main

August 12, 2021

1 Boxplots for permutations analysis (male downsampling to female level)

[1]:
```
library(tidyverse)
library(ggpubr)
```

```
  Attaching packages                                    tidyverse
1.3.1

 ggplot2 3.3.5      purrr   0.3.4
 tibble  3.1.2      dplyr   1.0.7
 tidyr   1.1.3      stringr 1.4.0
 readr   1.4.0      forcats 0.5.1

  Conflicts
tidyverse_conflicts()
 dplyr::filter() masks stats::filter()
 dplyr::lag()    masks stats::lag()
```

1.1 Function and configuration

[2]:
```
save_ggplots <- function(p, fn, w=6, h=6){
    for(ext in c('.svg', '.png', '.pdf')){
        ggsave(p, filename=paste0(fn, ext), width=w, height=h)
    }
}
```

[3]:
```
config = list(
    "Caudate"="../../../caudate/subsampling_male/deg_summary/_m/permutations.
 ↪csv",
    "DLPFC"="../../../dlpfc/subsampling_male/deg_summary/_m/permutations.csv",
    "Hippocampus"="../../../hippocampus/subsampling_male/deg_summary/_m/
 ↪permutations.csv"
)
```

1.2 Merge dataframes

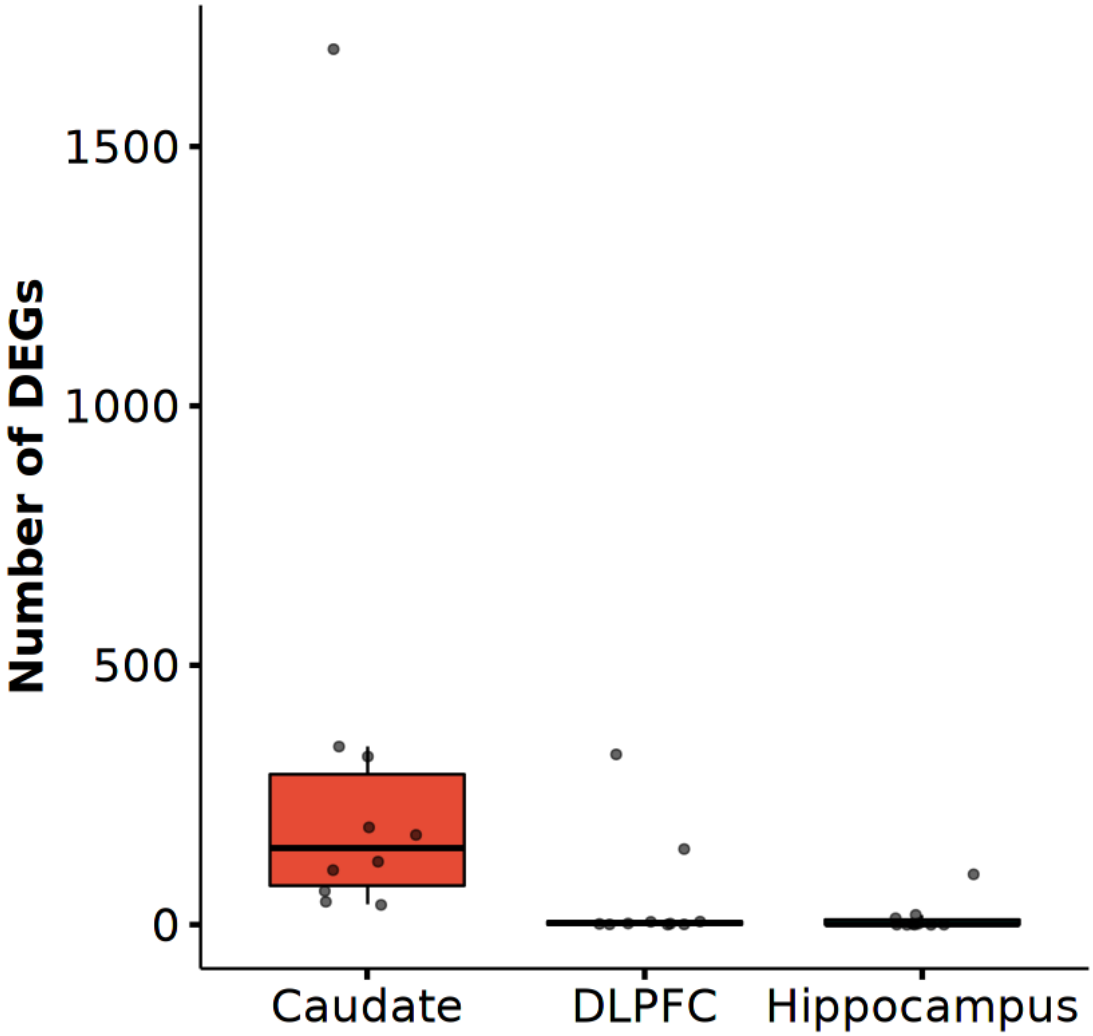
[4]:
```
datalist = list()
for(tissue in c("Caudate", "DLPFC", "Hippocampus")){
    df = data.table::fread(config[[tissue]])
    df$tissue <- tissue  # maybe you want to keep track of which iteration␣
 ↪produced it?
    datalist[[tissue]] <- df
}
big_df <- bind_rows(datalist) %>%
    select(Symbol, gencodeID, logFC, t, adj.P.Val, Permutation, tissue)
big_df %>% head()
```

