

# Project B-101-218

## Differential Expression Analysis Report

2022-06-30

Pipeline: RNASEQ\_DIFFERENTIAL\_GENE\_V1

Version: 6070

Pipeline run number: 1

PI: Horwitz Steven

PI email: horwitzs@mskcc.org

Investigator: Zachary Epstein-Peterson

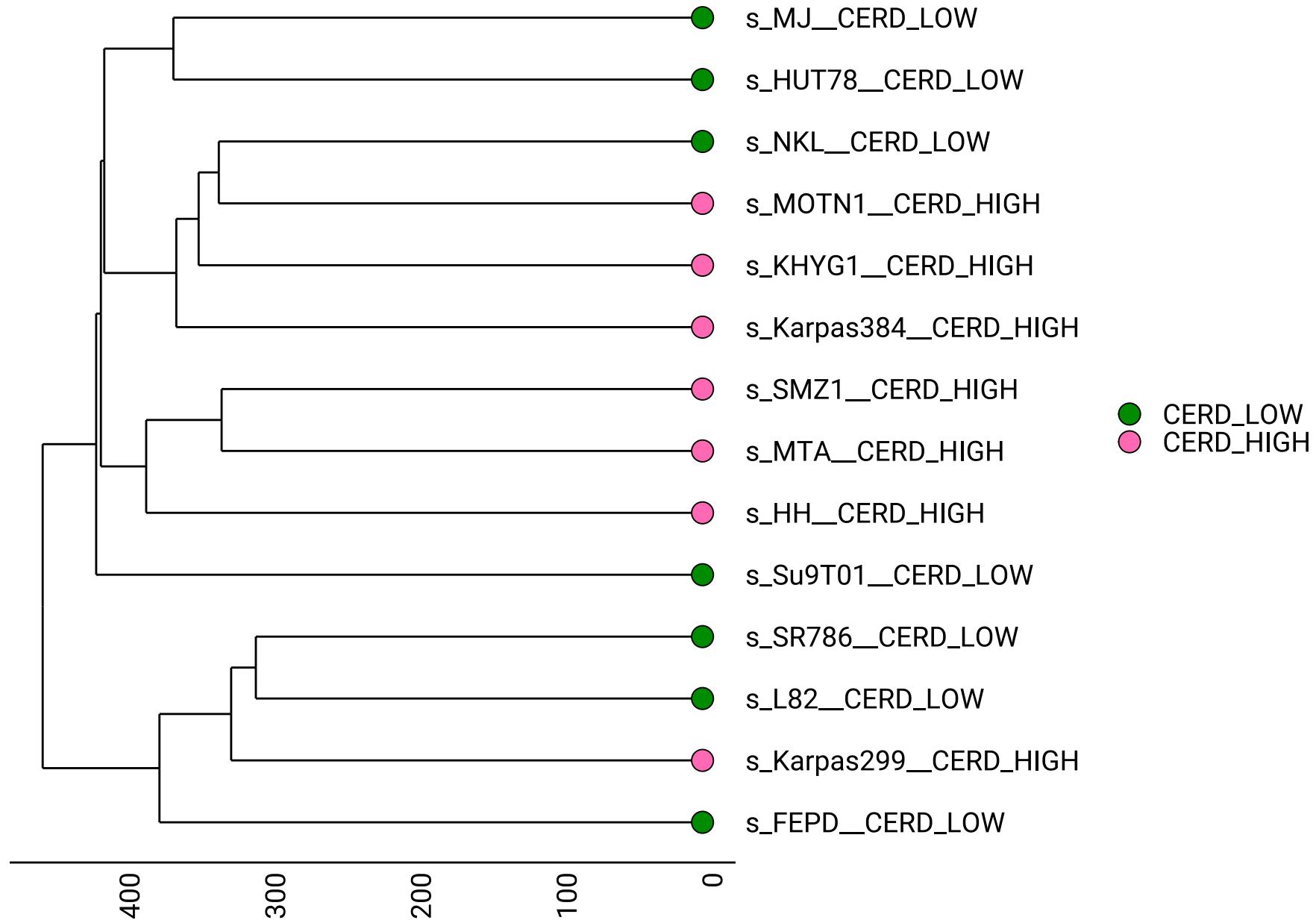
Investigator email: epsteinz@mskcc.org

# Report Contents

Hierarchical clustering	1
MDS	2
MDS (samples labeled)	3
PCA	4
PCA (samples labeled)	5
PC loadings	6
Expression of top 50 genes	7
Sample to sample distances	8
CERD_HIGH_vs_CERD_LOW Volcano plots	9
CERD_HIGH_vs_CERD_LOW DE heatmap	10
Dispersion Estimates	11
P-value histograms	12
MA plots	13

