R & OMICS

8/16/2022

Table of contents

1

This is a Quarto book to impress Jonathan.

This web resource covers basic R operations on gene expression tables such as data reading/writing, exploring wrangling and plotting. It is largely based on R *tidyverse* package and can be useful for Next Generation Sequencing functional data analysis.

The resource was created and is being maintained by Jordan RAMILOWSKI from the Bioinformatics Laboratory at Yokohama City University.

2 Required Data

- Download the data directory from Dropbox (link to come)

- setwd()

- Install and load all required R packages

```r
setwd('/Users/jordan/R_in_OMICS/input') # example
library(data.table)
```

3 Reading and Writing files in *R*

3.1 data.table::fread

Simple and efficient way of reading R files:

```r
library(data.table)
data = fread(file = 'connectome_1in.tsv')
data = fread(file = 'connectome_1in.tsv', nThread = 4)  # can use many processors 'nThread
```

'fread' vs 'read.delim'(base) speed

```r
# 'fread':
t0 = Sys.time()
data = fread(file = 'connectome_1in.tsv')
t1 = Sys.time()-t0

# 'read.delim':
t0 = Sys.time()
data = read.delim2(file = 'connectome_1in.tsv')
t2 = Sys.time()-t0
```

```
[1] "fread: Time difference of 0.18643 secs"
```

```
[1] "read.delim: Time difference of 4.577949 secs"
```

3.2 data.table::fwrite

Simple and efficient way of writing R files:

```r
fwrite(data, file = 'connectome_1out.csv')              # csv file (default)
fwrite(data, file = 'connectome_1out.tsv', sep = '\t') # tab-delim file (own specification
```

4 Exploring data in R

```r
library(data.table)
```

Read in and look at first 10 columns of **connectome** data set:

```r
connectome_1=fread(file='connectome_1in.tsv', select = 1:10) # first 10 columns
```

```r
dim(connectome_1) # number of rows / columns
```

```
[1] 17263     10
```

```r
colnames(connectome_1) # names of data
```

```
 [1] "ApprovedSymbol"        "Localization"          "Taxon"
 [4] "F5_min"                "F5_med"                "F5_max"
 [7] "Adipocyte Breast"      "Adipocyte Omental"     "Adipocyte Perirenal"
[10] "Adipocyte Subcutaneous"
```

Closer look using specific R packages/functions:

4.1 using 'str'

```r
str(connectome_1)
```

```
Classes 'data.table' and 'data.frame':  17263 obs. of  10 variables:
 $ ApprovedSymbol        : chr  "A1BG" "A1CF" "A2M" "A2ML1" ...
 $ Localization          : chr  "secreted" "multiple" "multiple" "secreted" ...
 $ Taxon                 : chr  "Eutheria" "Deuterostomia" "Euteleostomi" "Euteleostomi" ...
 $ F5_min                : num  0.264 0 0 0 0 ...
```

```
$ F5_med               : num  16.94 0 1.21 0 9.72 ...
$ F5_max               : num  2378 153 1515 206 80 ...
$ Adipocyte Breast     : num  36.8 0 90.3 0 24.4 ...
$ Adipocyte Omental    : num  35.2 0 121.4 0 10.8 ...
$ Adipocyte Perirenal  : num  49.7 0 50.6 0 17.4 ...
$ Adipocyte Subcutaneous: num  41.1 0 63.4 0 27.7 ...
- attr(*, ".internal.selfref")=<externalptr>
```

4.2 using 'skimr::skim'

```r
library(skimr)
skimr::skim(connectome_1)
```

Table 4.1: Data summary

Name	connectome_1
Number of rows	17263
Number of columns	10
Key	NULL
Column type frequency:	
character	3
numeric	7
Group variables	None

Variable type: character

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
ApprovedSymbol	0	1	1	10	0	17263	0
Localization	0	1	3	15	0	7	0
Taxon	0	1	0	16	614	17	0

Variable type: numeric

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100	hist
F5_min	0	1	3.88	17.74	0	0.00	0.00	2.55	1451.39	
F5_med	0	1	32.14	154.02	0	0.03	5.94	22.71	6876.78	

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100	hist
F5_max	0	1	476.92	4237.26	0	23.85	75.89	211.27	309580.87	
Adipocyte Breast	0	1	47.49	304.12	0	0.11	6.73	26.99	12064.67	
Adipocyte Omental	0	1	51.19	479.86	0	0.05	6.70	26.77	48575.26	
Adipocyte Perirenal	0	1	54.70	661.88	0	0.00	6.11	25.30	75911.77	
Adipocyte Subcutaneous	0	1	46.69	259.45	0	0.09	5.78	25.00	10252.56	

