

# main

July 15, 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()
```

```
      RNum  BrNum
   <chr> <chr>
1 R5527  Br822
2 R2725  Br823
3 R4805  Br836
4 R5528  Br845
5 R4817  Br846
6 R4793  Br848
```

A data.table: 6 × 2

```
[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 = covsTx0 %>% 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	Br822	Br823	Br836	Br845
	<chr>	<dbl>	<dbl>	<dbl>	<dbl>
1	DxSchizo	0.00000000	0.000000e+00	0.000000000	0.000000000
2	snpPC1	-0.03381080	8.703800e-02	0.101455000	0.093891100
3	snpPC2	-0.00312935	-1.717740e-03	-0.000642496	-0.002248600
4	snpPC3	-0.00154150	4.134530e-04	-0.000923211	0.000546358
5	PC1	9.44435325	-1.404276e+02	26.401230170	0.176087784
6	PC2	-34.30907719	-4.600464e+01	-20.957772720	19.222797240
7	PC3	-36.21790019	3.344817e+01	32.318500527	-34.721646225
8	PC4	-29.95281698	3.715611e+01	31.010473643	-9.989048731
9	PC5	31.56023835	-1.215883e+01	15.831920734	-37.298617746
10	PC6	-8.85540373	3.609151e+01	8.255974971	19.244053361

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

## 1.2 Reproducibility Information

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

```
[1] "2021-07-14 23:59:54 EDT"
```

```
user system elapsed
2.390 0.244 2.850
```

```
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