

DGEs_human_donortime_plots

Purpose:

To plot the human/HBV (with donor and time serving as factors in the design) as volcano plots.

Required libraries

```
library(dplyr)

##
## Attaching package: 'dplyr'
## The following objects are masked from 'package:stats':
##
##   filter, lag
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
library(ggplot2)
library(ggrepel)
library(stringr)
```

Visualizing the DGE output by volcano plots.

```
##The location of the DGE files just generated so that they can be read in
DGE_files <- "Human DGEs_donortime"
DGE_R <- basename(Sys.glob(file.path(DGE_files, "*.txt")))
DGE_R

## [1] "2018-03-13human_donor_timehumangenes-d28_vs_d8_HBV_analysis_results.txt"
## [2] "2018-03-13human_donor_timehumangenes-d28_vs_d8_coinf_analysis_results.txt"
## [3] "2018-03-13human_donor_timehumangenes-d28_vs_d8_mock_analysis_results.txt"

##Function to read in the DGE files
exptDGEs <- function(files) {
  d <- read.delim(files, header = TRUE)
}

##Read in all the files
all_DGEs <- lapply(file.path(DGE_files, DGE_R), exptDGEs)
names(all_DGEs) <- sub(".txt", "", DGE_R)
names(all_DGEs)

## [1] "2018-03-13human_donor_timehumangenes-d28_vs_d8_HBV_analysis_results"
## [2] "2018-03-13human_donor_timehumangenes-d28_vs_d8_coinf_analysis_results"
## [3] "2018-03-13human_donor_timehumangenes-d28_vs_d8_mock_analysis_results"

##Function to generate the volcano plots; here all the axes labels are intact
##and HBV genes are labeled for reference when making paper's figures.
for(i in 1:3) {