A data.table: 6 × 7

Symbol <chr>	gencodeID <chr>	logFC <dbl>	t <dbl>	adj.P.Val <dbl>	Permutation <int>
GDNF-AS1	ENSG00000248587.7	1.0159219	8.553727	1.566672e-09	2
GDNF	ENSG00000168621.14	0.8188297	6.045744	1.541499e-04	2
NSUN5P2	ENSG00000106133.17	-0.3648580	-6.031454	1.541499e-04	2
KCNH5	ENSG00000140015.19	0.5558007	5.980020	1.541499e-04	2
PAPPA	ENSG00000182752.9	0.7535862	5.862024	2.130430e-04	2
RMDN2	ENSG00000115841.19	-0.3193313	-5.796196	2.403024e-04	2

[5]:
```
bxp = big_df %>% group_by(Permutation, tissue) %>% summarize(Size = n()) %>%
    as.data.frame %>% pivot_wider(names_from = tissue, values_from = Size) %>%
    replace_na(list(DLPFC = 0, Hippocampus = 0)) %>%
    pivot_longer(-Permutation, names_to="Tissue", values_to="DEGs") %>%
    mutate_if(is.character, as.factor) %>%
    ggboxplot(x="Tissue", y="DEGs", fill="Tissue", add="jitter",
              xlab='', palette="npg", ylab="Number of DEGs",
              add.params=list(alpha=0.6), outlier.shape=NA,
              panel.labs.font=list(face='bold'), legend="",
              ggtheme=theme_pubr(base_size=20))+
        font("xy.title", face="bold")
save_ggplots(bxp, "permutation_boxplots_3regions", 6, 5)
bxp
```

