main

July 14, 2021

1 Prepare covariates from Matrix eQTL analysis

Kynon J Benjamin

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

```
 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 Load data

```
[2]: sample_df = data.table::fread("../../_m/sample_id_to_brnum.tsv")
     sample_df %>% head()
```

A data.table: 6 × 2

| RNum | BrNum |
<chr>	<chr>
R5527	Br822
R2725	Br823
R4805	Br836
R5528	Br845
R4817	Br846
R4793	Br848

```
[3]: load("/ceph/users/jbenja13/projects/sex_sz_ria/input/genotypes/covariates/_m/
     ↪hippocampus/covariates.rda")
     ls()
```

1. 'covsExon0' 2. 'covsGene0' 3. 'covsJxn0' 4. 'covsTx0' 5. 'sample_df'

```
[4]: covs = covsJxn0 %>% as.data.frame %>% select(sample_df$RNum) %>%
         t %>% as.data.frame %>% rownames_to_column("RNum") %>%
         inner_join(sample_df, by=c("RNum")) %>% select(-"RNum") %>%
         rename("ID"="BrNum") %>% column_to_rownames("ID") %>% t %>%
         as.data.frame %>% rownames_to_column("ID")

     covs[1:10, 1:5]
```

A data.frame: 10 × 5

	ID <chr>	Br822 <dbl>	Br823 <dbl>	Br836 <dbl>	Br845 <dbl>
1	DxSchizo	0.00000000	0.000000000	0.000000e+00	0.000000000
2	snpPC1	-0.03381080	0.087038000	1.014550e-01	0.093891100
3	snpPC2	-0.00312935	-0.001717740	-6.424960e-04	-0.002248600
4	snpPC3	-0.00154150	0.000413453	-9.232110e-04	0.000546358
5	PC1	-176.83664527	-54.469788615	8.709362e+00	17.104562532
6	PC2	18.79661049	-36.098134506	-1.333937e+02	101.956337819
7	PC3	-118.63179351	302.345385549	5.101843e+00	-33.346098172
8	PC4	-84.53598032	61.019584270	-5.190710e+01	88.616636507
9	PC5	-38.97815281	1.904365350	2.877393e+01	0.728018289
10	PC6	-34.31388373	44.258835729	5.564620e+01	82.479485160

```
[5]: covs %>% data.table::fwrite("junctions.combined_covariates.txt", sep='\t')
```

1.2 Repreducibility Information

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

```
[1] "2021-07-14 20:55:38 EDT"

   user  system elapsed
  2.442   0.189   2.939

 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     2021-07-14


 Packages
 package      * version date        lib source
```

```
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)
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)
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)
forcats      * 0.5.1    2021-01-27 [1] CRAN (R 4.0.2)
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)
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)
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)
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)
rlang          0.4.11   2021-04-30 [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        1.6.2   2021-05-17 [1] CRAN (R 4.0.3)
stringr      * 1.4.0   2019-02-10 [1] CRAN (R 4.0.2)
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)

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