main

August 1, 2021

1 Check sex chromosomes for caudate nucleus

```
[1]: suppressMessages({library(dplyr)
                        library(ggpubr)})
```

1.1 Get and annotated sex from genotypes

```
[2]: sexcheck_file = paste0('/ceph/projects/brainseq/genotype/download/topmed/',
                            'convert2plink/filter_maf_01/check_sex/_m/
      ↪LIBD_Brain_TopMed.sexcheck')
     pheno_file = '/ceph/users/jbenja13/projects/sex_sz_ria/input/phenotypes/_m/
      ↪phenotypes.csv'

     dt = data.table::fread(sexcheck_file) %>%
         inner_join(data.table::fread(pheno_file), by=c("FID"="BrNum")) %>%
         select(c("FID", "Dx", "Sex", "Age", "F", "SNPSEX")) %>%
         distinct() %>% mutate_if(is.character, as.factor)
     levels(dt$Sex) <- c('Female', 'Male')
     dt %>% head(2)
```
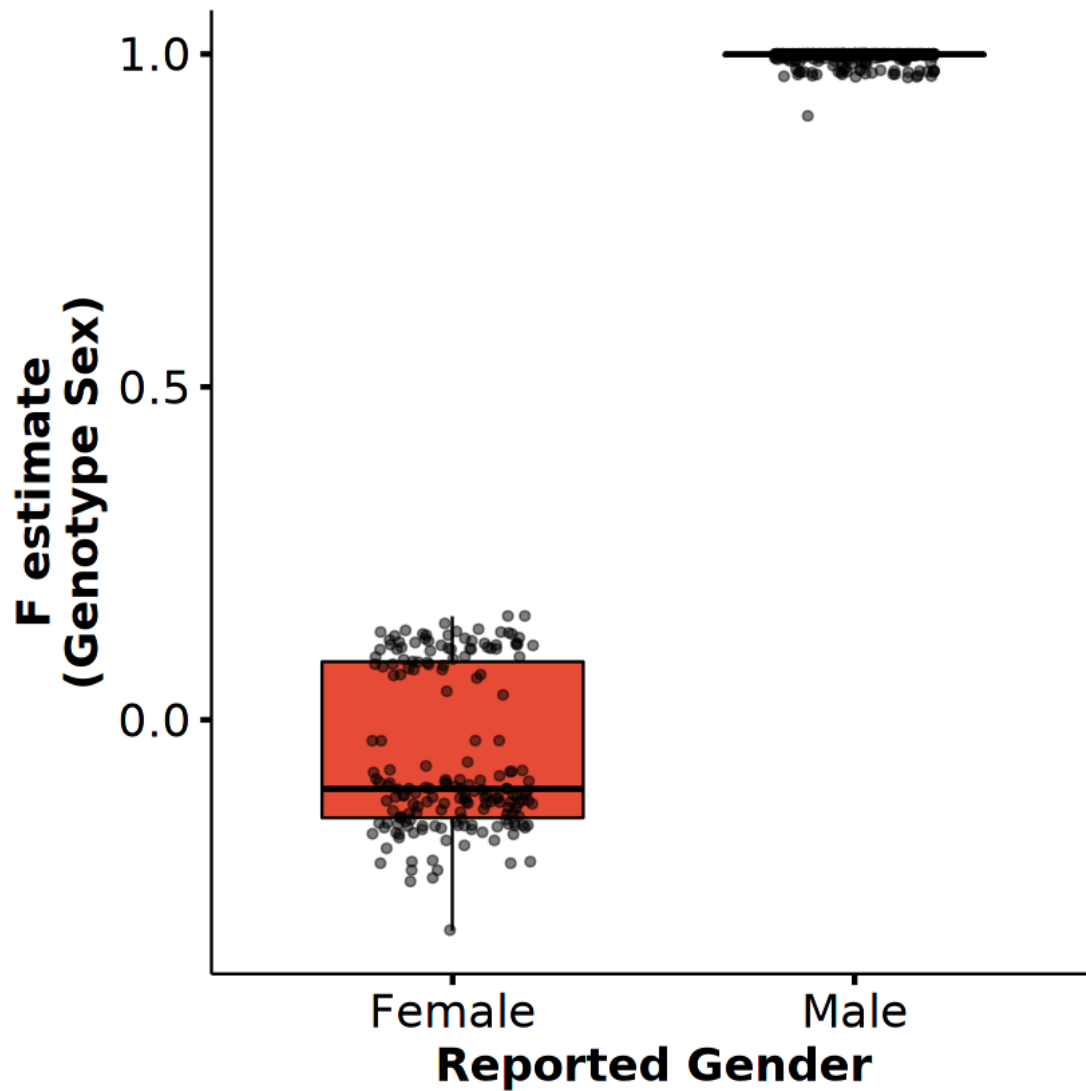
A data.table: 2 × 6

FID	Dx	Sex	Age	F	SNPSEX
<fct>	<fct>	<fct>	<dbl>	<dbl>	<int>
Br763	Schizo	Male	75.22	1	1
Br822	Control	Male	26.73	1	1

1.2 Plot F estimate

```
[3]: fig <- ggboxplot(dt, x = "Sex", y = "F", fill='Sex', width=0.65, palette='npg',
                      xlab='Reported Gender', ylab='F estimate\n(Genotype Sex)',
                      ggtheme=theme_pubr(base_size=20), add="jitter",
                      add.params=list(alpha=0.5), legend="None") +
         font("xy.title", face='bold')
     fig
```

```
[4]: ggsave("sex_check.png")
     ggsave("sex_check.pdf")
     ggsave("sex_check.svg")
```

Saving 6.67 x 6.67 in image

Saving 6.67 x 6.67 in image

Saving 6.67 x 6.67 in image

```
[5]: dt %>%
        group_by(Sex) %>%
```

```
    summarize(mean_size = mean(F, na.rm = TRUE), max = max(F, na.rm=TRUE))
```

A tibble: 2 × 3

Sex	mean_size	max
<fct>	<dbl>	<dbl>
Female	-0.05355674	0.156
Male	0.99623137	1.000

1.3 Check that reported gender matches genotype sex

```
[6]: sum(dt$Sex == "Male")
     sum(dt$SNPSEX == 1 & dt$Sex == "Male")
```

408

408

```
[7]: sum(dt$Sex == "Female")
     sum(dt$SNPSEX == 2 & dt$Sex == "Female")
```

187

187

1.3.1 Examine number of X chromosomes

```
[8]: table(dt$SNPSEX)
```

```
  1   2
408 187
```

```
[9]: table(dt$Sex)
```

```
Female   Male
   187    408
```

Summary: All samples have reported gender matching genotype sex

1.4 Reproducibility Information

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

```
[1] "2021-08-01 15:56:07 EDT"
```

```
   user  system elapsed
  3.079   0.137   3.472
```

```
 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     2021-08-01

 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)
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)
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)
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)
magrittr      2.0.1     2020-11-17 [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)
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)
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)
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)
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)
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)
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
```