5 Basic dplyr operations on data columns

```
-- Attaching packages -------------------------------------- tidyverse 1.3.2 --
v ggplot2 3.3.6      v purrr    0.3.4
v tibble  3.1.8      v dplyr    1.0.9
v tidyr   1.2.0      v stringr  1.4.0
v readr   2.1.2      v forcats  0.5.1
-- Conflicts ----------------------------------------- tidyverse_conflicts() --
x dplyr::between()   masks data.table::between()
x dplyr::filter()    masks stats::filter()
x dplyr::first()     masks data.table::first()
x dplyr::lag()       masks stats::lag()
x dplyr::last()      masks data.table::last()
x purrr::transpose() masks data.table::transpose()
```

```
        : 'magrittr'
```

```
    'package:purrr'         :
```

```
  set_names
```

```
    'package:tidyr'         :
```

```
  extract
```

```r
# the entire data has 17,263 rows and 150 columns
data = fread(file = 'connectome_1in.tsv')
dim(data)
```

```
[1] 17263    150
```

```r
# skim(data) #you can comment this out an run
```

5.1 'dplyr::select

```
# Simple
data %>% dplyr::select(ApprovedSymbol, F5_min, F5_med, F5_max) # selects columns by name
```

```
       ApprovedSymbol   F5_min  F5_med   F5_max
    1:           A1BG  0.263767  16.943 2377.989
    2:           A1CF  0.000000   0.000  153.024
    3:            A2M  0.000000   1.212 1515.493
    4:           A2ML1  0.000000   0.000  206.355
    5:          A4GALT  0.000000   9.725   80.009
   ---
17259:          ZYG11A  0.000000   0.264   11.104
17260:          ZYG11B  4.388704  16.580   65.872
17261:             ZYX 20.467403 468.361 1826.202
17262:           ZZEF1  1.744677   9.323   93.429
17263:            ZZZ3  4.360941  20.657   68.401
```

```
data %>% dplyr::select(1:4, 50) # select columns by position
```

```
       ApprovedSymbol Localization         Taxon    F5_min Eosinophils
    1:           A1BG      secreted      Eutheria  0.263767       2.636
    2:           A1CF      multiple Deuterostomia  0.000000       0.000
    3:            A2M      multiple  Euteleostomi  0.000000       0.000
    4:           A2ML1      secreted  Euteleostomi  0.000000       0.000
    5:          A4GALT      multiple      Tetrapoda  0.000000       0.000
   ---
17259:          ZYG11A       nucleus  Euteleostomi  0.000000       0.000
17260:          ZYG11B       nucleus  Euteleostomi  4.388704      41.867
17261:             ZYX      multiple         Theria 20.467403     742.793
17262:           ZZEF1         other        Amniota  1.744677      93.429
17263:            ZZZ3       nucleus Deuterostomia  4.360941       9.187
```

```
data %>% dplyr::select(CellCompartment=2, geneID=1, Minimum_expr=F5_min) ## dplyr::rename
```

```
    CellCompartment geneID Minimum_expr
 1:        secreted   A1BG     0.263767
 2:        multiple   A1CF     0.000000
 3:        multiple    A2M     0.000000
```

```
   4:        secreted  A2ML1     0.000000
   5:        multiple A4GALT     0.000000
  ---
17259:        nucleus ZYG11A     0.000000
17260:        nucleus ZYG11B     4.388704
17261:       multiple    ZYX    20.467403
17262:          other  ZZEF1     1.744677
17263:        nucleus   ZZZ3     4.360941
```

```r
head(data)[1:2,1:3] # here are the original names and order
```

```
   ApprovedSymbol Localization        Taxon
1:          A1BG     secreted      Eutheria
2:          A1CF     multiple Deuterostomia
```

```r
data %>% dplyr::select(contains("Macrophage")) # selects columns by pattern
```

```
      Granulocyte Macrophage progenitor Macrophage Monocyte derived
   1:                        34.492                        17.923
   2:                         0.000                         0.000
   3:                         0.000                       208.304
   4:                         0.000                         0.000
   5:                         0.000                         0.000
  ---
17259:                         0.000                         0.000
17260:                        28.104                        29.538
17261:                       226.111                       544.167
17262:                        77.925                        21.500
17263:                        22.994                        24.682
```

```r
data %>% dplyr::select(starts_with("CD34+")) # selects columns by starting pattern
```

```
      CD34+ Haematopoietic stem cell CD34+ progenitor
   1:                        23.968           34.219
   2:                         0.000            0.000
   3:                         1.317            0.280
   4:                         0.000            0.000
   5:                         2.370            0.280
```

```
        ---
17259:                          0.527               0.000
17260:                         14.750               6.768
17261:                         91.394             128.700
17262:                         12.379              26.067
17263:                         21.861              28.870
```

```
  data %>% dplyr::select(ends_with("Endothelial")) # select columns by ending pattern
```

```
      Hepatic Sinusoidal Endothelial Renal Glomerular Endothelial
   1:                         11.219                         4.962
   2:                          0.000                         0.000
   3:                         36.612                        47.414
   4:                          0.000                         0.000
   5:                         13.360                        36.860
  ---
17259:                          1.070                         1.272
17260:                         10.445                         9.821
17261:                        894.871                       826.222
17262:                          9.780                        13.692
17263:                         13.323                        14.051
```

```
  data %>%  dplyr::select(where(is.character))# selects columns by data type
```

```
      ApprovedSymbol Localization       Taxon
   1:           A1BG     secreted      Eutheria
   2:           A1CF     multiple Deuterostomia
   3:            A2M     multiple  Euteleostomi
   4:          A2ML1     secreted  Euteleostomi
   5:         A4GALT     multiple     Tetrapoda
  ---
17259:         ZYG11A      nucleus  Euteleostomi
17260:         ZYG11B      nucleus  Euteleostomi
17261:            ZYX     multiple        Theria
17262:          ZZEF1        other       Amniota
17263:           ZZZ3      nucleus Deuterostomia
```

```
  # Same as above, but works by excluding columns
  data %>%  dplyr::select(-c(6:150)) # excludes columns by position
```

```
       ApprovedSymbol Localization           Taxon     F5_min   F5_med
    1:           A1BG     secreted        Eutheria   0.263767   16.943
    2:           A1CF     multiple Deuterostomia   0.000000    0.000
    3:            A2M     multiple   Euteleostomi   0.000000    1.212
    4:          A2ML1     secreted   Euteleostomi   0.000000    0.000
    5:         A4GALT     multiple       Tetrapoda   0.000000    9.725
   ---
17259:         ZYG11A      nucleus   Euteleostomi   0.000000    0.264
17260:         ZYG11B      nucleus   Euteleostomi   4.388704   16.580
17261:            ZYX     multiple          Theria  20.467403  468.361
17262:          ZZEF1        other         Amniota   1.744677    9.323
17263:           ZZZ3      nucleus Deuterostomia   4.360941   20.657
```

```r
data %>%  dplyr::select(!where(is.numeric))# excludes columns by type using '!'
```

```
       ApprovedSymbol Localization           Taxon
    1:           A1BG     secreted        Eutheria
    2:           A1CF     multiple Deuterostomia
    3:            A2M     multiple   Euteleostomi
    4:          A2ML1     secreted   Euteleostomi
    5:         A4GALT     multiple       Tetrapoda
   ---
17259:         ZYG11A      nucleus   Euteleostomi
17260:         ZYG11B      nucleus   Euteleostomi
17261:            ZYX     multiple          Theria
17262:          ZZEF1        other         Amniota
17263:           ZZZ3      nucleus Deuterostomia
```

```r
# More complex

# Example:
# select and dplyr::rename column 1 and then,
# select all numeric columns starting with 'T' (character variable 'Taxon' will be exclude

data %>% dplyr::select(geneSymbol=1, starts_with("T") & where(is.numeric))
```

```
     geneSymbol Tenocyte Trabecular Meshwork Tracheal Epithelial
  1:       A1BG   18.775               30.337                10.808
  2:       A1CF    0.000                0.000                 0.000
  3:        A2M    5.266                0.000                 0.223
```

```
   4:     A2ML1     0.000                0.000              4.885
   5:    A4GALT    61.794                6.097             11.199
  ---
17259:    ZYG11A     0.000                2.393              0.096
17260:    ZYG11B    24.898               10.372              7.905
17261:       ZYX   489.985             1313.519            616.760
17262:     ZZEF1     6.676                8.898              6.903
17263:      ZZZ3    22.721               19.108             11.410
```

```
#A few more 'select' arguments are: 'num_range()','matches()', 'any_of(), 'all_of()'
```

