main

September 15, 2021

1 Examine sample make-up

1.1 Samples after quality control

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

1.1.1 Load Caudate data

1.1.2 Load DLPFC data

1.1.3 Load Hippocampus data

1.1.4 Load DG data

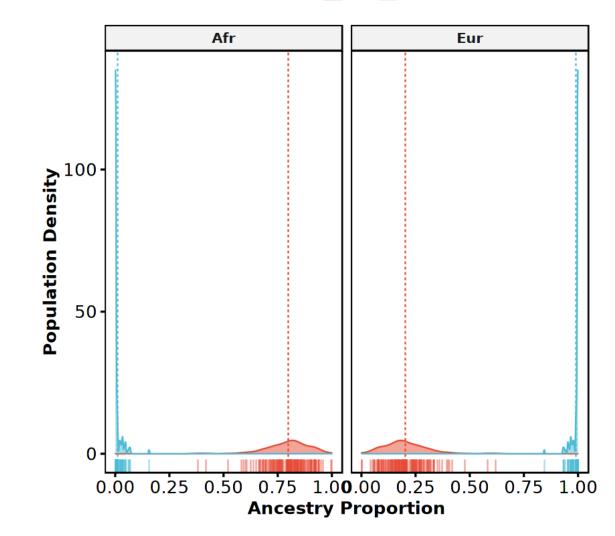
1.1.5 Merge data

1.2 STRUCTURE analysis

```
id
                          Afr
                                  Eur
                                          group
                          <dbl> <dbl>
                  <chr>
                                          <chr>
                  Br2374 0.007
                                  0.993
                                          CAUC
                  Br1857 0.001
                                  0.999
                                          CAUC
A data.table: 6 \times 4
                  Br1306 0.759
                                  0.241
                                          AA
                  Br2605 0.644
                                  0.356
                                          AA
                  Br1802 0.840
                                  0.160
                                          AA
                  Br2565 0.005
                                  0.995
                                          CAUC
```

```
AA_mean
                                      AA sd
                                                   AA_{max}
                                                             AA_{\underline{}}min EA_{\underline{}}mean
                                                                                      EA sd
                                                                                                    EA max
                group
                                                               <dbl>
                                                                                       <dbl>
                <fct>
                         <dbl>
                                      <dbl>
                                                   <dbl>
                                                                          <dbl>
                                                                                                    <dbl>
A tibble: 2 \times 9
                AA
                         0.78962609
                                      0.10611682
                                                   0.999
                                                               0.381
                                                                          0.2103739
                                                                                      0.10611682
                                                                                                    0.619
               CAUC
                        0.03087879
                                      0.02997578
                                                   0.156
                                                              0.001
                                                                          0.9691212
                                                                                      0.02997578
                                                                                                    0.999
```





1.3 eQTL analysis

```
[12]: pheno %>% dim
1. 1334 2. 21
[13]: print(paste("There are", unique(pheno$BrNum) %>% length, "unique BrNum."))
    [1] "There are 509 unique BrNum."
[14]: pheno %>% select(BrNum, Region) %>% distinct %>%
        mutate_if(is.character, as.factor) %>%
        group_by(Region) %>% count()
```

```
Region
                                           \mathbf{n}
                            <fct>
                                           <int>
                            Caudate
                                           400
      A grouped df: 4 \times 2
                            DentateGyrus
                                           161
                            DLPFC
                                           378
                            HIPPO
                                           395
[15]: pheno %>% select(BrNum, Race) %>% distinct %>%
           mutate_if(is.character, as.factor) %>%
           group_by(Race) %>% count()
                            Race
                                    n
                            <fct>
                                    <int>
      A grouped_df: 2 \times 2
                            \overline{AA}
                                    256
                            EA
                                    253
[16]: pheno %>% select(BrNum, Race, Region) %>% distinct %>%
           mutate if(is.character, as.factor) %>%
           group_by(Region, Race) %>% count()
                            Region
                                           Race
                                                   n
                            <fct>
                                           <fct>
                                                   <int>
                            Caudate
                                           \overline{AA}
                                                   206
                            Caudate
                                           EA
                                                   194
                            DentateGyrus
                                           AA
                                                   78
      A grouped df: 8 \times 3
                            DentateGyrus
                                           EA
                                                   83
                            DLPFC
                                           AA
                                                   204
                            DLPFC
                                           EA
                                                   174
                            HIPPO
                                                   213
                                           AA
                            HIPPO
                                           EA
                                                   182
[17]: pheno %>% select(BrNum, Sex, Region) %>% distinct %>%
           mutate_if(is.character, as.factor) %>%
           group_by(Region, Sex) %>% count()
                            Region
                                           Sex
                                                   \mathbf{n}
                            <fct>
                                           < fct >
                                                   <int>
                            Caudate
                                           \overline{\mathbf{F}}
                                                   126
                            Caudate
                                           Μ
                                                   274
                            DentateGyrus
                                           F
                                                   48
      A grouped df: 8 \times 3
                            DentateGyrus
                                           Μ
                                                   113
                            DLPFC
                                           F
                                                   121
                            DLPFC
                                           Μ
                                                   257
                            HIPPO
                                           F
                                                   126
                            HIPPO
                                           Μ
                                                   269
[18]: pheno %>% group_by(Region) %>%
```

