

main

August 25, 2021

1 Cell deconvolution comparison and quality control

Performed by Louise Huuki

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

```
Attaching packages
1.3.1

tidyverse

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 Functions

```
[2]: get_pheno <- function(){
      df = data.table::fread("../.../input/phenotypes/merged/_m/
      ↪merged_phenotypes.csv") %>%
        filter(Dx %in% c("Control"), Age > 17, Race %in% c("AA", "CAUC")) %>%
        mutate(Race = gsub("CAUC", "EA", Race)) %>% select(-V1)
      return(df)
    }

memPHENO <- memoise::memoise(get_pheno)

save_img <- function(image, fn, w, h){
  for(ext in c(".svg", ".pdf", ".png")){
    ggsave(file=paste0(fn, ext), plot=image, width=w, height=h)
  }
}
```

1.2 Prepare data

```
[3]: load("../_m/est_prop_Bisque.v2.Rdata", verbose = TRUE)
names(est_prop_bisque)
```

Loading objects:

est_prop_bisque

1. 'dg' 2. 'caudate' 3. 'hippo' 4. 'dlpfc'

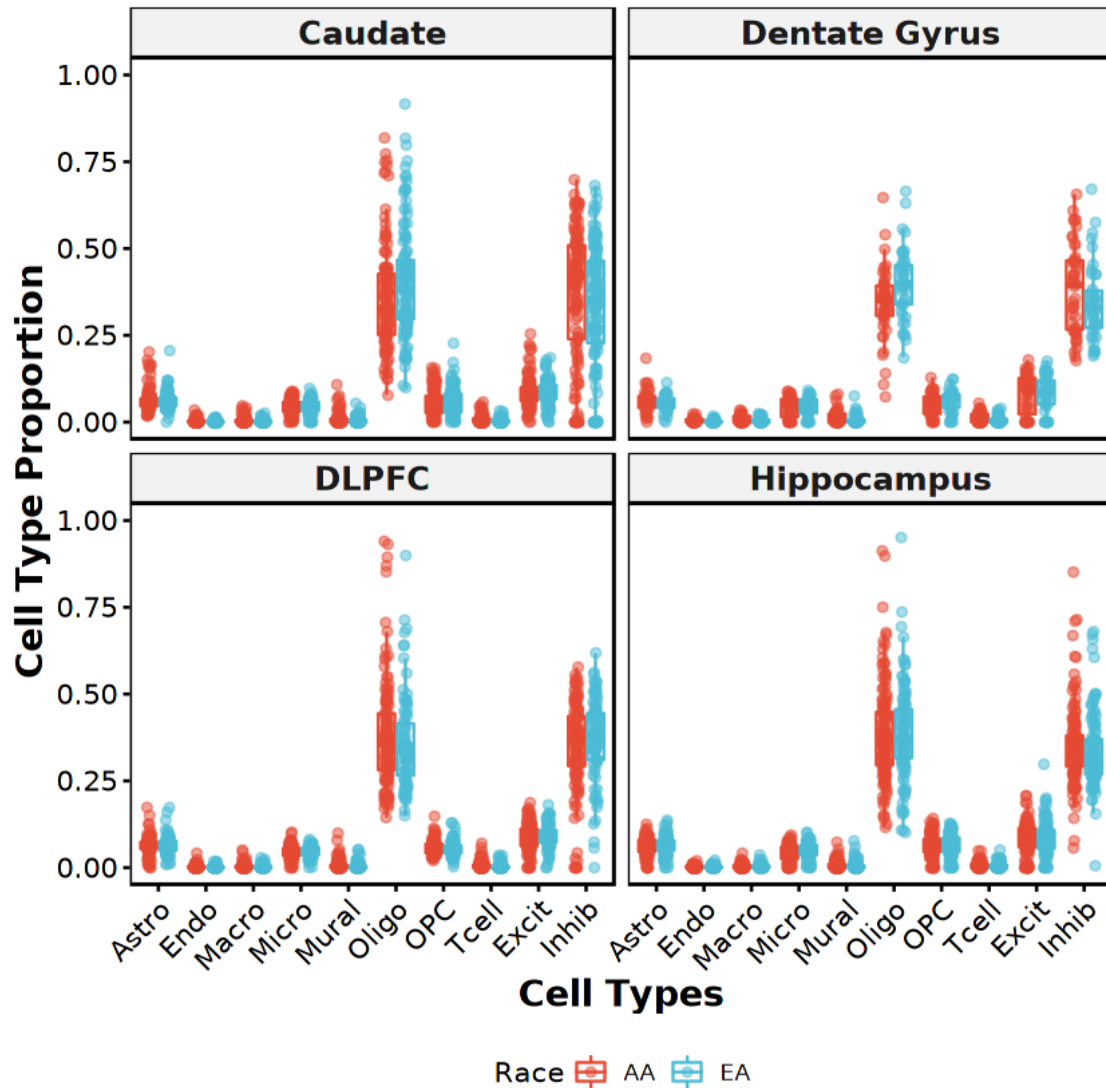
```
[4]: cc = est_prop_bisque$caudate$Est.prop.long %>%
      inner_join(memPHENO(), by=c("sample"="RNum")) %>%
      mutate_if(is.character, as.factor) %>%
      rename("Proportion"="prop") %>% mutate(Tissue="Caudate")
dd = est_prop_bisque$dlpfc$Est.prop.long %>%
      inner_join(memPHENO(), by=c("sample"="RNum")) %>%
      mutate_if(is.character, as.factor) %>%
      rename("Proportion"="prop") %>% mutate(Tissue="DLPFC")
hh = est_prop_bisque$hippo$Est.prop.long %>%
      inner_join(memPHENO(), by=c("sample"="RNum")) %>%
      mutate_if(is.character, as.factor) %>%
      rename("Proportion"="prop") %>% mutate(Tissue="Hippocampus")
gg = est_prop_bisque$dg$Est.prop.long %>%
      separate(sample, c("sample", "batch")) %>%
      inner_join(memPHENO(), by=c("sample"="RNum")) %>%
      mutate_if(is.character, as.factor) %>%
      rename("Proportion"="prop") %>% mutate(Tissue="Dentate Gyrus")