5.2 dplyr::rename / dplyr::rename_with

```
# renaming can also be done with 'select', but 'dplyr::rename' is easier

data=fread(file = 'connectome_1in.tsv', select = c(1,7:9))
data %>% dplyr::rename(geneSymbol = ApprovedSymbol)
```

```
       geneSymbol Adipocyte Breast Adipocyte Omental Adipocyte Perirenal
   1:        A1BG            36.847            35.174             49.723
   2:        A1CF             0.000             0.000              0.000
   3:         A2M            90.272           121.423             50.596
   4:       A2ML1             0.000             0.000              0.000
   5:      A4GALT            24.386            10.811             17.447
  ---
17259:      ZYG11A             0.091             0.000              0.000
17260:      ZYG11B            30.164            25.141             18.319
17261:         ZYX           473.814           442.646            389.935
17262:       ZZEF1             9.713             8.754              1.745
17263:        ZZZ3            18.076            20.905             19.191
```

```
data %>% dplyr::rename_with(tolower)
```

```
       approvedsymbol adipocyte breast adipocyte omental adipocyte perirenal
   1:            A1BG            36.847            35.174             49.723
   2:            A1CF             0.000             0.000              0.000
   3:             A2M            90.272           121.423             50.596
   4:           A2ML1             0.000             0.000              0.000
```

```
    5:         A4GALT       24.386          10.811            17.447
    ---
17259:         ZYG11A        0.091           0.000             0.000
17260:         ZYG11B       30.164          25.141            18.319
17261:            ZYX      473.814         442.646           389.935
17262:          ZZEF1        9.713           8.754             1.745
17263:           ZZZ3       18.076          20.905            19.191
```

```r
data %>% dplyr::rename_with(toupper, where(is.numeric)) # dplyr::renames to upper - here,
```

```
       ApprovedSymbol ADIPOCYTE BREAST ADIPOCYTE OMENTAL ADIPOCYTE PERIRENAL
    1:           A1BG           36.847           35.174             49.723
    2:           A1CF            0.000            0.000              0.000
    3:            A2M           90.272          121.423             50.596
    4:          A2ML1            0.000            0.000              0.000
    5:         A4GALT           24.386           10.811             17.447
    ---
17259:         ZYG11A            0.091            0.000              0.000
17260:         ZYG11B           30.164           25.141             18.319
17261:            ZYX          473.814          442.646            389.935
17262:          ZZEF1            9.713            8.754              1.745
17263:           ZZZ3           18.076           20.905             19.191
```