summarise_at(vars(c("Age")), list(mean = mean, sd = sd))

```
<chr>
                                      < dbl >
                                                <dbl>
                      Caudate
                                     49.12390
                                               16.05379
      A tibble: 4 \times 3
                      DentateGyrus
                                     50.06770
                                               15.43849
                      DLPFC
                                     45.83574
                                               16.49445
                      HIPPO
                                     45.49527
                                               16.41527
[19]: pheno %>% group_by(Region, Race) %>%
         summarise_at(vars(c("Age")), list(mean = mean, sd = sd))
                            Region
                                           Race
                                                   mean
                                                              \operatorname{sd}
                            <chr>
                                                              <dbl>
                                           <chr>
                                                    <dbl>
                            Caudate
                                           AA
                                                   48.81325
                                                              14.49676
                            Caudate
                                           EA
                                                   49.45376
                                                              17.58900
                            DentateGyrus
                                           AA
                                                   50.18423
                                                              15.53374
      A grouped df: 8 \times 4
                            DentateGyrus
                                           EA
                                                   49.95819
                                                              15.44210
                            DLPFC
                                           AA
                                                   46.97896
                                                              15.34261
                            DLPFC
                                           EA
                                                   44.49542
                                                              17.70090
                            HIPPO
                                           AA
                                                   46.34080
                                                              15.61922
                            HIPPO
                                           EA
                                                   44.50571
                                                              17.29140
[20]: pheno %>% filter(RIN != "NA") %>% mutate("RIN"=as.numeric(unlist(RIN))) %>%
           group_by(Region) %>% summarise_at(vars(c("RIN")), list(mean = mean, sd =__
        (sd)
                      Region
                                     mean
                                               \operatorname{sd}
                      <chr>
                                      <dbl>
                                                <dbl>
                      Caudate
                                     7.861000
                                               0.8648983
      A tibble: 4 \times 3
                      DentateGyrus
                                     5.208403
                                               1.1871187
                      DLPFC
                                     7.699471
                                               0.9340876
                      HIPPO
                                     7.616962
                                               1.0311104
[21]: pheno %>% filter(RIN != "NA") %>% mutate("RIN"=as.numeric(unlist(RIN))) %>%
           group_by(Region, Race) %>% summarise_at(vars(c("RIN")), list(mean = mean,__
       \rightarrowsd = sd))
                            Region
                                           Race
                                                              \operatorname{sd}
                                                   mean
                                                              <dbl>
                            <chr>
                                           <chr>
                                                    <dbl>
                            Caudate
                                           AA
                                                              0.8416464
                                                   7.859709
                            Caudate
                                           EA
                                                   7.862371
                                                              0.8911055
                            DentateGyrus
                                           AA
                                                   5.206349
                                                              1.2062837
      A grouped df: 8 \times 4
                            DentateGyrus
                                           EA
                                                   5.210714
                                                              1.1760765
                            DLPFC
                                           AA
                                                   7.678922
                                                              0.9445184
                            DLPFC
                                           EA
                                                   7.723563
                                                              0.9238440
                            HIPPO
                                           AA
                                                   7.604225
                                                              1.0509344
                            HIPPO
                                           EA
                                                   7.631868
                                                              1.0101014
```

