Introduction to Bulk RNAseq data analysis

Annotation and Visualisation of Differential Expression Results - Solutions

Abbi Edwards

Last modified: 01 Mar 2022

Exercise 1 - Volcano plot for 33 days

Now it's your turn! We just made the volcano plot for the 11 days contrast, you will make the one for the 33 days contrast.

If you haven't already make sure you load in our data and annotation. You can copy and paste the code below.

```
# First load data and annotations
results.interaction.33 <- readRDS("RObjects/DESeqResults.interaction_d33.rds")
ensemblAnnot <- readRDS("RObjects/Ensembl_annotations.rds")
```

(a) Shrink the results for the 33 days contrast.

```
#Shrink our values
ddsShrink.33 <- lfcShrink(ddsObj.interaction,
                          res = results.interaction.33,
                          type = "ashr")
```

```
## using 'ashr' for LFC shrinkage. If used in published research, please cite:
##     Stephens, M. (2016) False discovery rates: a new deal. Biostatistics, 18:2.
##     https://doi.org/10.1093/biostatistics/kxw041
```

```
shrinkTab.33 <- as.data.frame(ddsShrink.33) %>%
    rownames_to_column("GeneID") %>%
    left_join(ensemblAnnot, "GeneID") %>%
    rename(logFC=log2FoldChange, FDR=padj)
```

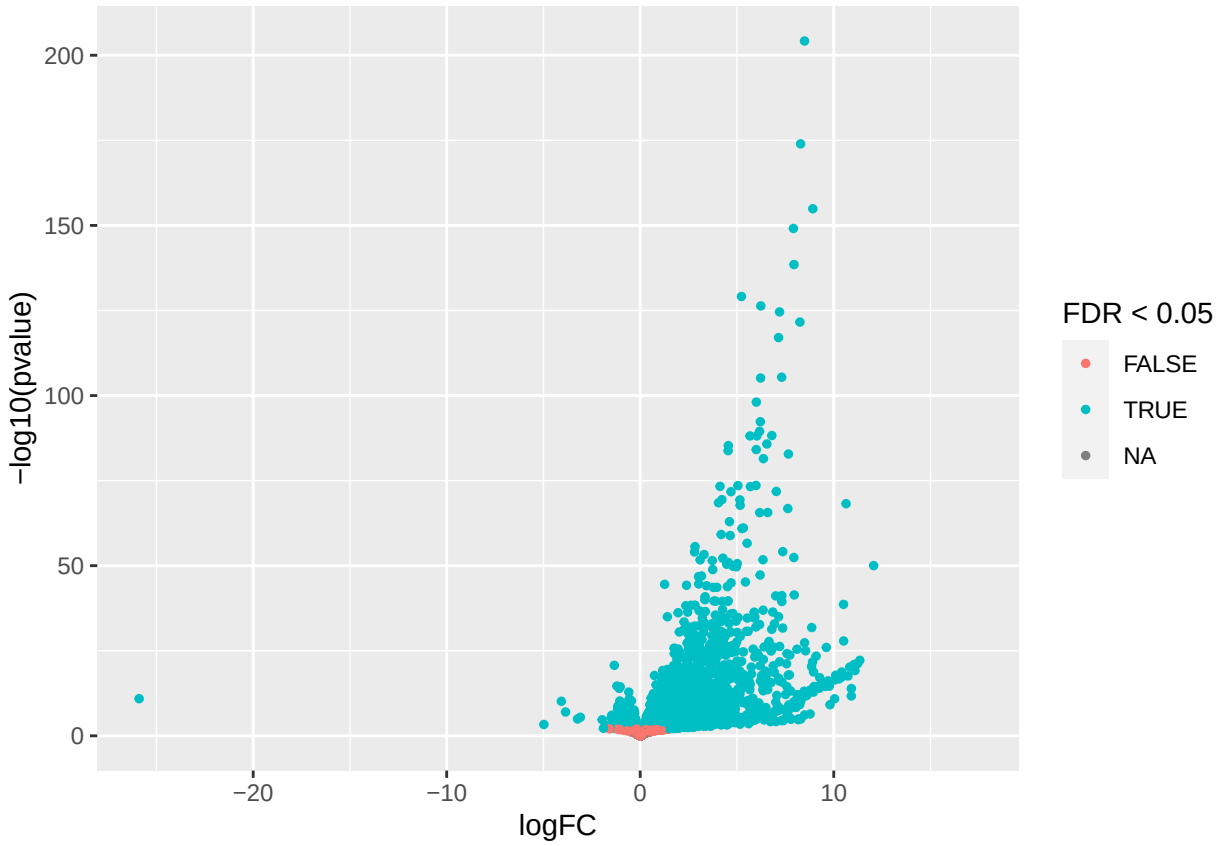
(b) Create a new column of -log10(pvalue) values in your shrinkTab for 33 days.

(c) Create a plot with points coloured by P-value < 0.05 similar to how we did in the first volcano plot

```
volcanoTab.33 <- shrinkTab.33 %>%
    mutate(`-log10(pvalue)` = -log10(pvalue))

ggplot(volcanoTab.33, aes(x = logFC, y=`-log10(pvalue)`)) +
    geom_point(aes(colour=FDR < 0.05), size=1)
```

```
## Warning: Removed 47 rows containing missing values (geom_point).
```



Exercise 2 - MA plot for day 33 with ggplot2

For this exercise create an MA plot for day 33 like the ones we plotted with `plotMA` from **DESeq2** but this time using ggplot2.

The x-axis (M) should be the log2 of the mean gene expression across all samples, and the y-axis should be the log2 of the fold change between Infected and Uninfected.

```
maTab.33 <- shrinkTab.33 %>%
    mutate(`M` = log2(baseMean))

ggplot(maTab.33, aes(x = M, y = logFC)) +
    geom_point(aes(colour=FDR < 0.05), size=1) +
    scale_y_continuous(limit=c(-4,4), oob = scales::squish)
```


4
2
0
−2
−4
logFC
0
5
10
15
20
M
FDR < 0.05
FALSE
TRUE
NA