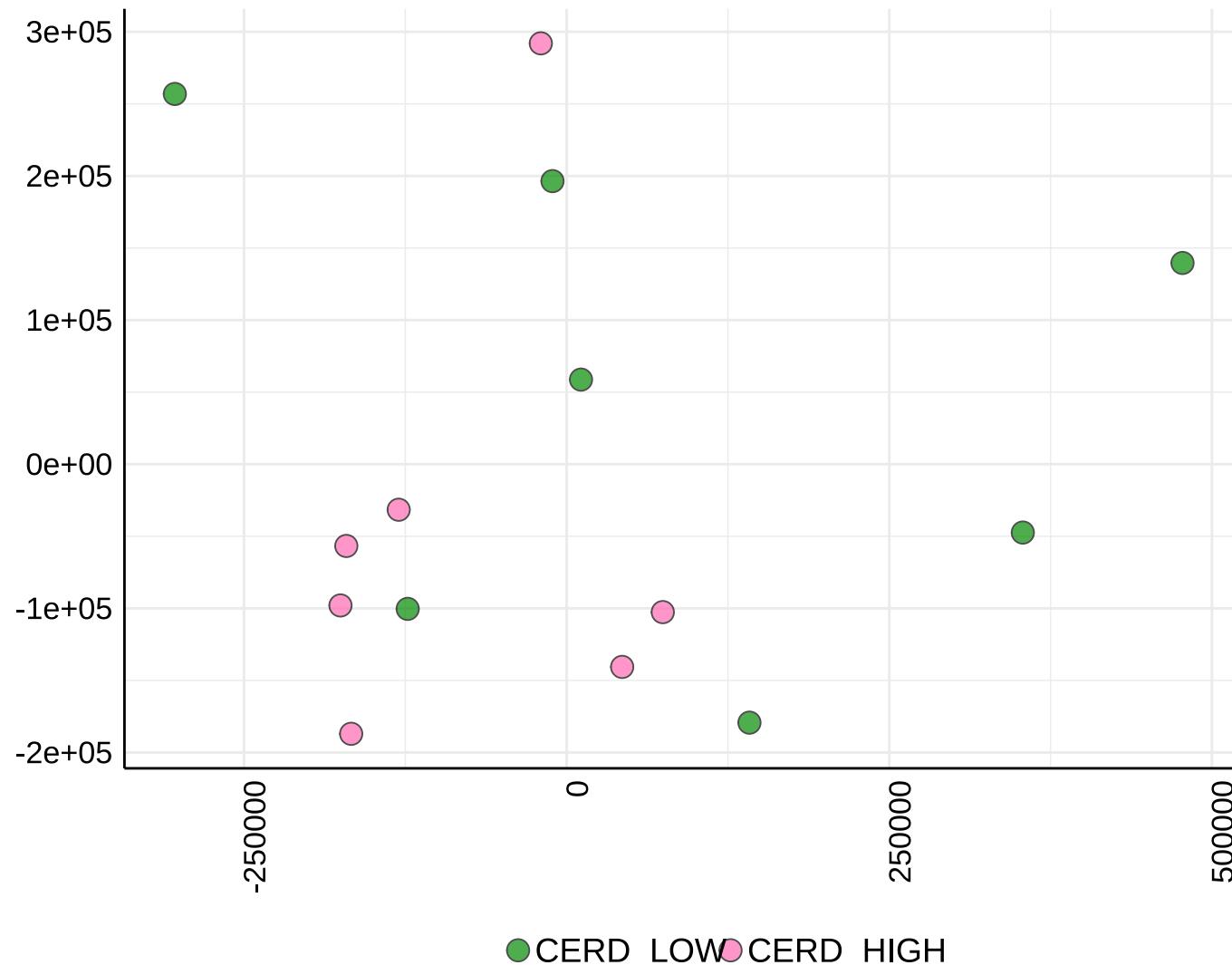
# Unsupervised clustering

All counts scaled using DESeq method



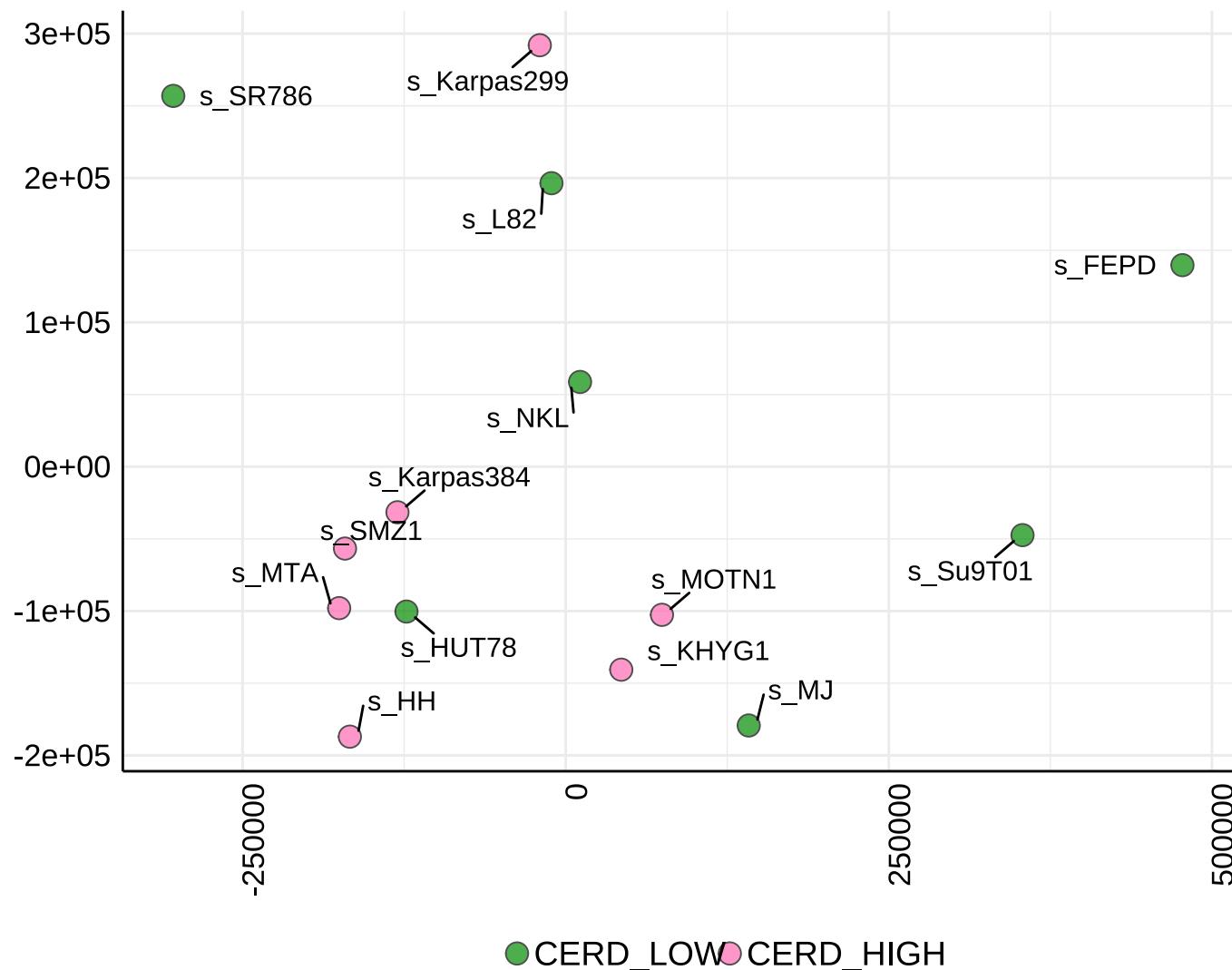
# Multidimensional Scaling (MDS)

