Sequencing: Bulk RNA-seq homework

Add your names here

8/6/2021

Contents

1 Default edgeR analysis

Analyze the data using edgeR, by using the code in the RNA-seq analysis intro lecture. **In all analyses, focus on the contrast comparing DPN treatment to control at 48h.**

2 Impact of blocking

Assess the difference in number of DE genes when not blocking on patient, i.e., removing the patient effect of the model. Compare the p-value distributions between these two models (i.e., with and without blocking).

3 Analyze dataset using full-quantile normalization

3.1 Implement and apply full-quantile normalization

```
### implement FQ normalization
FQnorm <- function(counts){
  ...
}

### normalize the data using FQ
```

3.2 Visualize effect of FQ normalization

Visualize the distributions of `log1p`-transformed counts (use the `density` function) to compare sample-specific count distributions before and after FQ normalization. What's the impact of FQ normalization on the differences in distribution between samples?

3.3 `edgeR` analysis using full-quantile normalized data

Don't forget to remove the `calcNormFactors` step in the `edgeR` analysis as data have already been normalized when using FQ-normalized counts as input!

```
### use FQ-normalized data as input to the edgeR analysis
```

3.4 Compare DE genes at 5% FDR

```
### Get list of DE genes using TMM and FQ normalization

### Compare list using, e.g., a Venn diagram
```