Dirichlet-multinomial framework for differential splicing and SQTL analysis.

Malgorzata Nowicka, Mark Robinson*

October 2, 2015

Contents

## 1 Dirichlet-multinomial workflow to detect differentiall splicing (DS) in RNA-seq data

Load *DM* package.

```
suppressPackageStartupMessages(library(DM))
```

Create a `dmDSdata` object that contains counts.

```
counts <- dataDS_counts[, -1]
group_id <- dataDS_counts[, 1]
group_split <- limma::strsplit2(group_id, ":")
gene_id <- group_split[, 1]
feature_id <- group_split[, 2]
sample_id = dataDS_metadata$sample_id
group = dataDS_metadata$group
d <- dmDSdata(counts = counts, gene_id = gene_id, feature_id = feature_id,
    sample_id = sample_id, group = group)
plotData(d)
```

---

*gosia.nowicka@uzh.ch
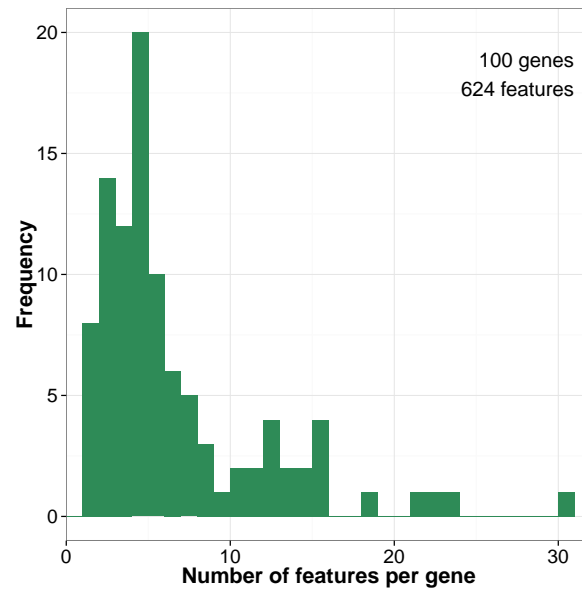
Filter genes and transcripts with low expression.

```
d <- dmFilter(d)
plotData(d)
```
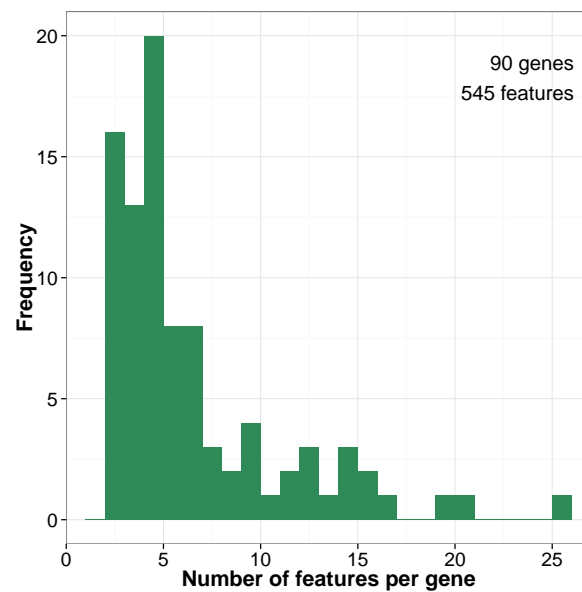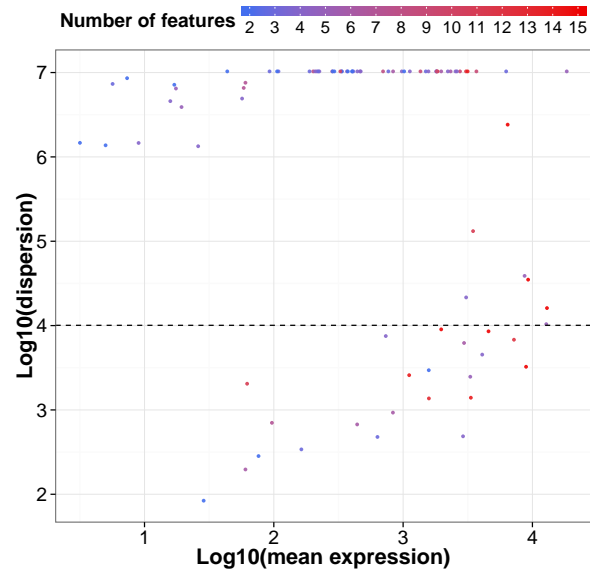