  index <- all_DGEs[[i]]
  output_name <- str_replace(names(all_DGEs[i]), "^([0-9]*-([0-9]*-([0-9]*\\s", "")) %>%
    str_replace("_*\\s*analysis", "")
  index$fillexp <- with(index, factor(ifelse(padj > 0.05, 0, 1)))
```

```

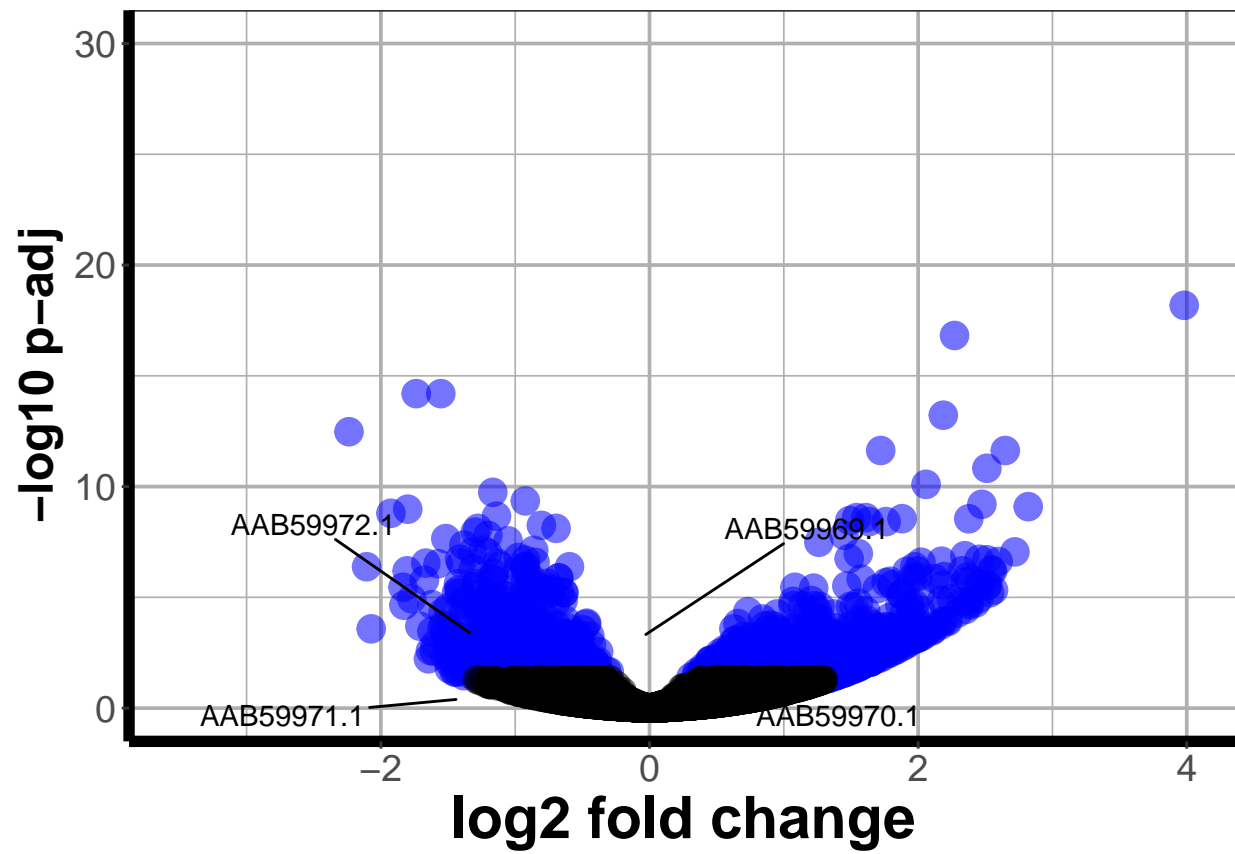
print(ggplot(data=index, aes(x=log2FoldChange, y=-log10(padj), color =filleffect)) +
  geom_point(aes(colour=filleffect, alpha=0.4, size=1)) +
  scale_colour_manual(values = c("black", "blue")) +
  theme(legend.position = "none") +
  xlab("log2 fold change") + ylab("-log10 p-adj") +
  coord_cartesian(xlim = c(-3.5, 4), ylim = c(0, 30)) +
  geom_text_repel(
    data = index[grepl("AAB59972.1|AAB59970.1|AAB59969.1|AAB59971.1", index$X),],
    mapping = aes(label = X),
    size = 4,
    box.padding = unit(0.4, "lines"),
    point.padding = unit(4, "lines")) +
  xlim(c(-4, 2.5)) + ylim(c(0, 5)) +
  theme_bw() +
  theme(legend.position = "none",
    axis.title.x = element_text(face = "bold", size = 22),
    axis.text = element_text(size = 14),
    plot.title = element_blank(),
    panel.grid.major = element_line(size = 0.65,
                                     color = "gray69"),
    panel.grid.minor = element_line(size = 0.3, color = "gray69"),
    axis.line = element_line(size = 2),
    axis.title.y = element_text(face = "bold", size = 18)))

ggsave(file.path("Human DGEs_donortime", paste(Sys.Date(), output_name,
"donortimeDGEs_unfiltered_labeled.png")), units = 'in', height = 10,
width = 10, dpi = 300, device = "png")
}