5.3 dplyr::relocate

```r
# moves columns to the front; this can also be done with 'select', but 'relocate' is easie

data=fread(file = 'connectome_1in.tsv', select = c(1,7:9))
data %>% relocate(starts_with('Adip'))
```

```
       Adipocyte Breast Adipocyte Omental Adipocyte Perirenal ApprovedSymbol
    1:           36.847           35.174             49.723           A1BG
    2:            0.000            0.000              0.000           A1CF
    3:           90.272          121.423             50.596            A2M
    4:            0.000            0.000              0.000          A2ML1
    5:           24.386           10.811             17.447         A4GALT
    ---
17259:            0.091            0.000              0.000         ZYG11A
17260:           30.164           25.141             18.319         ZYG11B
```

```
17261:          473.814          442.646          389.935          ZYX
17262:            9.713            8.754            1.745          ZZEF1
17263:           18.076           20.905           19.191          ZZZ3
```

```r
data %>% relocate(where(is.numeric))
```

```
       Adipocyte Breast Adipocyte Omental Adipocyte Perirenal ApprovedSymbol
   1:            36.847           35.174            49.723           A1BG
   2:             0.000            0.000             0.000           A1CF
   3:            90.272          121.423            50.596           A2M
   4:             0.000            0.000             0.000           A2ML1
   5:            24.386           10.811            17.447           A4GALT
 ---
17259:             0.091            0.000             0.000           ZYG11A
17260:            30.164           25.141            18.319           ZYG11B
17261:           473.814          442.646           389.935           ZYX
17262:             9.713            8.754             1.745           ZZEF1
17263:            18.076           20.905            19.191           ZZZ3
```

```r
# relocate can take other arguments '.before', '.after'
?relocate #check help
```

6 Basic dplyr operations on data rows

6.1 dplyr::filter

```r
# keeps rows based on desired criteria
# simple

data=fread(file = 'connectome_1in.tsv', select = c(1,5:8))
data %>%  dplyr::filter(F5_max>5) #keeps rows using a numeric value
```

	ApprovedSymbol	F5_med	F5_max	Adipocyte Breast	Adipocyte Omental
1:	A1BG	16.943	2377.989	36.847	35.174
2:	A1CF	0.000	153.024	0.000	0.000
3:	A2M	1.212	1515.493	90.272	121.423
4:	A2ML1	0.000	206.355	0.000	0.000
5:	A4GALT	9.725	80.009	24.386	10.811

15316:	ZYG11A	0.264	11.104	0.091	0.000
15317:	ZYG11B	16.580	65.872	30.164	25.141
15318:	ZYX	468.361	1826.202	473.814	442.646
15319:	ZZEF1	9.323	93.429	9.713	8.754
15320:	ZZZ3	20.657	68.401	18.076	20.905

```r
data=fread(file = 'connectome_1in.tsv', select = c(1:6))
my_genes=c('A1BG', 'CDKN3', 'ZPBP')
data %>%  dplyr::filter(ApprovedSymbol %in% my_genes) #keeps rows based on custom string
```

	ApprovedSymbol	Localization	Taxon	F5_min	F5_med	F5_max
1:	A1BG	secreted	Eutheria	0.263767	16.943	2377.989
2:	CDKN3	multiple	Amniota	0.000000	14.600	226.150
3:	ZPBP	secreted	Amniota	0.000000	0.000	0.651

```
# in conjunction with 'str_detect' from stringr package; (similar to 'grep')
data %>% dplyr::filter(str_detect(ApprovedSymbol, '^HIST2')) #keeps all HIST2 genes
```

```
    ApprovedSymbol Localization     Taxon      F5_min     F5_med      F5_max
 1:     HIST2H2AA3      nucleus Eukaryota    0.000000      0.431     671.790
 2:     HIST2H2AA4      nucleus Eukaryota    0.000000      0.431     671.790
 3:      HIST2H2AB      nucleus Eukaryota    9.958116     91.570    1240.587
 4:      HIST2H2BE     multiple Eukaryota   56.484107    416.034    2790.341
 5:      HIST2H2BF      nucleus Eukaryota    0.918399     21.239     457.095
 6:       HIST2H3A      nucleus Eukaryota    0.000000      0.174     797.346
 7:       HIST2H3C      nucleus Eukaryota    0.000000      0.174     797.346
 8:       HIST2H3D      nucleus Eukaryota    0.000000      0.174     797.346
 9:       HIST2H4A      nucleus Eukaryota  170.420082   1224.638   19898.334
10:       HIST2H4B      nucleus Eukaryota  170.420082   1224.638   19898.334
```

```
data %>% dplyr::filter(!str_detect(ApprovedSymbol, '^HIST2')) #removes (using '!') all HIS
```

```
        ApprovedSymbol Localization         Taxon     F5_min   F5_med     F5_max
     1:           A1BG     secreted      Eutheria   0.263767   16.943   2377.989
     2:           A1CF     multiple Deuterostomia   0.000000    0.000    153.024
     3:            A2M     multiple  Euteleostomi   0.000000    1.212   1515.493
     4:          A2ML1     secreted  Euteleostomi   0.000000    0.000    206.355
     5:         A4GALT     multiple      Tetrapoda   0.000000    9.725     80.009
    ---
 17249:         ZYG11A      nucleus  Euteleostomi   0.000000    0.264     11.104
 17250:         ZYG11B      nucleus  Euteleostomi   4.388704   16.580     65.872
 17251:            ZYX     multiple         Theria  20.467403  468.361   1826.202
 17252:          ZZEF1        other       Amniota   1.744677    9.323     93.429
 17253:           ZZZ3      nucleus Deuterostomia   4.360941   20.657     68.401
```

```
# more complex
# example:
# filter all HIST2 genes with a median expression >10TPM and localized to nucleus

data %>% dplyr::filter(str_detect(ApprovedSymbol, '^HIST2') &
                 F5_med>=10 & Localization == 'nucleus')
```

```
   ApprovedSymbol Localization     Taxon    F5_min    F5_med     F5_max
1:      HIST2H2AB      nucleus Eukaryota  9.958116    91.570   1240.587
```