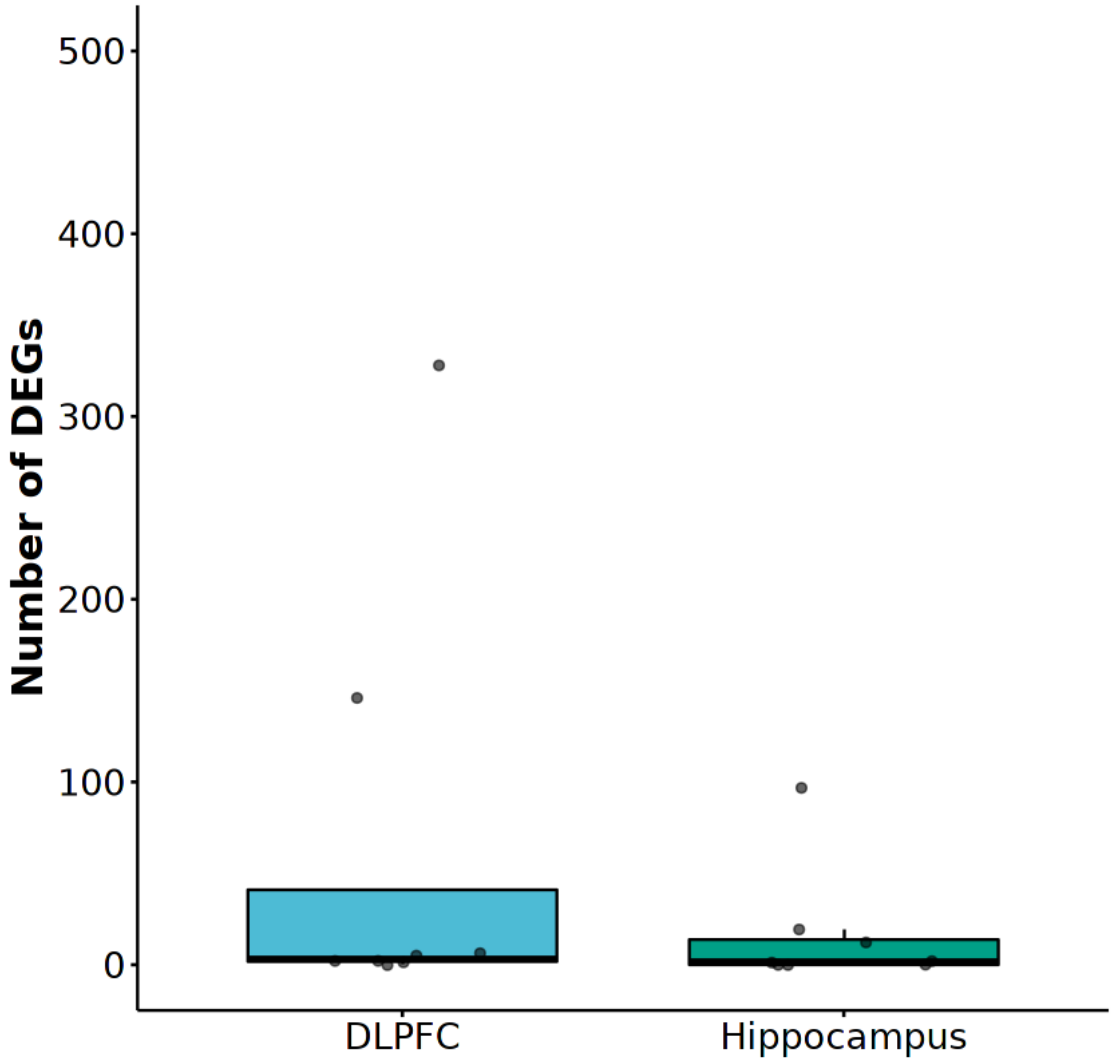
```
`summarise()` has grouped output by 'Permutation'. You can override using the
`.groups` argument.
```



[6]:
```
bxp = big_df %>% filter(tissue != "Caudate") %>%
    group_by(Permutation, tissue) %>% summarize(Size = n()) %>%
    as.data.frame %>% pivot_wider(names_from = tissue, values_from = Size) %>%
    replace_na(list(DLPFC = 0, Hippocampus = 0)) %>%
    pivot_longer(-Permutation, names_to="Tissue", values_to="DEGs") %>%
    mutate_if(is.character, as.factor) %>%
    ggboxplot(x="Tissue", y="DEGs", fill="Tissue", add="jitter",
              xlab='', palette=get_palette("npg", 3)[2:3], ylab="Number of␣
 ↪DEGs",
              add.params=list(alpha=0.6), outlier.shape=NA,
              panel.labs.font=list(face='bold', size = 18),
              legend="", ylim=c(0, 500))+
        font("xy.title", size=18, face="bold") +
```

```
        font("xy.text", size=16)
save_ggplots(bxp, "permutation_boxplots_2regions", 6, 5)
bxp
```

```
`summarise()` has grouped output by 'Permutation'. You can override using the
`.groups` argument.
```