*All counts scaled using DESeq method*

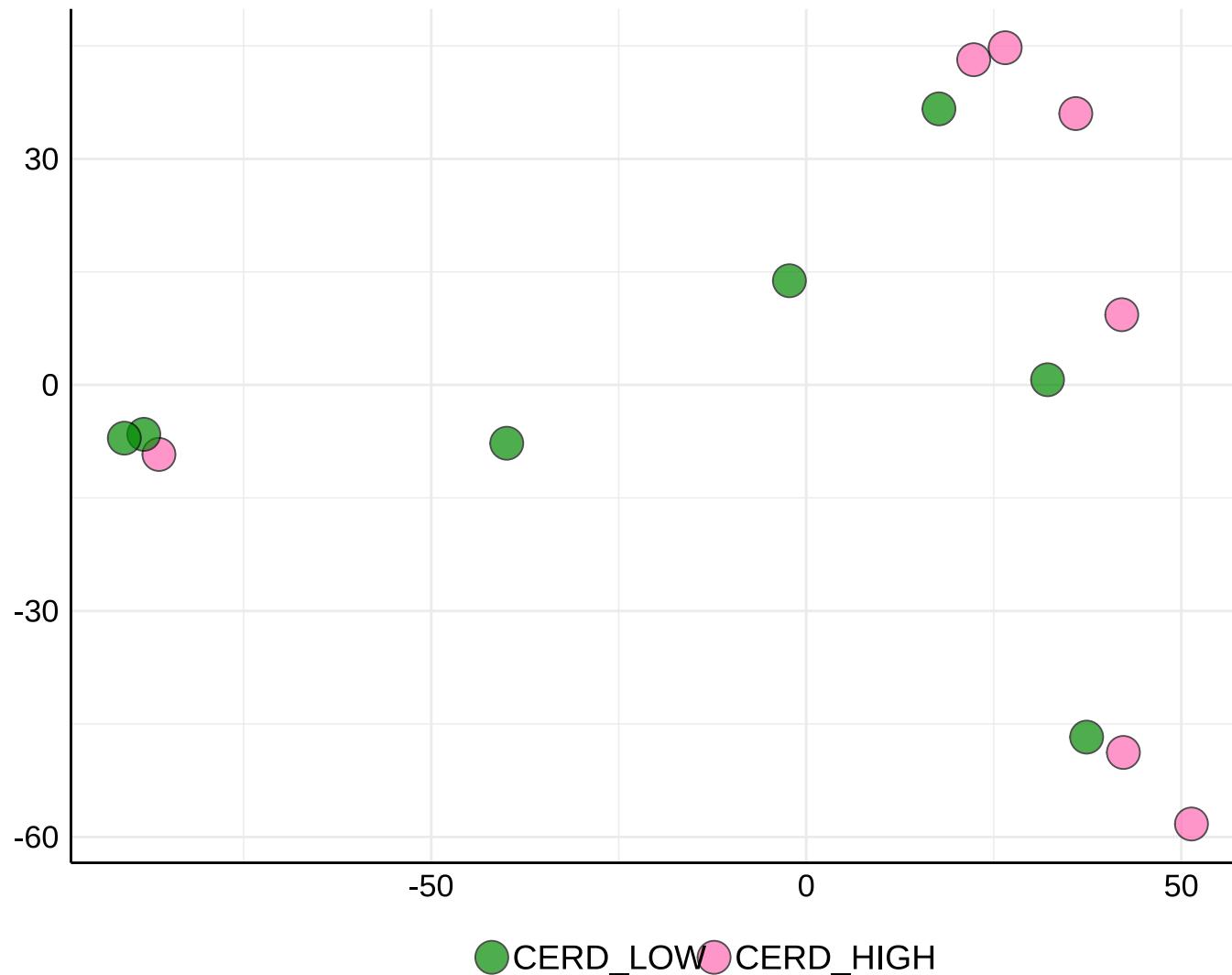


# Multidimensional Scaling (MDS)

All counts scaled using DESeq method