```

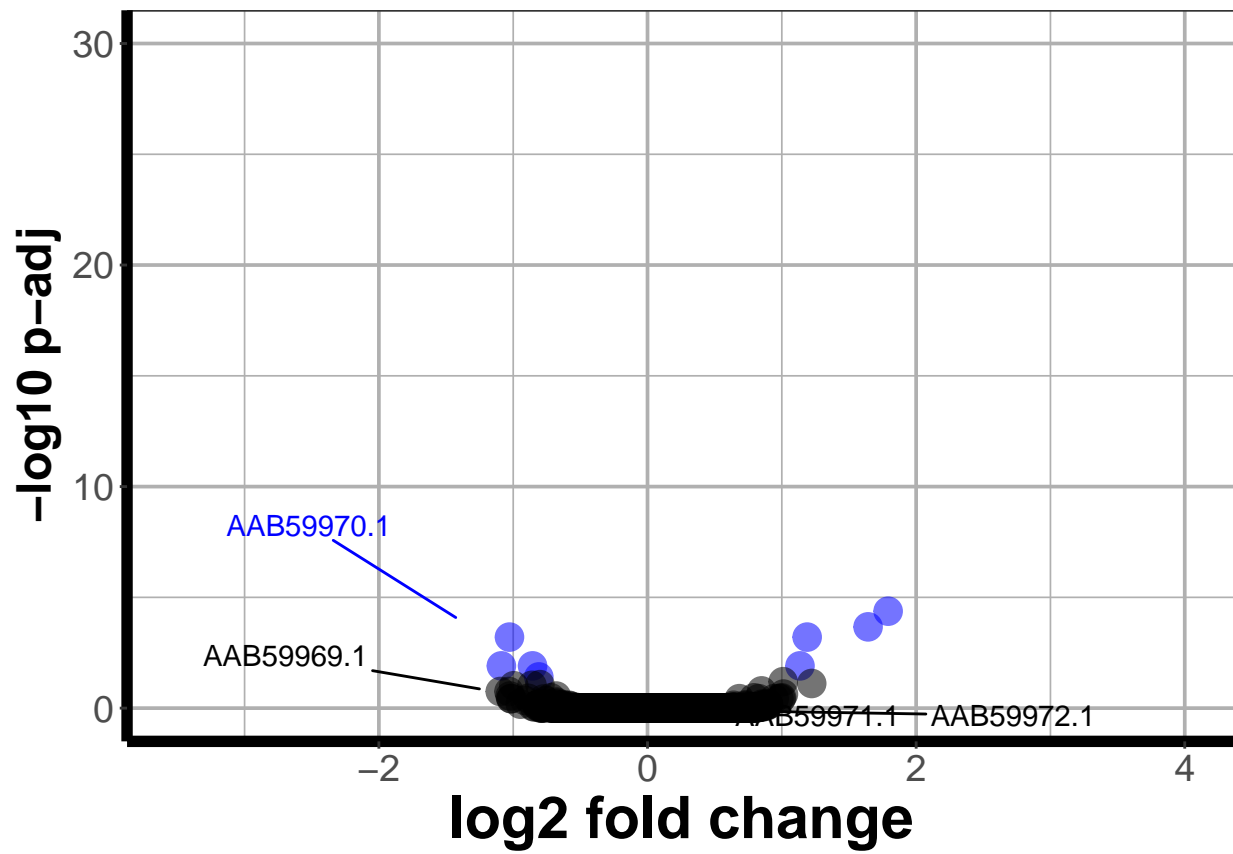
```
## Warning: Removed 44999 rows containing missing values (geom_point).
```

```
## Warning: Removed 44999 rows containing missing values (geom_point).
```