Region

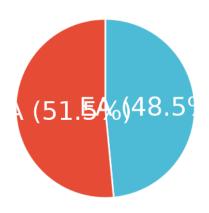
mean

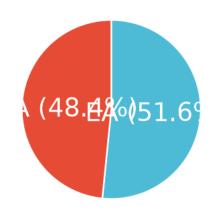
 sd

1.3.1 Pie chart

Caudate

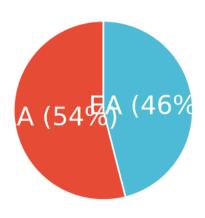
Dentate Gyrus

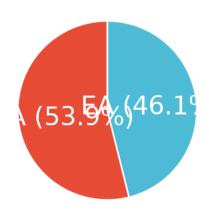




DLPFC

Hippocampus





1.4 Adult individuals for expression related analysis

```
[24]: pheno = pheno %>% filter(Age > 17, Dx == "Control") pheno %>% dim
```

1. 785 2. 21

[25]: print(paste("There are", unique(pheno\$BrNum) %>% length, "unique BrNum."))

[1] "There are 292 unique BrNum."

[26]: pheno %>% select(BrNum, Region) %>% distinct %>%
 mutate_if(is.character, as.factor) %>%
 group_by(Region) %>% count()

```
Region
                                           n
                            <fct>
                                            <int>
                            Caudate
                                            240
      A grouped df: 4 \times 2
                            DentateGyrus
                                           90
                            DLPFC
                                            212
                            HIPPO
                                           243
[27]: pheno %>% select(BrNum, Race) %>% distinct %>%
           mutate_if(is.character, as.factor) %>%
           group_by(Race) %>% count()
                            Race
                                    \mathbf{n}
                            <fct>
                                    <int>
      A grouped_df: 2 \times 2
                            \overline{AA}
                                    151
                            EA
                                    141
[28]: pheno %>% select(BrNum, Race, Region) %>% distinct %>%
           mutate if(is.character, as.factor) %>%
           group_by(Region, Race) %>% count()
                            Region
                                           Race
                                                    n
                            <fct>
                                            <fct>
                                                    <int>
                            Caudate
                                            \overline{AA}
                                                    122
                            Caudate
                                           EA
                                                    118
                            DentateGyrus
                                           AA
                                                    47
      A grouped df: 8 \times 3
                            DentateGyrus
                                           EA
                                                    43
                            DLPFC
                                            AA
                                                    123
                            DLPFC
                                           EA
                                                    89
                            HIPPO
                                            AA
                                                    133
                            HIPPO
                                           EA
                                                    110
[29]: pheno %>% select(BrNum, Sex, Region) %>% distinct %>%
           mutate_if(is.character, as.factor) %>%
           group_by(Region, Sex) %>% count()
                            Region
                                           Sex
                                                    \mathbf{n}
                            <fct>
                                            < fct >
                                                    <int>
                            Caudate
                                           \overline{\mathbf{F}}
                                                    71
                            Caudate
                                           Μ
                                                    169
                            DentateGyrus
                                           F
                                                    26
      A grouped df: 8 \times 3
                            DentateGyrus
                                           Μ
                                                    64
                            DLPFC
                                           F
                                                    66
                            DLPFC
                                           Μ
                                                    146
                            HIPPO
                                           F
                                                    74
                            HIPPO
                                           Μ
                                                    169
```

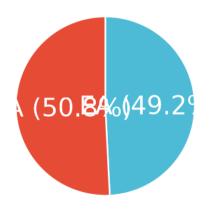
summarise_at(vars(c("Age")), list(mean = mean, sd = sd))