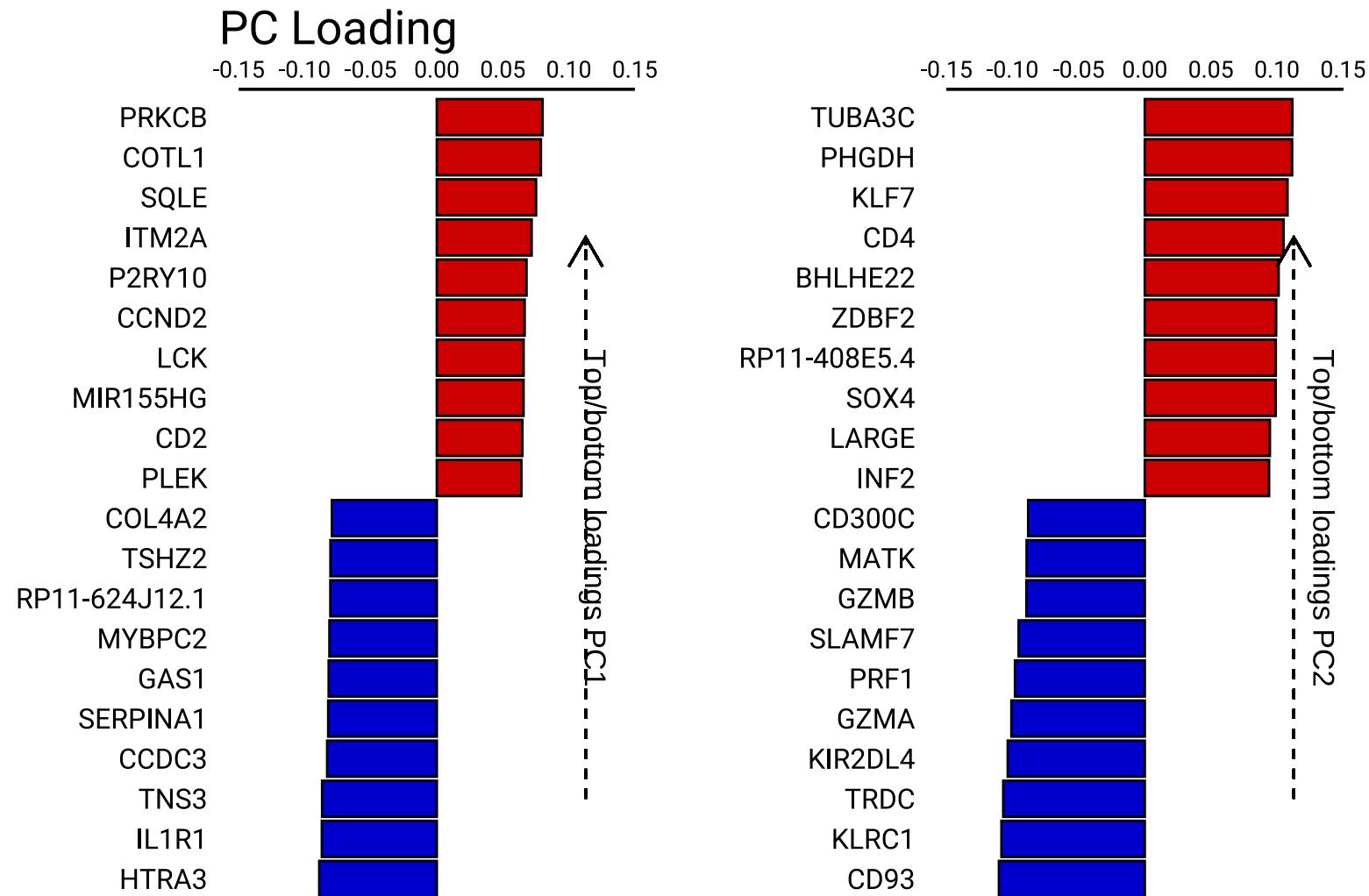


# PCA



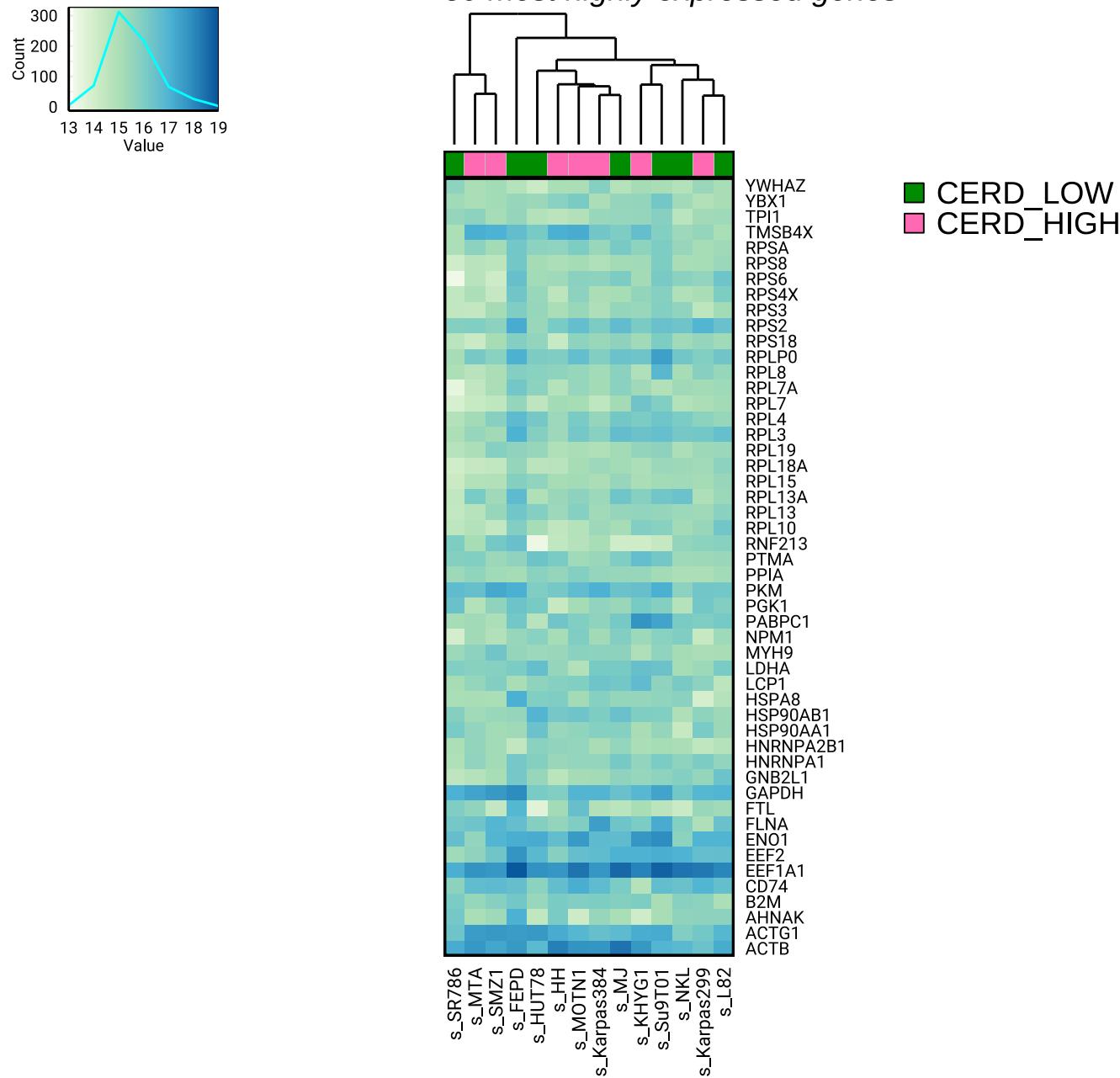
# PCA



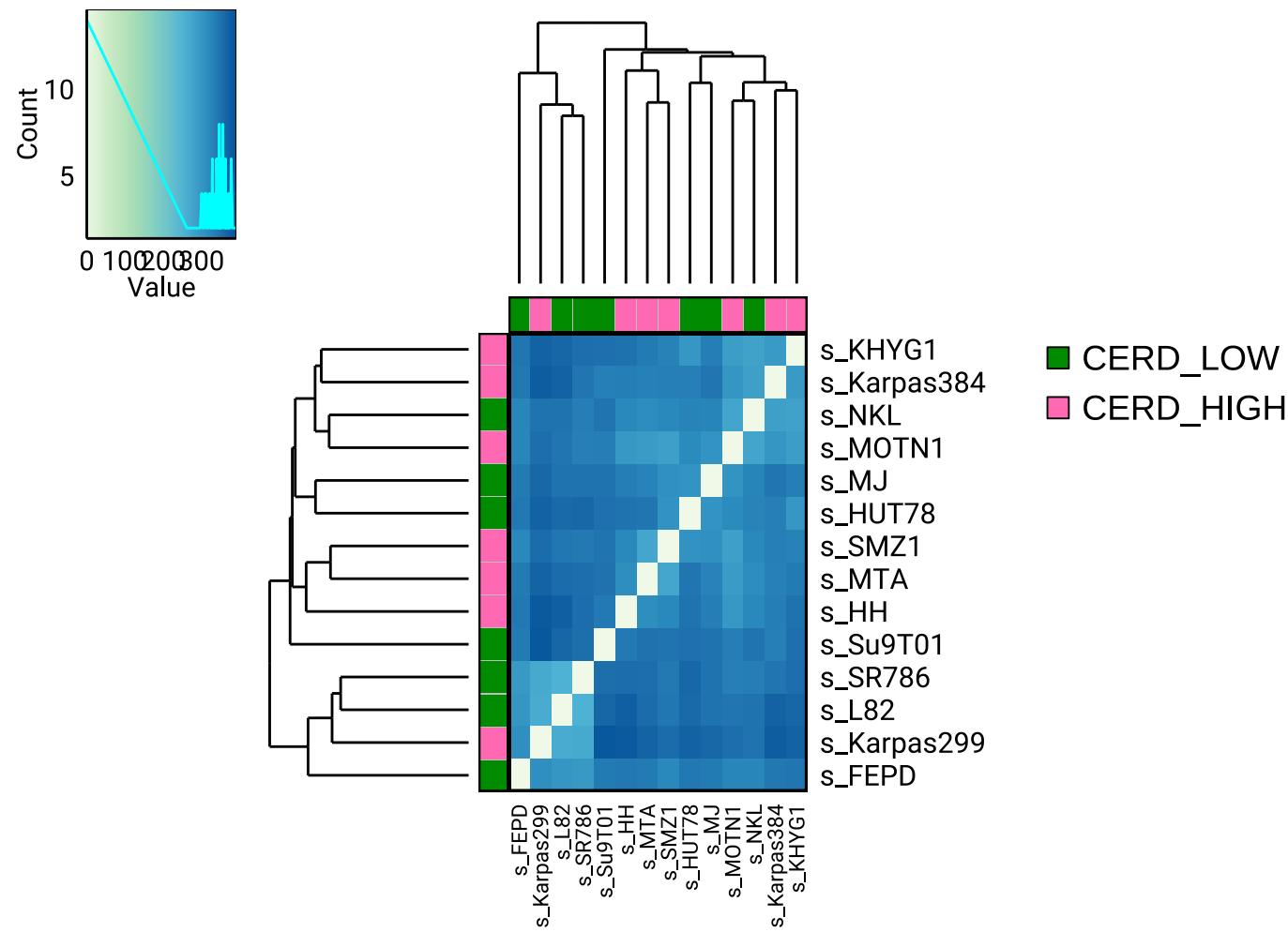