18

```
2:        HIST2H2BF      nucleus Eukaryota   0.918399   21.239    457.095
3:        HIST2H4A       nucleus Eukaryota 170.420082 1224.638 19898.334
4:        HIST2H4B       nucleus Eukaryota 170.420082 1224.638 19898.334
```

6.2 dplyr::arrange

```
# arranges row order
# descending - highest on top ; ascending highest at the bottom

data=fread(file = 'connectome_1in.tsv', select = 1:4)
data %>% dplyr::arrange(ApprovedSymbol) # arranges characters in descending order
```

```
       ApprovedSymbol Localization         Taxon    F5_min
    1:           A1BG      secreted      Eutheria  0.263767
    2:           A1CF      multiple Deuterostomia  0.000000
    3:            A2M      multiple  Euteleostomi  0.000000
    4:          A2ML1      secreted  Euteleostomi  0.000000
    5:         A4GALT      multiple     Tetrapoda  0.000000
   ---
17259:         ZYG11A       nucleus  Euteleostomi  0.000000
17260:         ZYG11B       nucleus  Euteleostomi  4.388704
17261:            ZYX      multiple        Theria 20.467403
17262:          ZZEF1         other       Amniota  1.744677
17263:           ZZZ3       nucleus Deuterostomia  4.360941
```

```
data %>% dplyr::arrange(desc(ApprovedSymbol)) # arranges characters in descending order
```

```
       ApprovedSymbol Localization         Taxon    F5_min
    1:           ZZZ3       nucleus Deuterostomia  4.360941
    2:          ZZEF1         other       Amniota  1.744677
    3:            ZYX      multiple        Theria 20.467403
    4:         ZYG11B       nucleus  Euteleostomi  4.388704
    5:         ZYG11A       nucleus  Euteleostomi  0.000000
   ---
17259:         A4GALT      multiple     Tetrapoda  0.000000
17260:          A2ML1      secreted  Euteleostomi  0.000000
17261:            A2M      multiple  Euteleostomi  0.000000
17262:           A1CF      multiple Deuterostomia  0.000000
17263:           A1BG      secreted      Eutheria  0.263767
```

```
data %>% dplyr::arrange(F5_min) # arranges numbers in descending order
```

```
       ApprovedSymbol Localization          Taxon     F5_min
    1:           A1CF     multiple  Deuterostomia     0.0000
    2:            A2M     multiple    Euteleostomi     0.0000
    3:          A2ML1      secreted    Euteleostomi     0.0000
    4:         A4GALT     multiple       Tetrapoda     0.0000
    5:          A4GNT        other       Tetrapoda     0.0000
   ---
17259:          RPS24    cytoplasm       Eukaryota   331.8715
17260:          PCBP2     multiple   Opisthokonta   344.3184
17261:           CFL1     multiple       Eukaryota   390.6739
17262:         EIF4G2    cytoplasm  Deuterostomia   402.0313
17263:           ACTB     multiple       Eukaryota 1451.3885
```

```
data %>% dplyr::arrange(-F5_min) # arranges numbers in ascending order
```

```
       ApprovedSymbol Localization          Taxon     F5_min
    1:           ACTB     multiple       Eukaryota 1451.3885
    2:         EIF4G2    cytoplasm  Deuterostomia   402.0313
    3:           CFL1     multiple       Eukaryota   390.6739
    4:          PCBP2     multiple   Opisthokonta   344.3184
    5:          RPS24    cytoplasm       Eukaryota   331.8715
   ---
17259:          ZUFSP      nucleus    Euteleostomi     0.0000
17260:          ZWINT     multiple        Eutheria     0.0000
17261:           ZXDA      nucleus    Euteleostomi     0.0000
17262:           ZXDB      nucleus    Euteleostomi     0.0000
17263:         ZYG11A      nucleus    Euteleostomi     0.0000
```

```
data %>% dplyr::arrange(Localization, -F5_min) # can be combined; order matters
```

```
       ApprovedSymbol Localization          Taxon   F5_min
    1:         EIF4G2    cytoplasm  Deuterostomia 402.0313
    2:          RPS24    cytoplasm       Eukaryota 331.8715
    3:           EIF1    cytoplasm       Eukaryota 301.4819
    4:            FTL    cytoplasm          Theria 149.4635
    5:            UBC    cytoplasm       Eukaryota 139.8899
   ---
```

```
17259:            XYLT1      secreted      Bilateria   0.0000
17260:            ZG16B      secreted     Catarrhini   0.0000
17261:           ZNF365      secreted      Tetrapoda   0.0000
17262:             ZPBP      secreted        Amniota   0.0000
17263:            ZPBP2      secreted        Amniota   0.0000
```

