Analysis of Rubenstein-Taybi mice

Kasper D. Hansen

Fall 2022

Data

Analysis is based on the paper "Leveraging the Mendelian disorders of the epigenetic machinery to systematically map functional epigenetic variation." by Luperchio, Boukas et al (eLife 2021).

Data is RNA-seq data from B- and T-cells from two groups of mice (sorting by magnetic beads). One group is a mouse model for Rubenstein-Taybi syndrome, the other is wild-type. The Rubenstein-Taybi mutation is a heterozygous loss-of-function.

Notes

Our focus here is on the steps performed in a typical RNA-seq analysis with voom. This is not a publication level analysis.

We only focus on 2 genotypes.

We do not control for batch effects.

Loading the data

We load the edgeR package and the data in the form of a RangedSummarizedExperiment which might be familiar. edgeR uses its own data structure, a DGEList.

```
library(edgeR)
```

```
## Loading required package: limma
```

```
library(SummarizedExperiment)
```

```
## Loading required package: MatrixGenerics
```

```
## Loading required package: matrixStats
```

```
##
## Attaching package: 'MatrixGenerics'
```

```
## The following objects are masked from 'package:matrixStats':
##
##     colAlls, colAnyNAs, colAnys, colAvgsPerRowSet, colCollapse,
##     colCounts, colCummaxs, colCummins, colCumprods, colCumsums,
##     colDiffs, colIQRDiffs, colIQRs, colLogSumExps, colMadDiffs,
##     colMads, colMaxs, colMeans2, colMedians, colMins, colOrderStats,
```

Looking at the data

```
plotMDS(dge, label =  dge$samples$genotype)
```

Analysis

First, filtering followed by library size estimation / normalization.

```
keep.exprs <- filterByExpr(dge, group=dge$samples$genotype)
dge <- dge[keep.exprs,, keep.lib.sizes=FALSE]
dge <- calcNormFactors(dge, method = "TMM")
```

Then we set of the design matrix. Note that the design matrix uses CBP as reference and therefore high fold-change are down-regulated in the mutant.

```
design <- model.matrix( ~ genotype, data = dge$samples)
colnames(design) <- gsub("genotype", "", colnames(design))
design
```

```
##           (Intercept) WT
## 8_B_CBP             1  0
## 5_B_WT              1  1
## 16_B_WT             1  1
## 15_B_WT             1  1
## 21_B_CBP            1  0
## 6_B_WT              1  1
## 10_B_CBP            1  0
```