# Normalized Gene Expression

50 most highly expressed genes

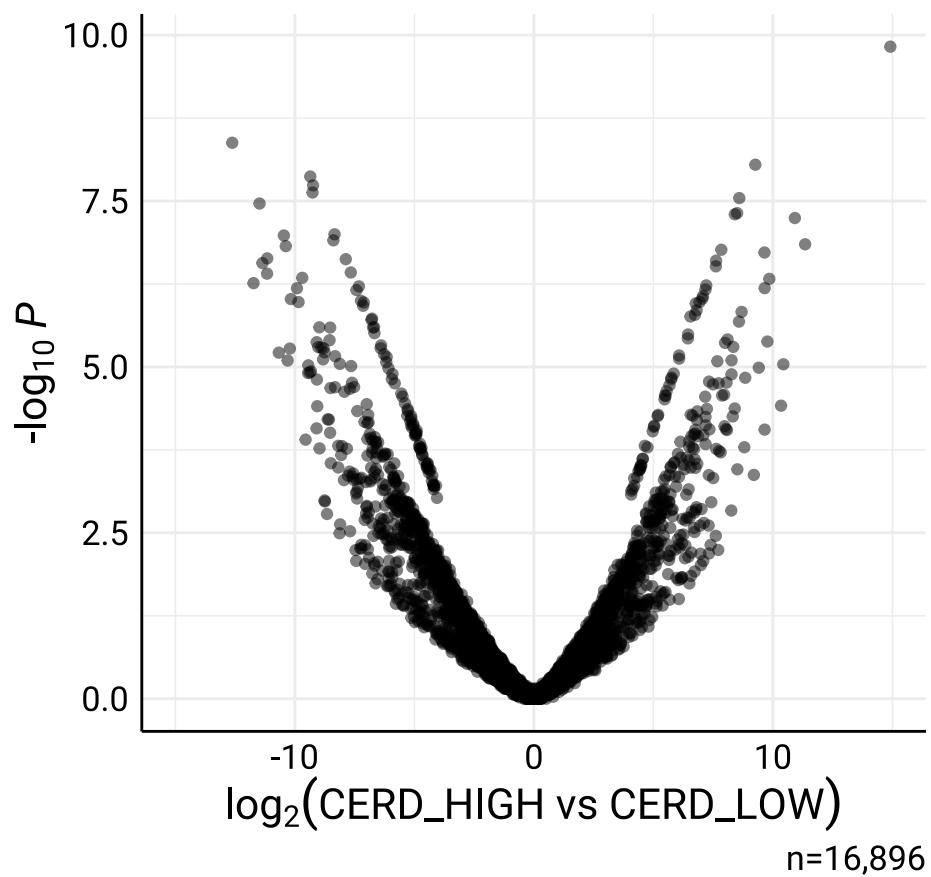