6.3 dplyr::mutate / dplyr::case_when

```
# 'mutate': adds a new variable as a column

data=fread(file = 'connectome_1in.tsv', select = c(1,2,5))
data %>% dplyr::mutate(my_new_variable = 'FANTOM_5_data') # added as last column
```

```
       ApprovedSymbol Localization   F5_med my_new_variable
    1:           A1BG     secreted   16.943   FANTOM_5_data
    2:           A1CF     multiple    0.000   FANTOM_5_data
    3:            A2M     multiple    1.212   FANTOM_5_data
    4:          A2ML1     secreted    0.000   FANTOM_5_data
    5:         A4GALT     multiple    9.725   FANTOM_5_data
   ---
17259:         ZYG11A      nucleus    0.264   FANTOM_5_data
17260:         ZYG11B      nucleus   16.580   FANTOM_5_data
17261:            ZYX     multiple  468.361   FANTOM_5_data
17262:          ZZEF1        other    9.323   FANTOM_5_data
17263:           ZZZ3      nucleus   20.657   FANTOM_5_data
```

```
# 'case_when': now row values depend on the pattern

data %>% dplyr::mutate(expr_level = case_when(F5_med<10 ~'lowly-expressed gene',
                                        TRUE ~ 'highly-expressed gene'))
```

```
       ApprovedSymbol Localization  F5_med            expr_level
    1:           A1BG     secreted   16.943 highly-expressed gene
    2:           A1CF     multiple    0.000  lowly-expressed gene
    3:            A2M     multiple    1.212  lowly-expressed gene
    4:          A2ML1     secreted    0.000  lowly-expressed gene
    5:         A4GALT     multiple    9.725  lowly-expressed gene
   ---
17259:         ZYG11A      nucleus    0.264  lowly-expressed gene
```

```
17260:        ZYG11B      nucleus  16.580 highly-expressed gene
17261:           ZYX     multiple 468.361 highly-expressed gene
17262:          ZZEF1       other   9.323  lowly-expressed gene
17263:           ZZZ3      nucleus  20.657 highly-expressed gene
```

```r
data=fread(file = 'connectome_1in.tsv', select = c(1,2,4))
data %>% dplyr::mutate(expr_level =
                 case_when(F5_min>0 & Localization == 'secreted'~
                           'highly-expressed secreted protein',
                           TRUE ~ 'not so interesting')) %>%
       dplyr:: arrange(desc(Localization))
```

```
     ApprovedSymbol Localization   F5_min                        expr_level
  1:          A1BG     secreted 0.263767 highly-expressed secreted protein
  2:         A2ML1     secreted 0.000000                not so interesting
  3:        AADACL2     secreted 0.000000                not so interesting
  4:         ABHD12     secreted 5.284597 highly-expressed secreted protein
  5:        ABHD12B     secreted 0.000000                not so interesting
 ---
17259:        ZNF503     cytoplasm 0.000000                not so interesting
17260:        ZNF706     cytoplasm 6.065311                not so interesting
17261:         ZNFX1     cytoplasm 6.826872                not so interesting
17262:         ZNRF4     cytoplasm 0.000000                not so interesting
17263:        ZSWIM8     cytoplasm 1.970653                not so interesting
```

7 Combining 'dplyr' functions

Example:
1. find all highly expressed in CD14+ Monocytes (>10TPM) putative ligands (secreted genes)
2. order the genes from most highly expressed down, annotate them and save to a file
3. ApprovedSymbol (dplyr::rename to geneSymbol), Localization and the cell of interest
4. after excluding all 'orf' genes (ex: C1orf162, etc)

```
data=fread(file = 'connectome_1in.tsv')
data %>% dplyr::filter(Localization=='secreted' & `CD14+ Monocytes` > 10) %>%
        dplyr::arrange(-`CD14+ Monocytes`) %>%
        dplyr::select(ApprovedSymbol, Localization, `CD14+ Monocytes`) %>%
        dplyr::rename(geneSymbol=ApprovedSymbol) %>%
        dplyr::mutate(Class='Monocyte Ligand') %>%
        dplyr::filter(!str_detect(geneSymbol, 'orf'))
```

```
     geneSymbol Localization CD14+ Monocytes          Class
  1:        LYZ     secreted      13533.670 Monocyte Ligand
  2:        IL8     secreted      12834.926 Monocyte Ligand
  3:        B2M     secreted       7159.262 Monocyte Ligand
  4:       CCL3     secreted       3027.740 Monocyte Ligand
  5:      CXCL2     secreted       2427.316 Monocyte Ligand
 ---
169:        PF4     secreted         10.554 Monocyte Ligand
170:       NID1     secreted         10.457 Monocyte Ligand
171:      TCTN3     secreted         10.359 Monocyte Ligand
172:      FOLR3     secreted         10.122 Monocyte Ligand
173:      THBS3     secreted         10.065 Monocyte Ligand
```