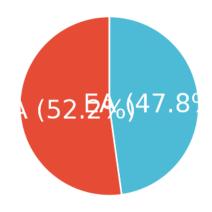
[30]: pheno %>% group_by(Region) %>%

```
Region
                                     mean
                                                \operatorname{sd}
                      <chr>
                                      <dbl>
                                                <dbl>
                      Caudate
                                     48.31150
                                                \overline{15.84692}
      A tibble: 4 \times 3
                      DentateGyrus
                                     47.88311
                                                15.02380
                      DLPFC
                                     45.16991
                                                14.76717
                      HIPPO
                                     44.56724
                                                14.73045
[31]: pheno %>% group_by(Region, Race) %>%
         summarise_at(vars(c("Age")), list(mean = mean, sd = sd))
                            Region
                                           Race
                                                    mean
                                                              \operatorname{sd}
                            <chr>
                                                              <dbl>
                                           <chr>
                                                    <dbl>
                            Caudate
                                           AA
                                                    45.63770
                                                              14.72979
                            Caudate
                                           EA
                                                    51.07593
                                                              16.53588
                            DentateGyrus
                                           AA
                                                    45.85043
                                                              16.32827
      A grouped df: 8 \times 4
                            DentateGyrus
                                           EA
                                                    50.10488
                                                              13.28980
                            DLPFC
                                           AA
                                                    44.12511
                                                              14.97092
                            DLPFC
                                           EA
                                                    46.61386
                                                              14.43996
                            HIPPO
                                           AA
                                                    43.30015
                                                              14.73609
                            HIPPO
                                           EA
                                                    46.09927
                                                              14.64404
[32]: pheno %>% filter(RIN != "NA") %>% mutate("RIN"=as.numeric(unlist(RIN))) %>%
           group_by(Region) %>% summarise_at(vars(c("RIN")), list(mean = mean, sd =__
        ⇒sd))
                      Region
                                     mean
                                                \operatorname{sd}
                      <chr>
                                      <dbl>
                                                <dbl>
                      Caudate
                                     7.850000
                                                0.7956997
      A tibble: 4 \times 3
                      DentateGyrus
                                     5.315152
                                                1.2186048
                      DLPFC
                                     7.699057
                                                0.8803807
                      HIPPO
                                     7.735391
                                                0.9668378
[33]: pheno %>% filter(RIN != "NA") %>% mutate("RIN"=as.numeric(unlist(RIN))) %>%
           group_by(Region, Race) %>% summarise_at(vars(c("RIN")), list(mean = mean,__
       \rightarrowsd = sd))
                            Region
                                           Race
                                                              \operatorname{sd}
                                                    mean
                                                              <dbl>
                            <chr>
                                           <chr>
                                                    <dbl>
                            Caudate
                                           AA
                                                              0.7993477
                                                    7.829508
                            Caudate
                                           EA
                                                    7.871186
                                                              0.7947587
                            DentateGyrus
                                           AA
                                                    5.447368
                                                              1.2173824
      A grouped df: 8 \times 4
                            DentateGyrus
                                           EA
                                                    5.135714
                                                              1.2190507
                            DLPFC
                                           AA
                                                    7.696748
                                                              0.8851169
                            DLPFC
                                           EA
                                                    7.702247
                                                              0.8787876
                            HIPPO
                                           AA
                                                    7.715038
                                                              0.9754173
                                           EA
                            HIPPO
                                                    7.760000
                                                              0.9602370
[34]: ## Get and annotate plot
      cc_pie = annotate_figure(plot_pie("Caudate"),
```