## Sample to Sample Distances

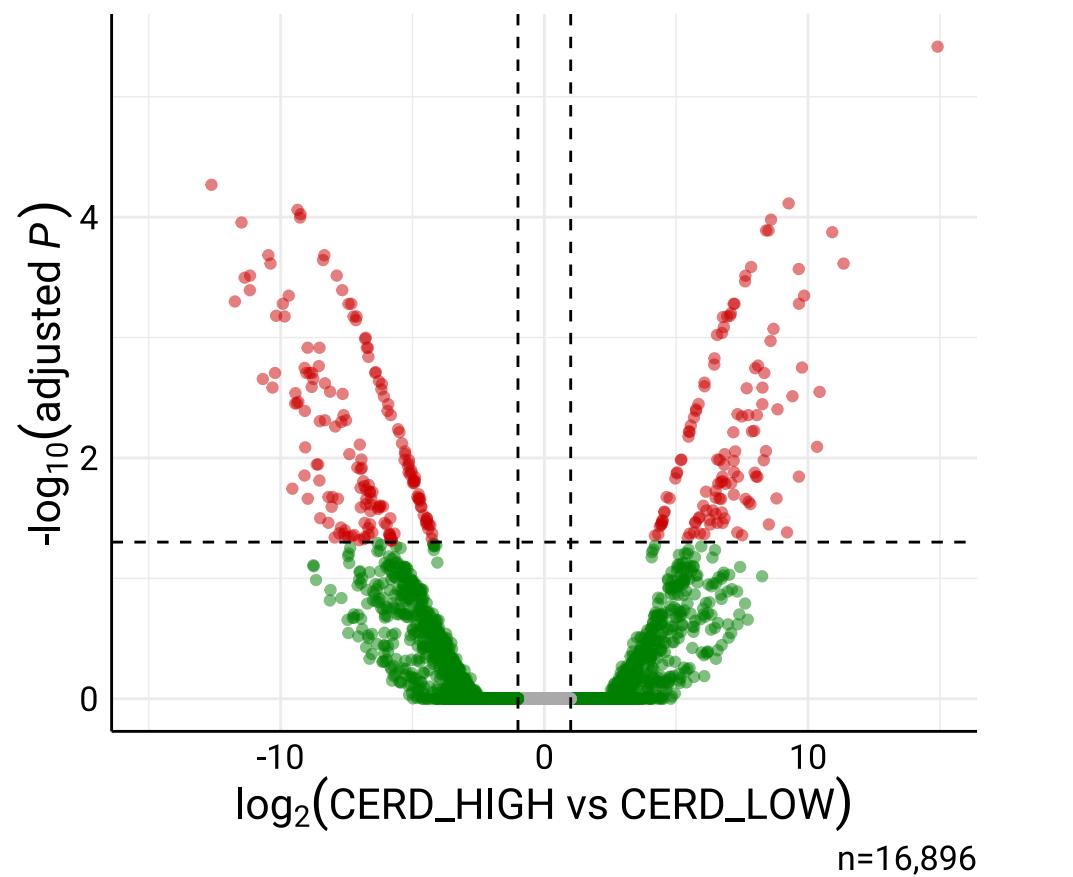


## CERD\_HIGH vs CERD\_LOW

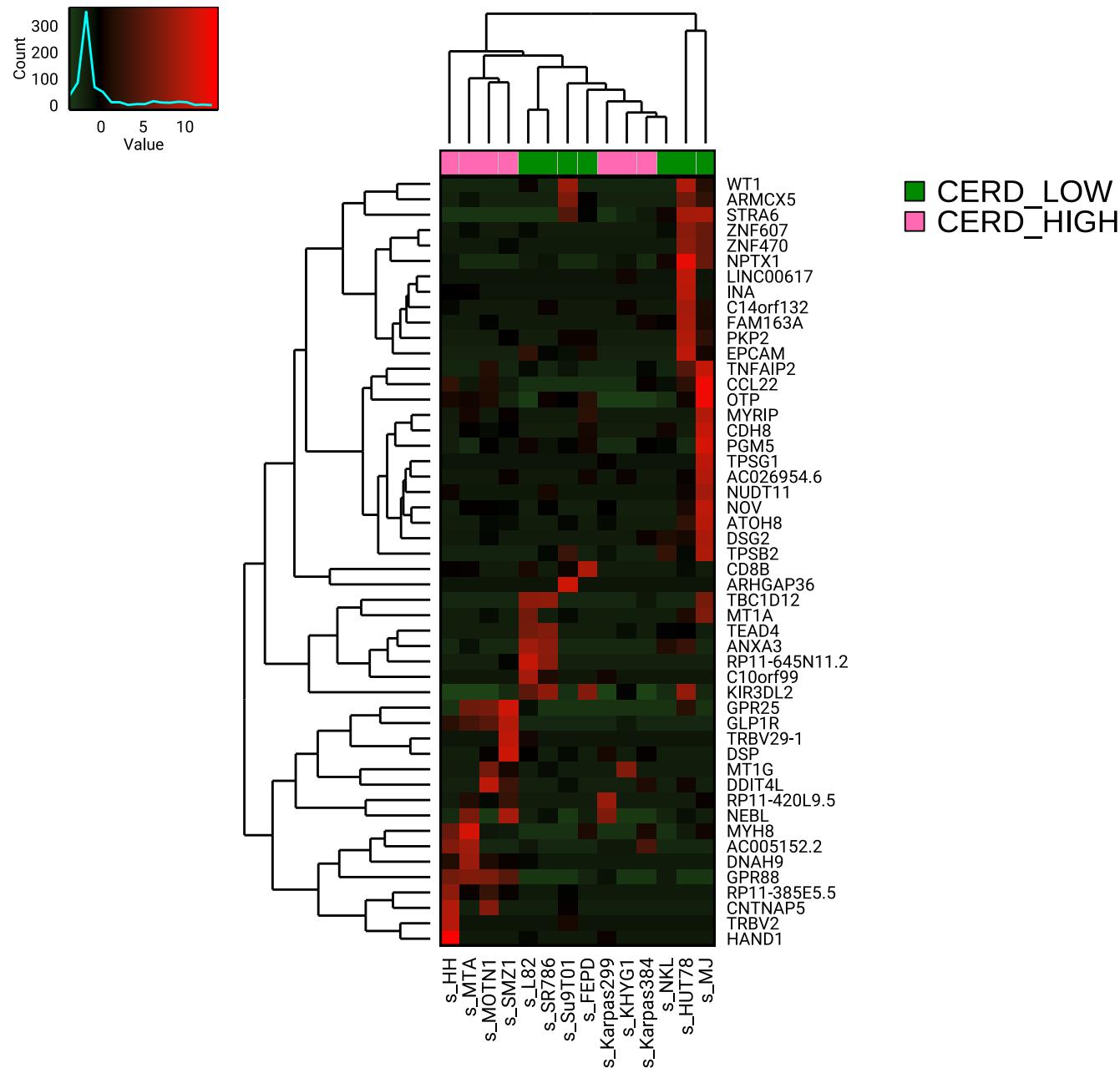
• NS