df = bind_rows(cc, dd, hh, gg)
df %>% head(2)
```

A data.frame: 2 × 15	sample	cell_type	Proportion	BrNum	Region	RIN	Age	Sex	Race
	<fct>	<fct>	<dbl>	<fct>	<fct>	<dbl>	<dbl>	<fct>	<fct>
1	R12872	Astro	0.05615308	Br949	Caudate	9.2	18.05	M	AA
2	R12873	Astro	0.05615373	Br991	Caudate	9.4	35.24	M	AA

1.3 Cell type proportion plots

```
[5]: bxp = df %>% ggboxplot(x="cell_type", y="Proportion", color="Race", facet.
      ~by="Tissue",
      panel.labs.font=list(face='bold', size = 14), palette="npg",
      outlier.shape=NA, ylab='Cell Type Proportion', add='jitter',
      add.params=list(alpha=0.5), ylim=c(0, 1), xlab="Cell Types",
      legend="bottom") +
      font("xy.text", size=12) + font("xy.title", size=16, face="bold") +
      rotate_x_text(45)
print(bxp)
save_img(bxp, paste0("boxplot_celltypes_ancestry"), w=7, h=6)
```



```
[6]: for(tissue in c("Caudate", "DLPFC", "Hippocampus", "Dentate Gyrus")){
  bxp = df %>% filter(Tissue == tissue) %>%
    ggboxplot(x="Race", y="Proportion", color="Race", facet.by="cell_type",
              panel.labs.font=list(face='bold', size = 14), palette="npg",
              outlier.shape=NA, ylab='Cell Type Proportion', add='jitter',
              add.params=list(alpha=0.5), ylim=c(0, 1.2), ncol=3,
              xlab="Ancestry", legend="None") +