Caudate

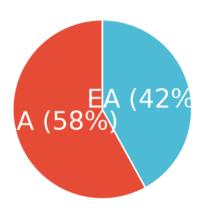
Dentate Gyrus

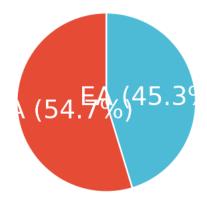




DLPFC

Hippocampus





1.5 Reproducibility Information

```
[35]: Sys.time()
      proc.time()
      options(width = 120)
      sessioninfo::session_info()
     [1] "2021-09-15 11:54:23 EDT"
        user system elapsed
      22.678
                1.593 25.373
       Session info
      setting value
      version R version 4.0.3 (2020-10-10)
      os
               Arch Linux
               x86_64, linux-gnu
      system
      ui
               X11
      language (EN)
               en_US.UTF-8
      collate
      ctype
               en_US.UTF-8
               America/New_York
      tz
      date
               2021-09-15
       Packages
      package
                            * version date
                                                   lib source
      abind
                              1.4-5
                                       2016-07-21 [1] CRAN (R 4.0.2)
                              0.2.1
                                       2019-03-21 [1] CRAN (R 4.0.2)
      assertthat
      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)
      Biobase
                            * 2.50.0
                                       2020-10-27 [1] Bioconductor
      BiocGenerics
                            * 0.36.1
                                       2021-04-16 [1] Bioconductor
                              1.0 - 7
                                       2021-04-24 [1] CRAN (R 4.0.3)
      bitops
      broom
                              0.7.9
                                       2021-07-27 [1] CRAN (R 4.0.3)
                              1.5-12.2 2020-07-07 [1] CRAN (R 4.0.2)
      Cairo
                                       2021-06-27 [1] CRAN (R 4.0.3)
      car
                              3.0 - 11
      carData
                              3.0-4
                                       2020-05-22 [1] CRAN (R 4.0.2)
                                       2016-07-27 [1] CRAN (R 4.0.2)
      cellranger
                              1.1.0
                              3.0.1
                                       2021-07-17 [1] CRAN (R 4.0.3)
      cli
                              2.0-2
                                       2021-06-24 [1] CRAN (R 4.0.3)
      colorspace
                                       2020-12-30 [1] CRAN (R 4.0.2)
      cowplot
                              1.1.1
      crayon
                              1.4.1
                                       2021-02-08 [1] CRAN (R 4.0.3)
                              4.3.2
                                       2021-06-23 [1] CRAN (R 4.0.3)
      curl
      data.table
                              1.14.0
                                       2021-02-21 [1] CRAN (R 4.0.3)
      DBI
                                       2021-01-15 [1] CRAN (R 4.0.2)
                              1.1.1
      dbplyr
                              2.1.1
                                       2021-04-06 [1] CRAN (R 4.0.3)
      DelayedArray
                              0.16.3
                                       2021-03-24 [1] Bioconductor
```

0.6.27

* 1.0.7

digest

dplyr

2020-10-24 [1] CRAN (R 4.0.2)