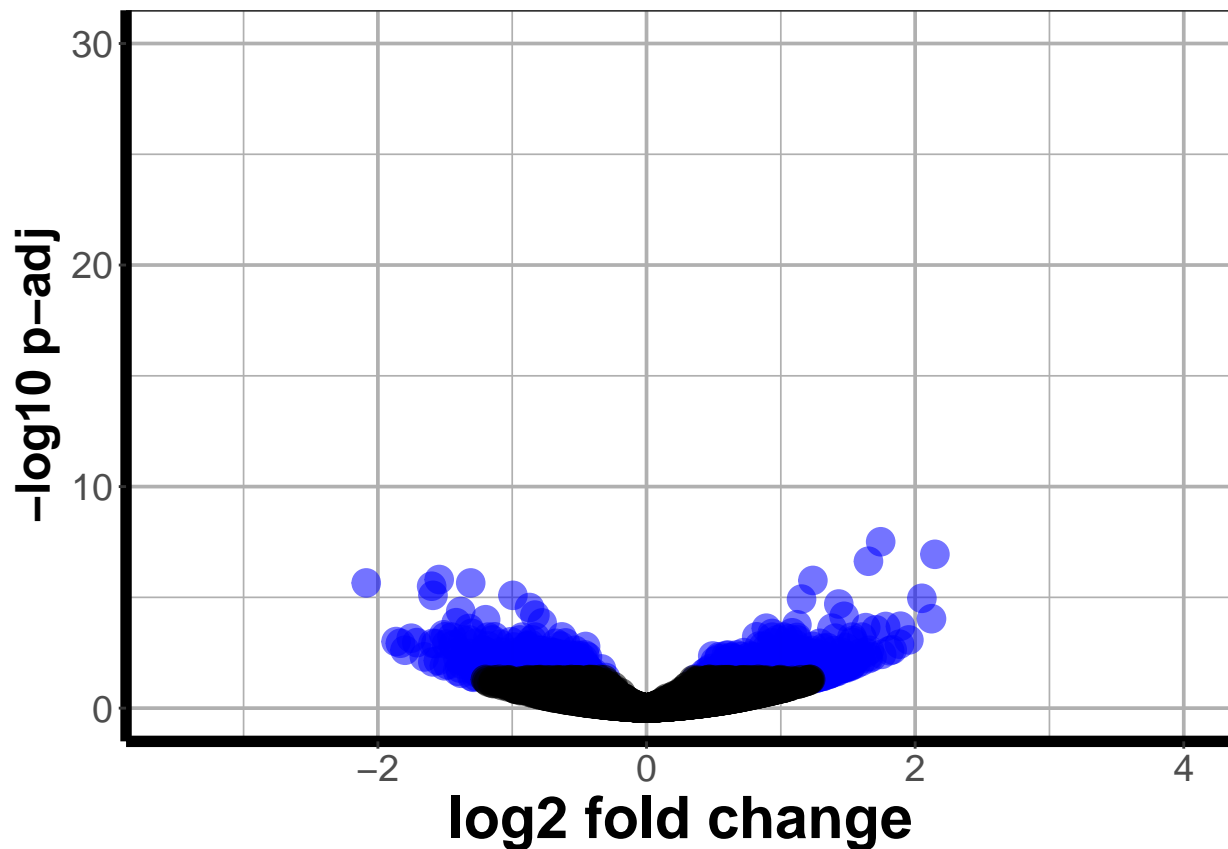


Warning: Removed 46605 rows containing missing values (geom_point).

Warning: Removed 46605 rows containing missing values (geom_point).



```
## Warning: Removed 46030 rows containing missing values (geom_point).
## Warning: Removed 4 rows containing missing values (geom_text_repel).
## Warning: Removed 46030 rows containing missing values (geom_point).
## Warning: Removed 4 rows containing missing values (geom_text_repel).
```



Session Info

```
sessionInfo()
```

```
## R version 3.3.3 (2017-03-06)
## Platform: x86_64-apple-darwin13.4.0 (64-bit)
## Running under: macOS Sierra 10.12.6
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
## [1] stats      graphics  grDevices  utils      datasets  methods   base
##
## other attached packages:
## [1] stringr_1.2.0 ggrepel_0.6.5 ggplot2_2.2.1 dplyr_0.7.3
##
## loaded via a namespace (and not attached):
## [1] Rcpp_0.12.10    knitr_1.16      bindr_0.1       magrittr_1.5
## [5] munsell_0.4.3   colorspace_1.3-2 R6_2.2.0        rlang_0.1.2
## [9] plyr_1.8.4      tools_3.3.3     grid_3.3.3      gtable_0.2.0
## [13] htmltools_0.3.5 lazyeval_0.2.0  yaml_2.1.14     assertthat_0.2.0
## [17] rprojroot_1.2   digest_0.6.12   tibble_1.3.3    bindrcpp_0.2
## [21] glue_1.1.1      evaluate_0.10   rmarkdown_1.4   labeling_0.3
## [25] stringi_1.1.5   scales_0.4.1    backports_1.0.5 pkgconfig_2.0.1
```