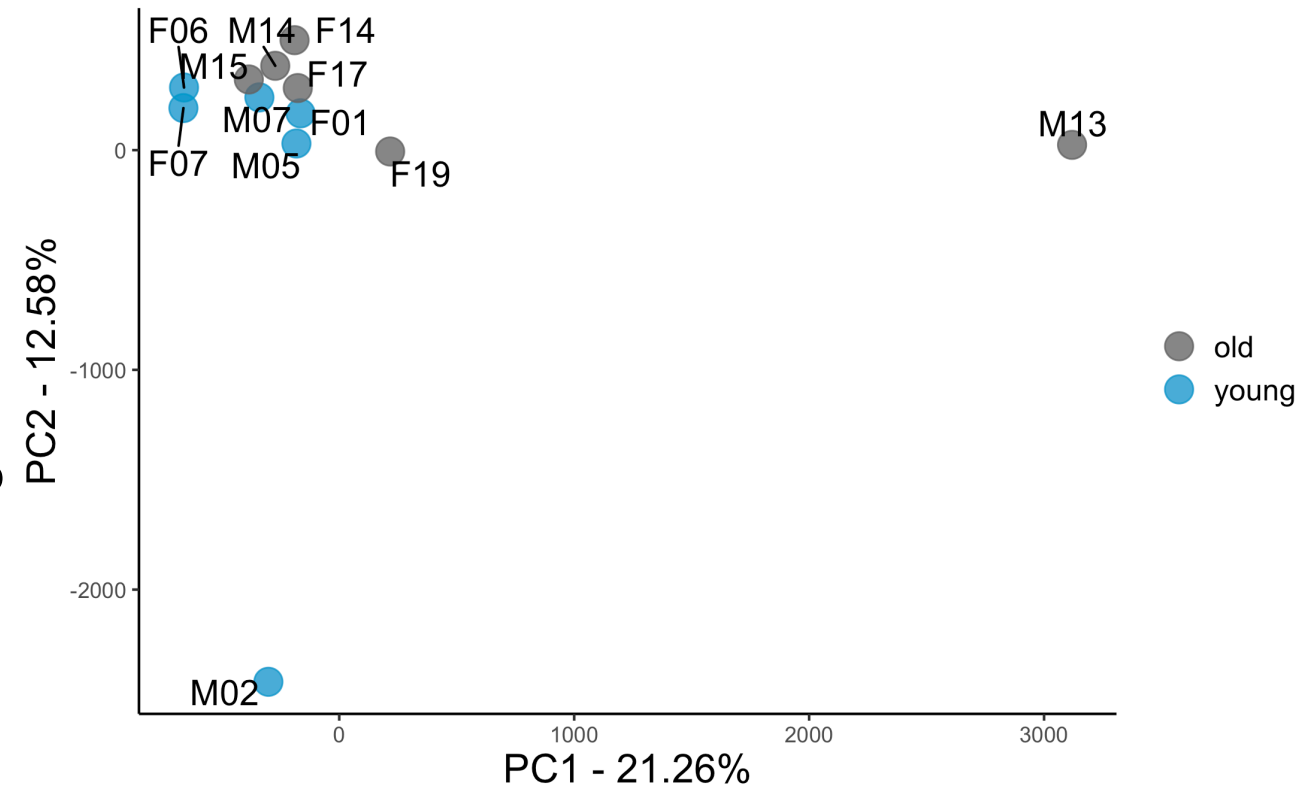# Analyze RNA-seq

2020-07-28

Clustering

- Quality control
  - Is there a batch effect?
  - Is one sample different from the others?
- Do I see expected differences?
- Do I see unexpected differences?

Clustering

- Quality control
  - Is there a batch effect?
  - Is one sample different from the others?