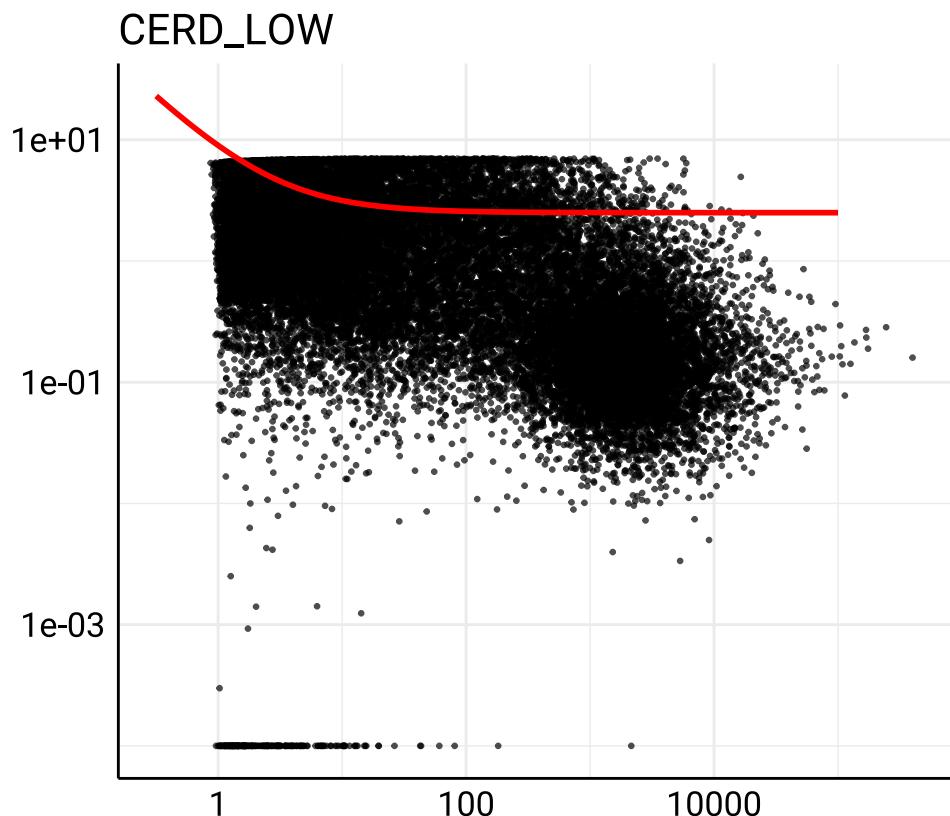
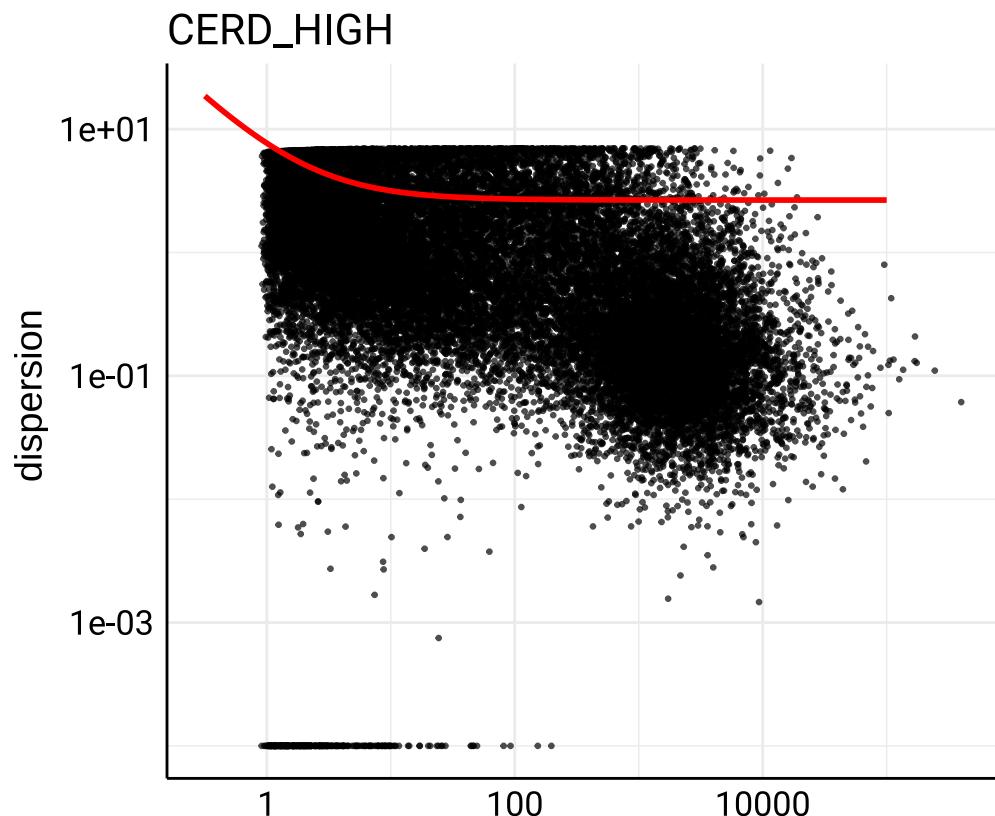
•  $\log_2(\text{CERD\_HIGH/CERD\_LOW})$  and  $-\log_{10} P$