[7]:
```
big_df %>% group_by(Permutation, tissue) %>% summarize(Size = n()) %>%
    as.data.frame %>% pivot_wider(names_from = tissue, values_from = Size) %>%
    replace_na(list(DLPFC = 0, Hippocampus = 0)) %>%
    pivot_longer(-Permutation, names_to="Tissue", values_to="DEGs") %>%
    mutate_if(is.character, as.factor) %>%
```

```
    group_by(Tissue) %>% summarize(Mean = mean(DEGs), Median = median(DEGs),␣
↪Std = sd(DEGs))
```

```
`summarise()` has grouped output by 'Permutation'. You can override using the
`.groups` argument.
```

A tibble: 3 × 4

Tissue <fct>	Mean <dbl>	Median <dbl>	Std <dbl>
Caudate	308.8	147.0	496.27094
DLPFC	49.0	2.0	107.99177
Hippocampus	13.1	0.5	30.18995

1.3 Reproducibility Information

[8]:
```
Sys.time()
proc.time()
options(width = 120)
sessioninfo::session_info()
```

```
[1] "2021-08-12 19:05:34 EDT"
```

```
   user  system elapsed
  4.731   0.222   5.261
```

```
 Session info
 setting  value
 version  R version 4.0.3 (2020-10-10)
 os       Arch Linux
 system   x86_64, linux-gnu
 ui       X11
 language (EN)
 collate  en_US.UTF-8
 ctype    en_US.UTF-8
 tz       America/New_York
 date     2021-08-12

 Packages
 package     * version  date       lib source
 abind         1.4-5    2016-07-21 [1] CRAN (R 4.0.2)
 assertthat    0.2.1    2019-03-21 [1] CRAN (R 4.0.2)
 backports     1.2.1    2020-12-09 [1] CRAN (R 4.0.2)
 base64enc     0.1-3    2015-07-28 [1] CRAN (R 4.0.2)
 broom         0.7.8    2021-06-24 [1] CRAN (R 4.0.3)
 Cairo         1.5-12.2 2020-07-07 [1] CRAN (R 4.0.2)
 car           3.0-11   2021-06-27 [1] CRAN (R 4.0.3)
 carData       3.0-4    2020-05-22 [1] CRAN (R 4.0.2)
 cellranger    1.1.0    2016-07-27 [1] CRAN (R 4.0.2)
 cli           3.0.0    2021-06-30 [1] CRAN (R 4.0.3)
```

```
colorspace    2.0-2    2021-06-24 [1] CRAN (R 4.0.3)
crayon        1.4.1    2021-02-08 [1] CRAN (R 4.0.3)
curl          4.3.2    2021-06-23 [1] CRAN (R 4.0.3)
data.table    1.14.0   2021-02-21 [1] CRAN (R 4.0.3)
DBI           1.1.1    2021-01-15 [1] CRAN (R 4.0.2)
dbplyr        2.1.1    2021-04-06 [1] CRAN (R 4.0.3)
digest        0.6.27   2020-10-24 [1] CRAN (R 4.0.2)
dplyr       * 1.0.7    2021-06-18 [1] CRAN (R 4.0.3)
ellipsis      0.3.2    2021-04-29 [1] CRAN (R 4.0.3)
evaluate      0.14     2019-05-28 [1] CRAN (R 4.0.2)
fansi         0.5.0    2021-05-25 [1] CRAN (R 4.0.3)
farver        2.1.0    2021-02-28 [1] CRAN (R 4.0.3)
forcats     * 0.5.1    2021-01-27 [1] CRAN (R 4.0.2)
foreign       0.8-80   2020-05-24 [2] CRAN (R 4.0.3)
fs            1.5.0    2020-07-31 [1] CRAN (R 4.0.2)
generics      0.1.0    2020-10-31 [1] CRAN (R 4.0.2)
ggplot2     * 3.3.5    2021-06-25 [1] CRAN (R 4.0.3)
ggpubr      * 0.4.0    2020-06-27 [1] CRAN (R 4.0.2)