- Do I see expected differences?
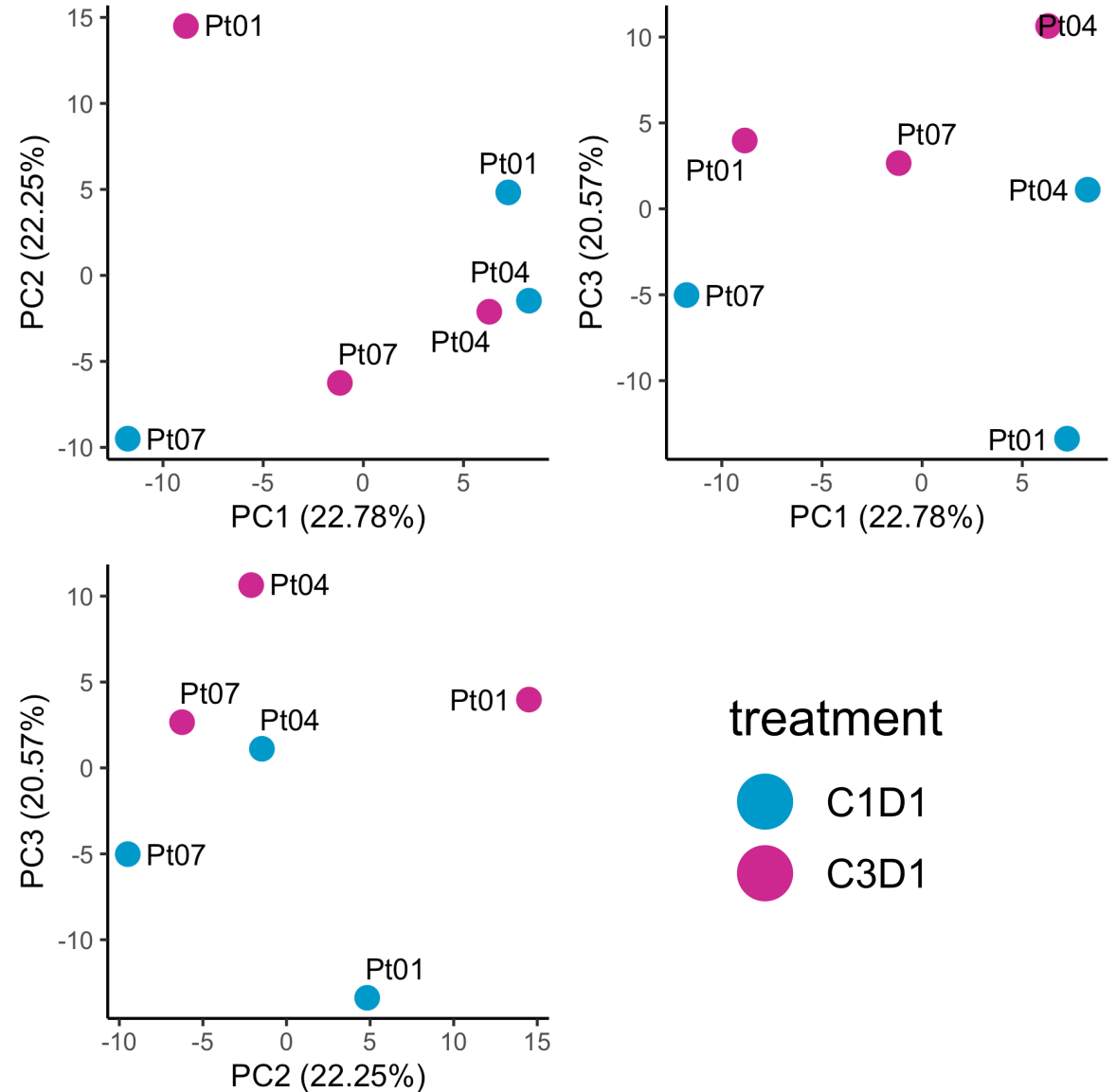
- Do I see unexpected differences?

Clustering

- Quality control
  - Is there a batch effect?
  - Is one sample different from the others?
- Do I see expected differences?
- Do I see unexpected differences?

Differential Expression

- Which genes changed expression between control and (whatever variable you manipulated)?

- Test every gene for a difference with a specific type of linear model for RNA-seq
  - log2 fold change between conditions
  - p-value

