

# main

September 27, 2021

## 1 Boxplot of region specific ancestry-related DEGs

```
[1]: suppressPackageStartupMessages({  
  library(tidyverse)  
  library(ggpubr)  
})
```

### 1.1 Functions

#### 1.1.1 Cached functions

```
[2]: get_biomart_df <- function(){  
  biomart = data.table::fread("../_h/biomart.csv")  
}  
memMART <- memoise::memoise(get_biomart_df)  
  
get_degs <- function(){  
  mashr_file = "../_m/mashr/_m/genes/lfsr_feature_4tissues.txt.gz"  
  return(data.table::fread(mashr_file))  
}  
memDEG <- memoise::memoise(get_degs)  
  
get_select_degs <- function(){  
  geneids = c("NKAIN1", "LRRK1", "WIF1", "GRIN3A", "PDE1A", "EMX2OS", "PSPH",  
    ↪ "PLRG1")  
  gene_df = memDEG() %>% pivot_longer(-Effect, names_to="Tissue",  
    ↪ values_to="lfsr") %>%  
    mutate(Tissue=gsub("\\.", " ", Tissue), ensembl_gene_id=gsub("\\..*",  
    ↪ "", Effect)) %>%  
    inner_join(memMART(), by="ensembl_gene_id") %>%  
    filter(external_gene_name %in% geneids) %>% select(-V1, -description,  
    ↪ -entrezgene) %>%  
    mutate(New_ID=paste(external_gene_name, paste("lfsr:", signif(lfsr,2)),  
    ↪ sep="\n")) %>%  
    mutate_if(is.character, as.factor)  
  return(gene_df)  
}
```

```

memGENES <- memoise::memoise(get_select_degs)

get_residualized_df <- function(){
  expr_file = "../_m/genes/residualized_expression.tsv"
  return(data.table::fread(expr_file) %>%
    filter(Geneid %in% memGENES()$Effect) %>%
    column_to_rownames("Geneid"))
}
memRES <- memoise::memoise(get_residualized_df)

get_pheno_df <- function(){
  phenotype_file = paste0('.../input/phenotypes/merged/_m/
↳merged_phenotypes.csv')
  return(data.table::fread(phenotype_file) %>% select(RNum, Race, Region))
}
memPHENO <- memoise::memoise(get_pheno_df)

merge_dt <- function(){
  bigdf = memRES() %>% t %>% as.data.frame %>% rownames_to_column("RNum") %>%
    inner_join(memPHENO(), by="RNum") %>%
    mutate(Race=gsub("CAUC", "EA", Race),
      Region=gsub("HIPPO", "Hippocampus", Region),
      Region=gsub("DentateGyrus", "Dentate Gyrus", Region)) %>%
    pivot_longer(~c(RNum, Race, Region), names_to="Geneid",
↳values_to="Res") %>%
    inner_join(memGENES(), by=c("Geneid"="Effect"))
  return(bigdf)
}
memDF <- memoise::memoise(merge_dt)

```

### 1.1.2 Simple functions

```

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

```

## 1.2 Plot DEGs

### 1.2.1 Group plot

```

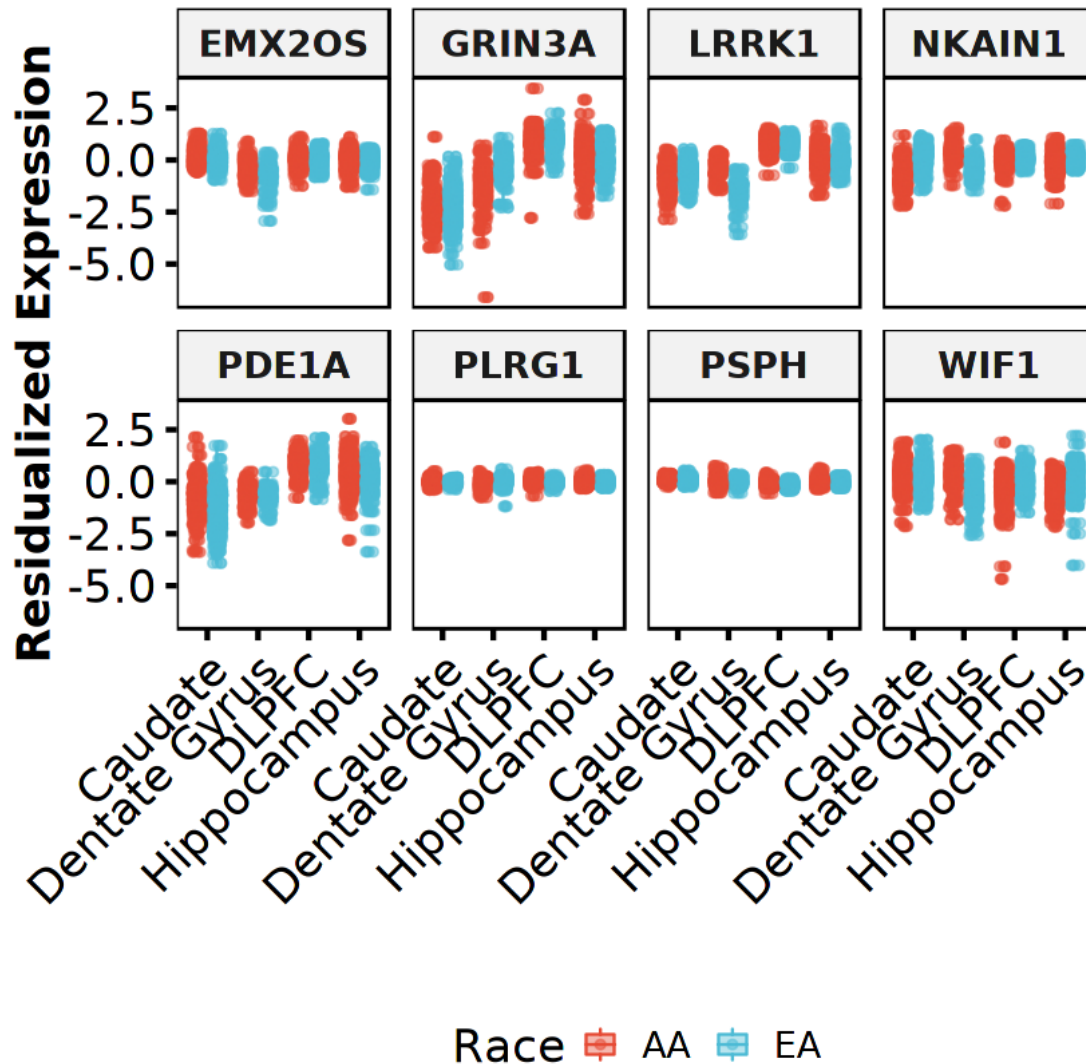
[4]: tissues = c("Caudate", "Dentate Gyrus", "DLPFC", "Hippocampus")
byp = memDF() %>% mutate(Region = factor(Region, levels=c(tissues))) %>%
  ggboxplot(x="Region", y="Res", fill="Race", color="Race", add="jitter",
    palette="npg", facet.by="external_gene_name", ncol=4, xlab="",