## CERD\_HIGH vs CERD\_LOW

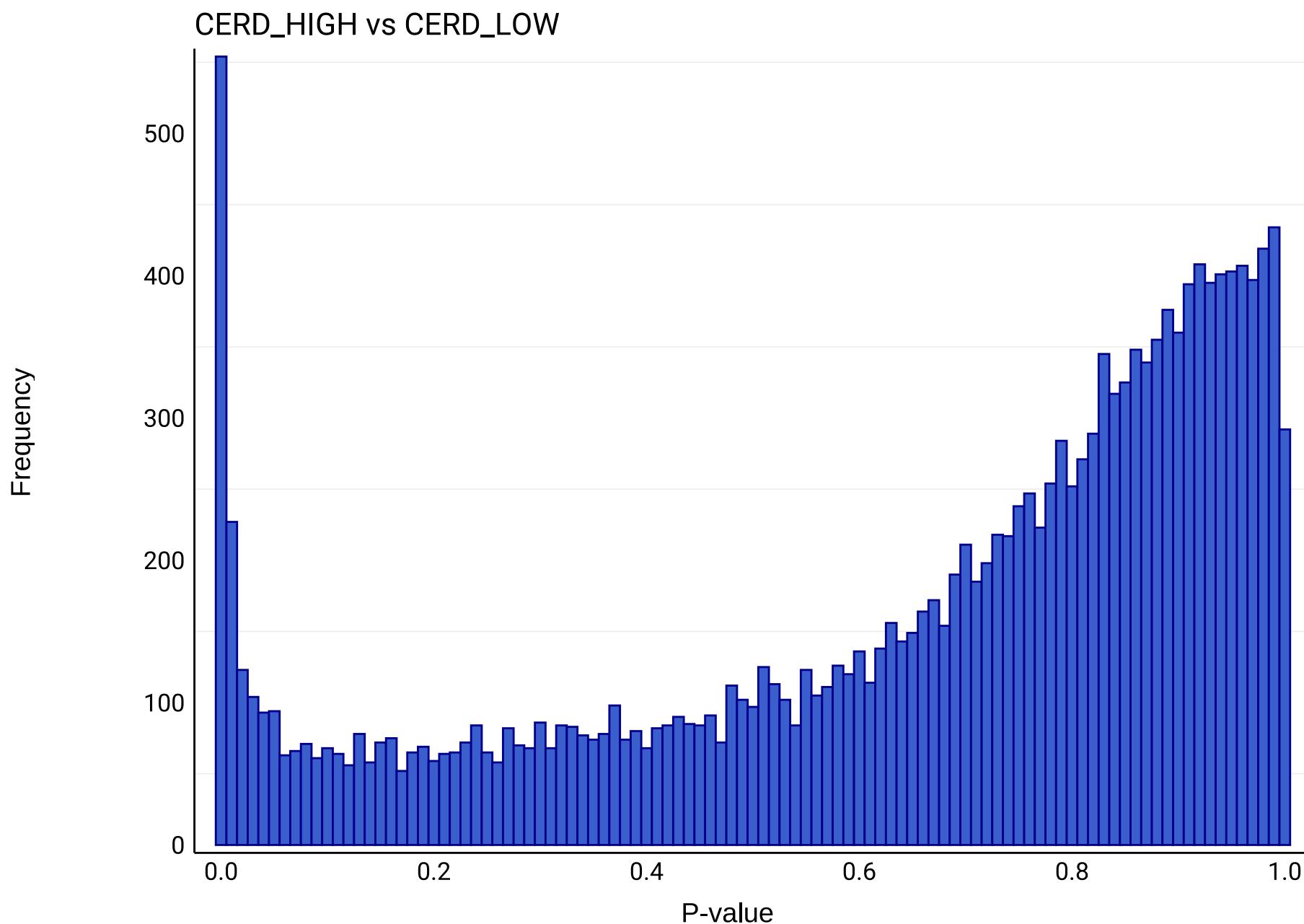


# Dispersion Estimates



mean of normalized counts

# P-value Histograms



## MA plots