ggsci         2.9      2018-05-14 [1] CRAN (R 4.0.2)
ggsignif      0.6.2    2021-06-14 [1] CRAN (R 4.0.3)
glue          1.4.2    2020-08-27 [1] CRAN (R 4.0.2)
gtable        0.3.0    2019-03-25 [1] CRAN (R 4.0.2)
haven         2.4.1    2021-04-23 [1] CRAN (R 4.0.3)
hms           1.1.0    2021-05-17 [1] CRAN (R 4.0.3)
htmltools     0.5.1.1  2021-01-22 [1] CRAN (R 4.0.2)
httr          1.4.2    2020-07-20 [1] CRAN (R 4.0.2)
IRdisplay     1.0      2021-01-20 [1] CRAN (R 4.0.2)
IRkernel      1.2      2021-05-11 [1] CRAN (R 4.0.3)
jsonlite      1.7.2    2020-12-09 [1] CRAN (R 4.0.2)
labeling      0.4.2    2020-10-20 [1] CRAN (R 4.0.2)
lifecycle     1.0.0    2021-02-15 [1] CRAN (R 4.0.3)
lubridate     1.7.10   2021-02-26 [1] CRAN (R 4.0.3)
magrittr      2.0.1    2020-11-17 [1] CRAN (R 4.0.2)
modelr        0.1.8    2020-05-19 [1] CRAN (R 4.0.2)
munsell       0.5.0    2018-06-12 [1] CRAN (R 4.0.2)
openxlsx      4.2.4    2021-06-16 [1] CRAN (R 4.0.3)
pbdZMQ        0.3-5    2021-02-10 [1] CRAN (R 4.0.3)
pillar        1.6.1    2021-05-16 [1] CRAN (R 4.0.3)
pkgconfig     2.0.3    2019-09-22 [1] CRAN (R 4.0.2)
purrr       * 0.3.4    2020-04-17 [1] CRAN (R 4.0.2)
R6            2.5.0    2020-10-28 [1] CRAN (R 4.0.2)
Rcpp          1.0.7    2021-07-07 [1] CRAN (R 4.0.3)
readr       * 1.4.0    2020-10-05 [1] CRAN (R 4.0.2)
readxl        1.3.1    2019-03-13 [1] CRAN (R 4.0.2)
repr          1.1.3    2021-01-21 [1] CRAN (R 4.0.2)
reprex        2.0.0    2021-04-02 [1] CRAN (R 4.0.3)
rio           0.5.27   2021-06-21 [1] CRAN (R 4.0.3)
rlang         0.4.11   2021-04-30 [1] CRAN (R 4.0.3)
```

```
 rstatix       0.7.0   2021-02-13 [1] CRAN (R 4.0.3)
 rstudioapi    0.13    2020-11-12 [1] CRAN (R 4.0.2)
 rvest         1.0.0   2021-03-09 [1] CRAN (R 4.0.3)
 scales        1.1.1   2020-05-11 [1] CRAN (R 4.0.2)
 sessioninfo   1.1.1   2018-11-05 [1] CRAN (R 4.0.2)
 stringi       1.7.3   2021-07-16 [1] CRAN (R 4.0.3)
 stringr     * 1.4.0   2019-02-10 [1] CRAN (R 4.0.2)
 svglite       2.0.0   2021-02-20 [1] CRAN (R 4.0.3)
 systemfonts   1.0.2   2021-05-11 [1] CRAN (R 4.0.3)
 tibble      * 3.1.2   2021-05-16 [1] CRAN (R 4.0.3)
 tidyr       * 1.1.3   2021-03-03 [1] CRAN (R 4.0.3)
 tidyselect    1.1.1   2021-04-30 [1] CRAN (R 4.0.3)
 tidyverse   * 1.3.1   2021-04-15 [1] CRAN (R 4.0.3)
 utf8          1.2.1   2021-03-12 [1] CRAN (R 4.0.3)
 uuid          0.1-4   2020-02-26 [1] CRAN (R 4.0.2)
 vctrs         0.3.8   2021-04-29 [1] CRAN (R 4.0.3)
 withr         2.4.2   2021-04-18 [1] CRAN (R 4.0.3)
 xml2          1.3.2   2020-04-23 [1] CRAN (R 4.0.2)
 zip           2.2.0   2021-05-31 [1] CRAN (R 4.0.3)

[1] /home/jbenja13/R/x86_64-pc-linux-gnu-library/4.0
[2] /usr/lib/R/library
```