    stat_compare_means(comparisons=list(c("AA", "EA")), aes(label=..p.
    ↪signif..),
                      method="t.test") +
    font("xy.text", size=12) + font("xy.title", size=16, face="bold")
  print(bxp)
  fn = paste0("boxplot_ancestryBYcelltype_",gsub(" ", "_", tolower(tissue)))
```

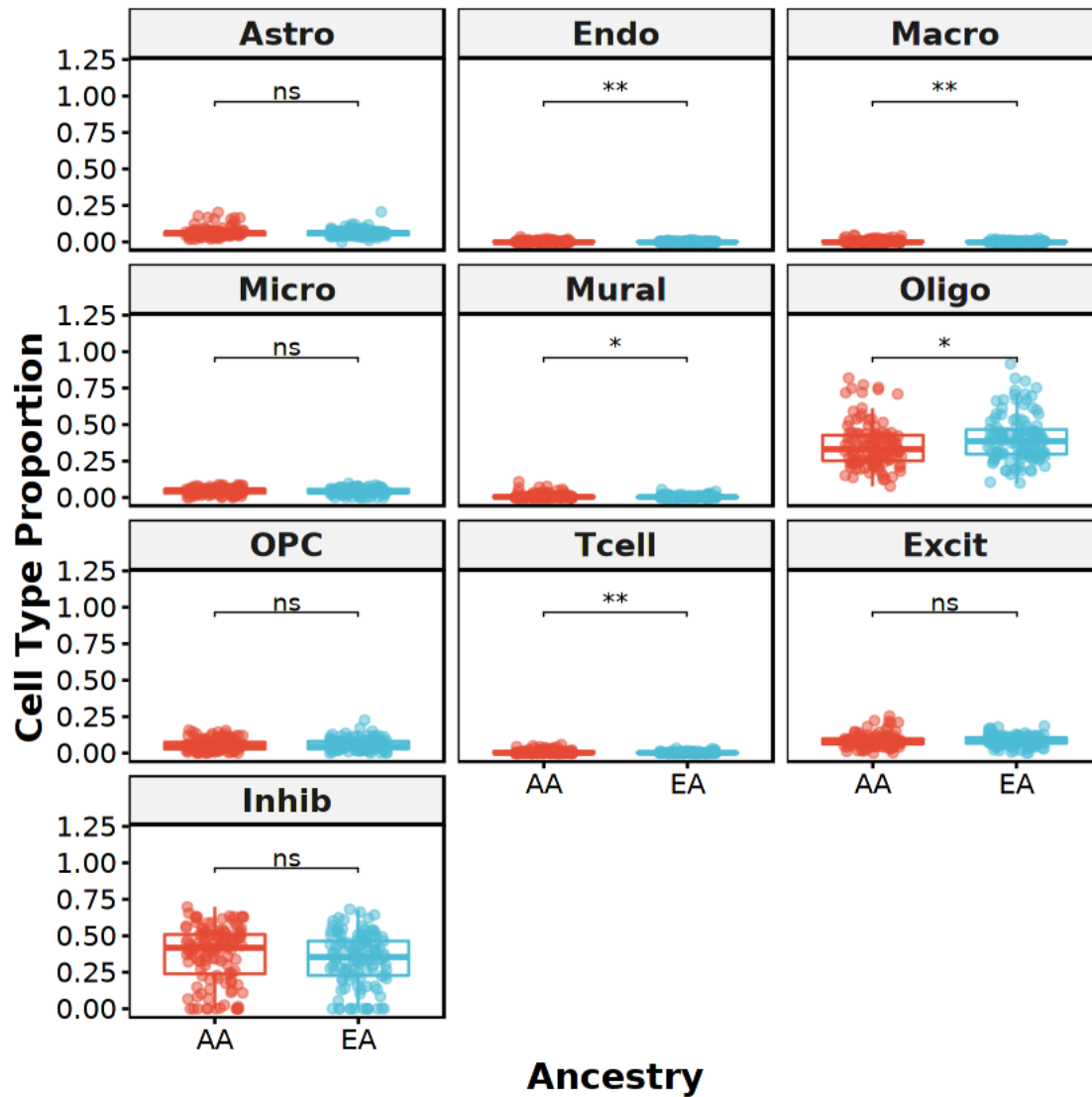
```
save_img(bxp, fn, w=7, h=6)
}
```

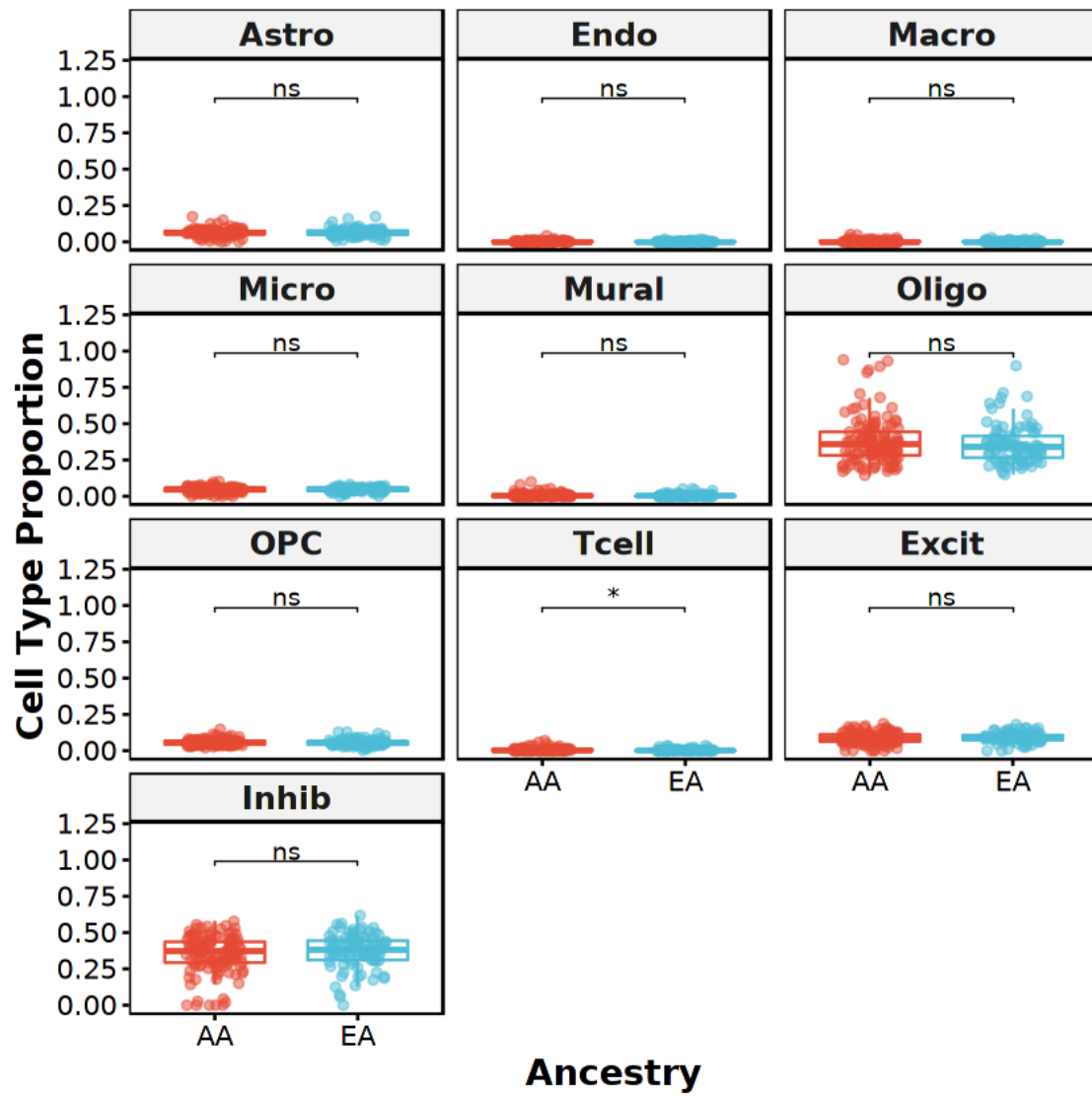
Warning message:

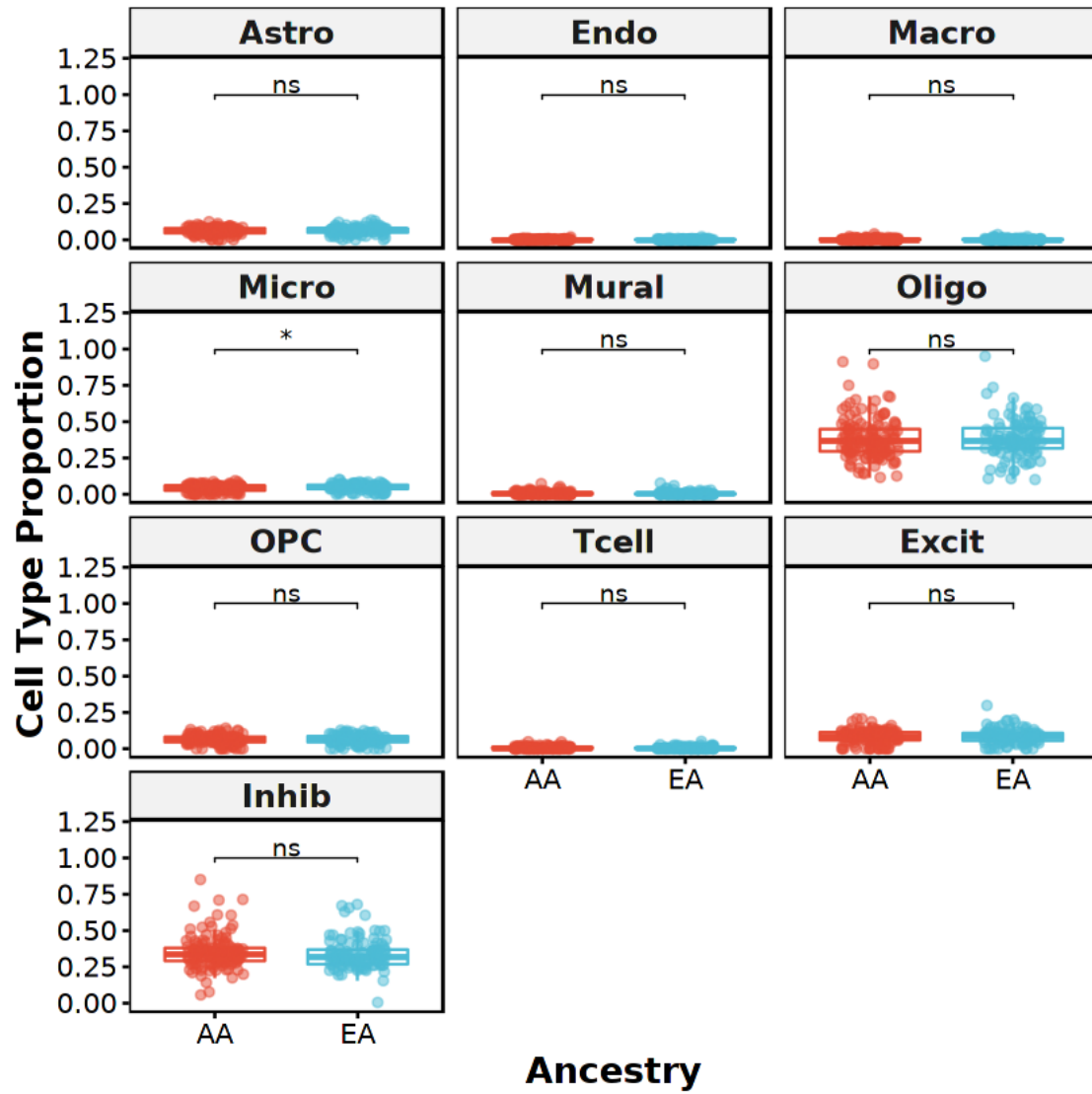
"Using `as.character()` on a quosure is deprecated as of rlang 0.3.0.

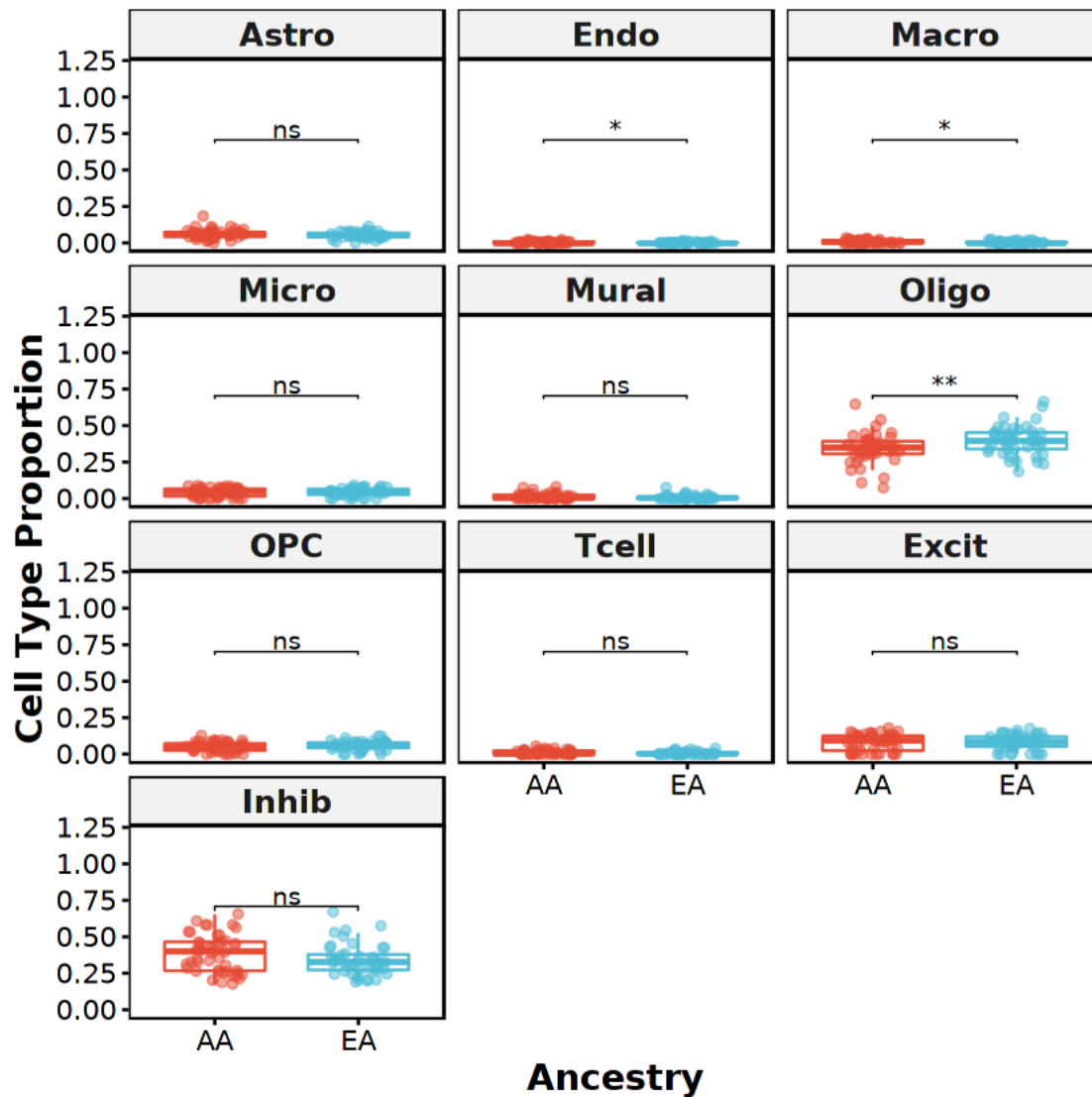
Please use `as_label()` or `as_name()` instead.

This warning is displayed once per session."









1.4 Reproducibility Information

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

```
[1] "2021-08-25 14:56:34 EDT"
```

```
user system elapsed
28.506  0.677  28.290
```

```
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    2021-08-25

```

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)
cachem	1.0.5	2021-05-15	[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)
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	2.4.1	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)
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	1.6.1	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