8 Working with groupped data

```
-- Attaching packages --------------------------------------- tidyverse 1.3.2 --
v ggplot2 3.3.6      v purrr   0.3.4
v tibble  3.1.8      v dplyr   1.0.9
v tidyr   1.2.0      v stringr 1.4.0
v readr   2.1.2      v forcats 0.5.1
-- Conflicts ------------------------------------------ tidyverse_conflicts() --
x dplyr::between()   masks data.table::between()
x dplyr::filter()    masks stats::filter()
x dplyr::first()     masks data.table::first()
x dplyr::lag()       masks stats::lag()
x dplyr::last()      masks data.table::last()
x purrr::transpose() masks data.table::transpose()


         : 'magrittr'


       'package:purrr'         :

     set_names


       'package:tidyr'         :

     extract
```

8.1 dplyr::add_count / dplyr::count

```r
data = fread(file = 'connectome_1in.tsv')
data %>% dplyr::add_count(ApprovedSymbol) %>% dplyr::arrange(-n)
```

```
     ApprovedSymbol Localization          Taxon    F5_min  F5_med   F5_max
  1:           A1BG      secreted      Eutheria  0.263767  16.943 2377.989
```

```
   2:          A1CF    multiple Deuterostomia 0.000000    0.000  153.024
   3:           A2M    multiple   Euteleostomi 0.000000    1.212 1515.493
   4:         A2ML1    secreted   Euteleostomi 0.000000    0.000  206.355
   5:        A4GALT    multiple       Tetrapoda 0.000000    9.725   80.009
  ---
17259:       ZYG11A     nucleus   Euteleostomi 0.000000    0.264   11.104
17260:       ZYG11B     nucleus   Euteleostomi 4.388704   16.580   65.872
17261:          ZYX    multiple          Theria 20.467403 468.361 1826.202
17262:        ZZEF1       other         Amniota 1.744677    9.323   93.429
17263:         ZZZ3     nucleus Deuterostomia 4.360941   20.657   68.401
      Adipocyte Breast Adipocyte Omental Adipocyte Perirenal
   1:          36.847           35.174             49.723
   2:           0.000            0.000              0.000
   3:          90.272          121.423             50.596
   4:           0.000            0.000              0.000
   5:          24.386           10.811             17.447
  ---
17259:          0.091            0.000              0.000
17260:         30.164           25.141             18.319
17261:        473.814          442.646            389.935
17262:          9.713            8.754              1.745
17263:         18.076           20.905             19.191
      Adipocyte Subcutaneous Alveolar Epithelial Amniotic Epithelial
   1:               41.076             4.143               0.264
   2:                0.000             0.000               0.000
   3:               63.397             0.000               0.000
   4:                0.000             0.000              21.168
   5:               27.730            11.608              19.070
  ---
17259:                0.000             0.123               1.733
17260:               18.911             4.978              15.027
17261:              663.910          1012.799             495.516
17262:                9.769             6.561              12.725
17263:               11.938            10.109               6.001
      Amniotic Membrane Annulus Pulposus Astrocyte Cerebellum
   1:             4.915          31.099               16.943
   2:             0.000           0.000                0.000
   3:             0.819           0.000                1.364
   4:           206.355           0.000                0.000
   5:            30.631          42.725                6.146
  ---
17259:             0.396           0.000                1.195
17260:            21.576          17.402                7.212
```

```
17261:            81.118            638.182            597.896
17262:            12.789              8.083              9.124
17263:            25.542             26.437             10.307
      Astrocyte Cerebral Cortex Basophils Bronchial Epithelial
   1:                     7.498   17.908                  9.793
   2:                     0.000    0.103                  0.000
   3:                     8.824    0.128                  0.000
   4:                     0.000    0.619                  3.187
   5:                     0.655    0.334                 16.745
  ---
17259:                     1.115    0.128                  0.100
17260:                    27.474   14.586                  6.326
17261:                   251.642  440.993                641.173
17262:                     8.411   33.169                  9.038
17263:                    27.943   17.550                 13.487
      CD133+ Haematopoietic stem cell
   1:                         50.206
   2:                          0.000
   3:                          0.000
   4:                          0.000
   5:                          0.000
  ---
17259:                          0.000
17260:                          8.514
17261:                        116.266
17262:                         11.450
17263:                         31.122
      CD14+ Monocyte derived Endothelial progenitor CD14+ Monocytes
   1:                                         21.596          16.290
   2:                                          0.000           0.000
   3:                                        144.359           0.185
   4:                                          0.000           0.000
   5:                                          0.075           0.217
  ---
17259:                                          0.000           0.000
17260:                                         43.930          17.754
17261:                                        270.955         476.539
17262:                                         59.266          49.950
17263:                                         68.401          23.115
      CD14+CD16+ Monocytes CD14+CD16- Monocytes CD14-CD16+ Monocytes
   1:            16.181               23.447               10.928
   2:             0.000                0.000                0.000
   3:             0.141                0.050                0.127
```

```
   4:             0.000           0.000              0.000
   5:             0.000           0.037              0.000
   ---
17259:            0.000           0.000              0.000
17260:           15.776          20.266             16.119
17261:          359.648         580.106            329.579
17262:           19.920          23.946             30.071
17263:           15.776          17.156             13.928
      CD19+ B cells CD34+ Haematopoietic stem cell CD34+ progenitor