- Have to correct the results of the tests for multiple testing

Gene	Mean Expression Control	Mean Expression Treatment
Gene A	2	4
Gene B	2	1

Gene A

$$fold\ change\ (FC) = \frac{treatment}{control} = \frac{4}{2} = 2$$
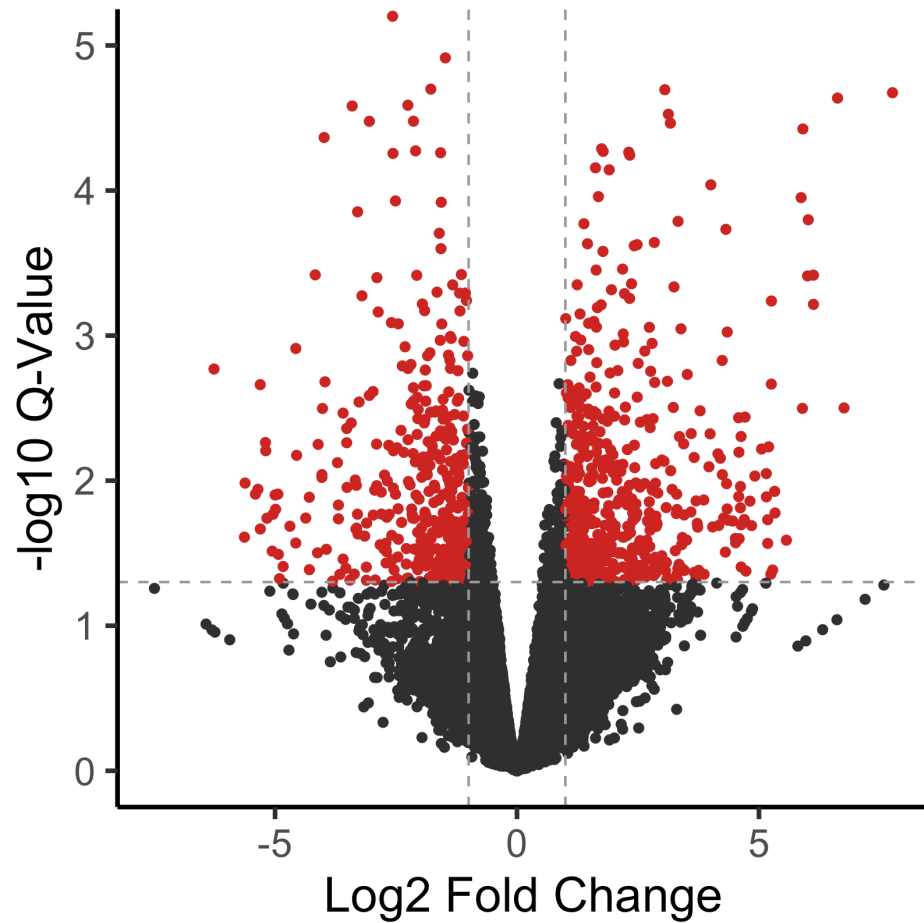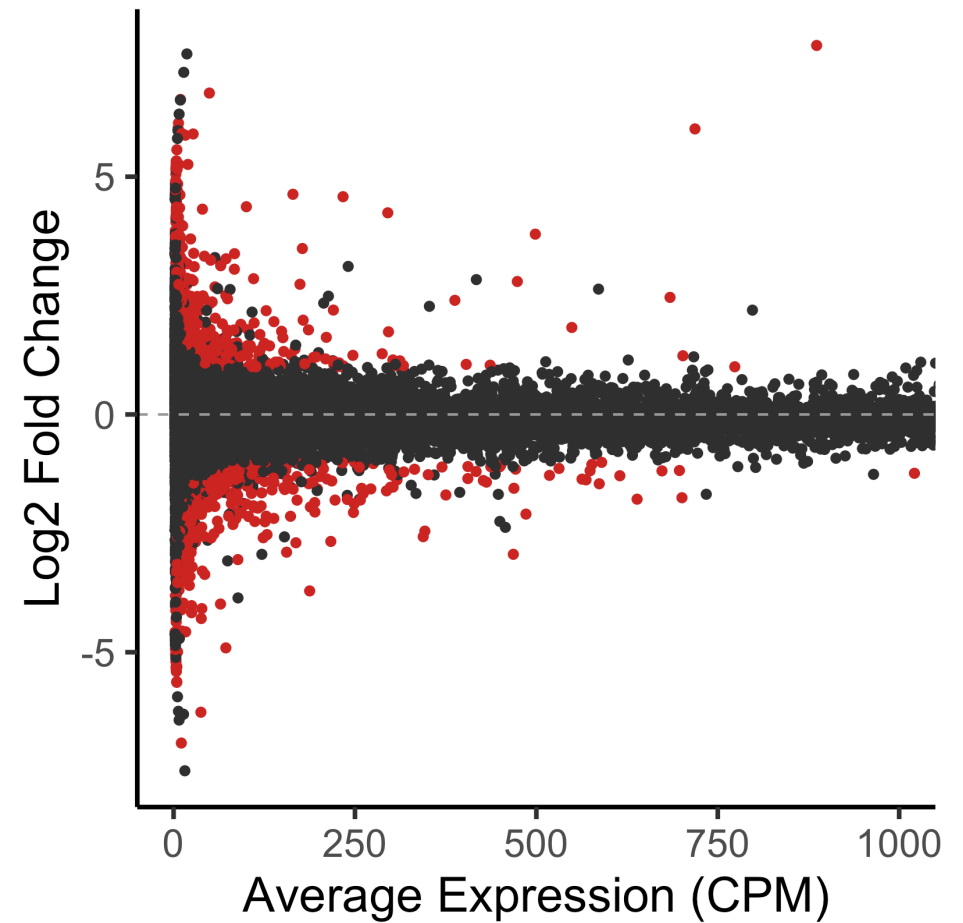
$$log_2 FC = log_2 2 = 1$$

Gene B

$$log_2 FC = log_2 \frac{1}{2} = -1$$

Visualize Results

Volcano Plot

MA Plot

Pathway/Gene Set Enrichment Analysis

- Summarize data

- What sets of genes/signalling pathways/cell circuits have changed?

- Testing for whether you see more genes that belong to the pathway or gene set than expected.