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)
GenomeInfoDb	*	1.26.7	2021-04-08	[1]	Bioco	ndı	ıctor
GenomeInfoDbData		1.2.4	2021-02-02	[1]	Bioconductor		
GenomicRanges	*	1.42.0	2020-10-27	[1]	Bioconductor		
ggplot2	*	3.3.5	2021-06-25	[1]	CRAN	(R	4.0.3)
ggpubr	*	0.4.0	2020-06-27	[1]	CRAN		4.0.2)
ggsci		2.9	2018-05-14	[1]	CRAN		4.0.2)
ggsignif		0.6.2	2021-06-14				4.0.3)
glue		1.4.2	2020-08-27		CRAN		4.0.2)
gridExtra		2.3	2017-09-09	[1]	CRAN		4.0.2)
gtable		0.3.0	2019-03-25	[1]	CRAN		4.0.2)
haven		2.4.3	2021-08-04	[1]	CRAN		4.0.3)
hms		1.1.0	2021-05-17	[1]	CRAN		4.0.3)
htmltools		0.5.2	2021-08-25	[1]	CRAN		4.0.3)
httr		1.4.2	2020-07-20	[1]	CRAN		4.0.2)
IRanges	*	2.24.1	2020-12-12				
IRdisplay		1.0	2021-01-20	[1]			4.0.2)
IRkernel		1.2	2021-05-11	[1]	CRAN		4.0.3)
jsonlite		1.7.2	2020-12-09	[1]	CRAN		4.0.2)
labeling		0.4.2	2020 12 03	[1]	CRAN		4.0.2)
lattice		0.20-41	2020-10-20	[2]	CRAN		4.0.3)
		1.0.0	2020-04-02	[1]	CRAN		4.0.3)
lifecycle lubridate		1.7.10	2021-02-15	[1]	CRAN		4.0.3)
			2021-02-20				4.0.2)
magrittr		2.0.1					
Matrix		1.3-4	2021-06-01	[1]	CRAN		4.0.3)
MatrixGenerics		1.2.1	2021-01-30		Bioco		
matrixStats	*	0.60.1	2021-08-23				
modelr		0.1.8	2020-05-19				4.0.2)
munsell		0.5.0	2018-06-12				4.0.2)
openxlsx		4.2.4	2021-06-16		CRAN		4.0.3)
pbdZMQ		0.3-5	2021-02-10	[1]			4.0.3)
pillar		1.6.2	2021-07-29				4.0.3)
pkgconfig		2.0.3	2019-09-22	[1]	CRAN		4.0.2)
purrr	*	0.3.4	2020-04-17	[1]			4.0.2)
R6		2.5.1	2021-08-19				
Rcpp		1.0.7	2021-07-07				4.0.3)
RCurl			2021-08-17		CRAN		4.0.3)
readr	*	2.0.1	2021-08-10	[1]	CRAN		4.0.3)
readxl		1.3.1	2019-03-13				4.0.2)
repr		1.1.3	2021-01-21	[1]	CRAN	(R	4.0.2)

```
2.0.1
                                 2021-08-05 [1] CRAN (R 4.0.3)
reprex
                       0.5.27
                                 2021-06-21 [1] CRAN (R 4.0.3)
rio
                                 2021-04-30 [1] CRAN (R 4.0.3)
                       0.4.11
rlang
rstatix
                       0.7.0
                                 2021-02-13 [1] CRAN (R 4.0.3)
                       0.13
                                 2020-11-12 [1] CRAN (R 4.0.2)
rstudioapi
                       1.0.1
                                 2021-07-26 [1] CRAN (R 4.0.3)
rvest
S4Vectors
                     * 0.28.1
                                 2020-12-09 [1] Bioconductor
                                 2020-05-11 [1] CRAN (R 4.0.2)
scales
                       1.1.1
sessioninfo
                       1.1.1
                                 2018-11-05 [1] CRAN (R 4.0.2)
                                 2021-08-25 [1] CRAN (R 4.0.3)
stringi
                       1.7.4
                     * 1.4.0
                                 2019-02-10 [1] CRAN (R 4.0.2)
stringr
SummarizedExperiment * 1.20.0
                                 2020-10-27 [1] Bioconductor
                                 2021-02-20 [1] CRAN (R 4.0.3)
svglite
                       2.0.0
systemfonts
                                 2021-05-11 [1] CRAN (R 4.0.3)
                       1.0.2
                     * 3.1.4
                                 2021-08-25 [1] CRAN (R 4.0.3)
tibble
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)
                                 2021-04-15 [1] CRAN (R 4.0.3)
tidyverse
                     * 1.3.1
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)
                       0.1 - 4
                                 2020-02-26 [1] CRAN (R 4.0.2)
uuid
                       0.3.8
                                 2021-04-29 [1] CRAN (R 4.0.3)
vctrs
                                 2021-04-18 [1] CRAN (R 4.0.3)
withr
                       2.4.2
xml2
                       1.3.2
                                 2020-04-23 [1] CRAN (R 4.0.2)
XVector
                       0.30.0
                                 2020-10-27 [1] Bioconductor
                       2.2.0
                                 2021-05-31 [1] CRAN (R 4.0.3)
zip
                                 2020-10-27 [1] Bioconductor
zlibbioc
                       1.36.0
```

^{[1] /}home/jbenja13/R/x86_64-pc-linux-gnu-library/4.0

^{[2] /}usr/lib/R/library