Estimate dispersion.

```
d <- dmDispersion(d, BPPARAM = BiocParallel::MulticoreParam(workers = 5))

## * Calculating mean gene expression..
## Took  0.459  seconds.
## * Estimating common dispersion..
## Took  16.767  seconds.

## !  Using common_dispersion = 10077.27 as disp_init !

## * Estimating genewise dispersion..
## Took  4.43  seconds.
```
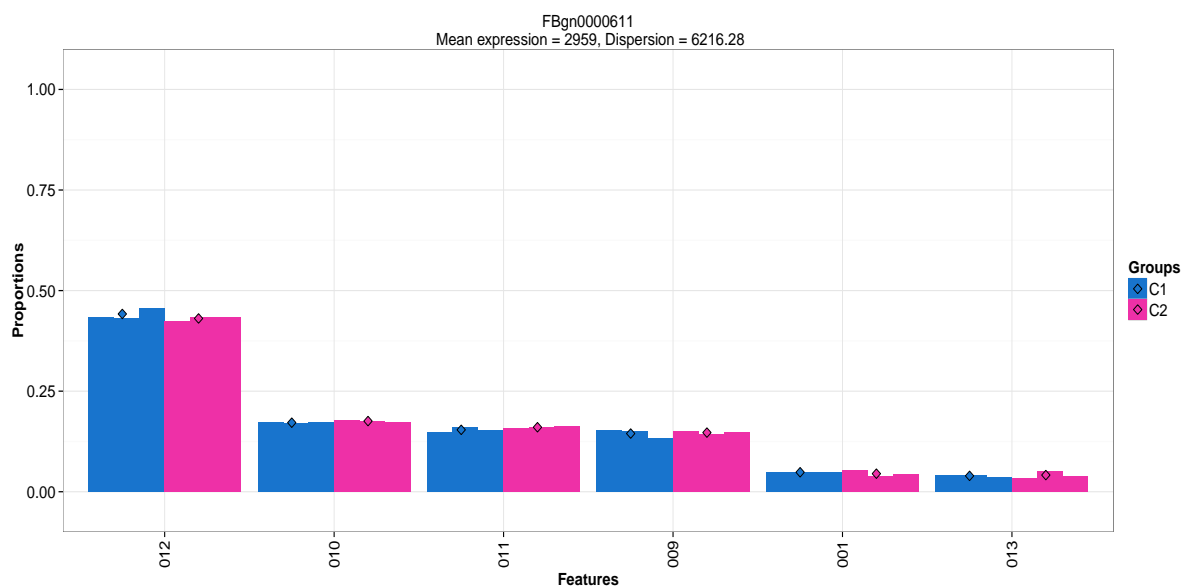
```
plotDispersion(d)
```


Estimate proportions.

```
d <- dmFit(d)

## * Fitting full model..
## Took  10.777  seconds.

gene_id <- names(d)[1]
plotFit(d, gene_id = gene_id, plot_type = "barplot")

## Plot gene 1: FBgn0000611
```


Use likelihood ratio to test for differential splicing.

```
d <- dmTest(d)

## Running comparison  between groups:  C1, C2

## * Fitting null model..
## Took  10.853  seconds.
## * Calculating likelihood ratio statistics..
## Took  0.001286983  seconds.

plotTest(d)
```
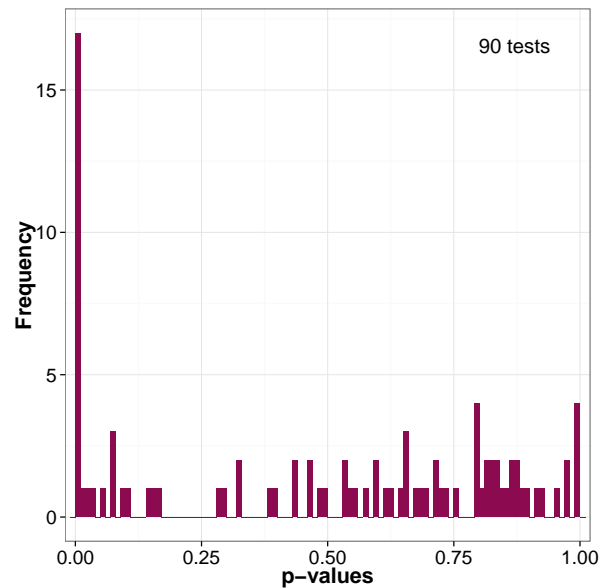


```
res <- results(d)
```

```
gene_id <- res$gene_id[1]
plotFit(d, gene_id = gene_id)

## Plot gene 1: FBgn0000611
```