```

```

ylab="Residualized Expression", panel.labs.font=list(face='bold'),
outlier.shape=NA, add.params=list(alpha=0.5), alpha=0.4,
legend="bottom", ggtheme=theme_pubr(base_size=20, border=TRUE)) +
font("xy.title", face="bold") + rotate_x_text(45)
save_ggplots("boxplot_interaction_ancestry_degs", bxp, 10, 8)
bxp

```



### 1.2.2 Individual genes

```

[5]: geneids = c("NKAIN1", "LRRK1", "WIF1", "GRIN3A", "PDE1A", "EMX2OS", "PSPH",
  ↪ "PLRG1")
for(geneid in geneids){

```

```

dt = memDF() %>% filter(external_gene_name == geneid)
y1 = min(dt$Res) - 0.25; y2 = max(dt$Res) + 0.25
bxp = dt %>% mutate(Region = factor(Region, levels=c(tissues))) %>%
  ggboxplot(x="Region", y="Res", fill="Race", color="Race", add="jitter",
    palette="npg", facet.by="external_gene_name", ncol=4, xlab="",
    ylab="Residualized Expression", panel.labs.
    ↪font=list(face='bold'),
    outlier.shape=NA, add.params=list(alpha=0.5), alpha=0.4,
    ylim=c(y1, y2), ggtheme=theme_pubr(base_size=20, border=TRUE),
    legend="bottom") +
    font("xy.title", face="bold") + rotate_x_text(45)
save_ggplots(paste0("boxplot_interaction_ancestry_",geneid), bxp, 4, 6)
}

```

### 1.3 Session Info

```

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

```

```
[1] "2021-09-27 14:08:26 EDT"
```

```

  user  system elapsed
32.873   4.294   26.974

```

```
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     2021-09-27
```

```
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	0.7.9	2021-07-27	[1]	CRAN (R 4.0.3)
cachem	1.0.6	2021-08-19	[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	3.0.1	2021-07-17	[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)
fastmap	1.1.0	2021-01-25	[1]	CRAN	(R 4.0.2)
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	2.4.3	2021-08-04	[1]	CRAN	(R 4.0.3)
hms	1.1.0	2021-05-17	[1]	CRAN	(R 4.0.3)
htmltools	0.5.2	2021-08-25	[1]	CRAN	(R 4.0.3)
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)
memoise	2.0.0	2021-01-26	[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	1.6.2	2021-07-29	[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)
R.methodsS3	1.8.1	2020-08-26	[1]	CRAN	(R 4.0.3)
R.oo	1.24.0	2020-08-26	[1]	CRAN	(R 4.0.3)
R.utils	2.10.1	2020-08-26	[1]	CRAN	(R 4.0.3)
R6	2.5.1	2021-08-19	[1]	CRAN	(R 4.0.3)

Rcpp	1.0.7	2021-07-07	[1]	CRAN	(R 4.0.3)
readr	* 2.0.1	2021-08-10	[1]	CRAN	(R 4.0.3)
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	2.0.1	2021-08-05	[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	1.0.1	2021-07-26	[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	1.7.4	2021-08-25	[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	* 3.1.4	2021-08-25	[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)
tzdb	0.1.2	2021-07-20	[1]	CRAN	(R 4.0.3)
utf8	1.2.2	2021-07-24	[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