   1:       24.777                     23.968            34.219
   2:        0.000                      0.000             0.000
   3:        0.000                      1.317             0.280
   4:        0.000                      0.000             0.000
   5:        0.217                      2.370             0.280
   ---
17259:        1.290                      0.527             0.000
17260:       17.912                     14.750             6.768
17261:       26.779                     91.394           128.700
17262:       27.082                     12.379            26.067
17263:       35.293                     21.861            28.870
      CD4+ T cells CD4+CD25+CD45RA+ naive regulatory T cells
   1:      15.572                             13.057
   2:       0.000                              0.000
   3:       0.000                              0.000
   4:       0.000                              0.000
   5:       0.000                              0.000
   ---
17259:       0.262                              0.118
17260:      25.652                             16.660
17261:     156.419                             76.060
17262:      45.869                             23.929
17263:      51.262                             38.660
      CD4+CD25+CD45RA- memory regulatory T cells
   1:                              15.473
   2:                               0.000
   3:                               0.000
   4:                               0.000
   5:                               0.341
   ---
17259:                             0.062
17260:                            15.991
17261:                           154.997
17262:                            20.953
```

```
17263:                                  31.162
     CD4+CD25-CD45RA+ naive conventional T cells
   1:                                  16.530
   2:                                   0.000
   3:                                   0.000
   4:                                   0.000
   5:                                   0.000
   ---
17259:                                   0.101
17260:                                  20.752
17261:                                 111.837
17262:                                  26.076
17263:                                  40.911
     CD4+CD25-CD45RA- memory conventional T cells CD8+ T cells
   1:                                  17.132       25.467
   2:                                   0.000        0.000
   3:                                   0.000        0.000
   4:                                   0.000        0.000
   5:                                   0.028        0.000
   ---
17259:                                   0.152        0.149
17260:                                  15.595       17.183
17261:                                 115.866      130.069
17262:                                  24.326       29.854
17263:                                  34.308       42.319
     Cardiac Myocyte Chondroblast Chondrocyte Chorionic Membrane
   1:          8.595        9.231      17.668              7.071
   2:          0.000        0.000       0.000              0.000
   3:         17.161        4.198      28.522             30.713
   4:          0.000        0.000       0.000             18.666
   5:         12.142       31.842      19.033             27.820
   ---
17259:          0.309        0.241       0.000              0.153
17260:         22.829       18.431      23.253             18.167
17261:        432.366      341.704     127.448            287.212
17262:         10.602        9.115       9.840             11.830
17263:         24.197       20.446      22.374             21.327
     Ciliary Epithelial Corneal Epithelial
   1:              5.169           10.818
   2:              0.000            0.000
   3:              1.474            0.000
   4:              0.029           16.614
   5:              3.882           62.267
```

```
        ---
17259:              1.246             1.103
17260:             22.091             5.705
17261:            425.145           416.233
17262:              9.560             6.434
17263:             26.001             4.983
      Dendritic Monocyte Immature derived Dendritic Plasmacytoid ES cells
   1:                              19.102                  93.768   10.198
   2:                               0.000                   0.000    0.000
   3:                             622.499                   0.966    0.044
   4:                               0.000                   0.000    0.125
   5:                               3.082                   1.999    5.704
   ---
17259:                               0.044                   0.000   11.104
17260:                              30.346                   4.601   18.448
17261:                             345.054                 103.197  177.084
17262:                              22.906                  23.909    4.880
17263:                              17.190                  50.876   34.292
      Endothelial Aortic Endothelial Artery Endothelial Lymphatic
   1:            14.159             12.190                    2.268
   2:             0.000              0.000                    0.000
   3:             0.849              7.317                  112.864
   4:             0.000              0.000                    0.000
   5:            16.708             22.830                   17.733
   ---
17259:            0.514              0.737                    0.291
17260:           16.580             10.450                   25.263
17261:          760.351            951.401                  430.343
17262:           14.401             10.641                   16.899
17263:           12.937             11.703                   37.392
      Endothelial Microvascular Endothelial Thoracic
   1:                    15.821                   16.751
   2:                     0.000                    0.000
   3:                    22.919                    4.527
   4:                     0.000                    0.000
   5:                    48.154                   23.226
   ---
17259:                     0.760                    0.000
17260:                    14.217                   12.139
17261:                   678.548                  810.633
17262:                    12.125                   12.915
17263:                    12.092                   14.897
      Endothelial Umbilical Vein Endothelial Vein Eosinophils
```

```
     1:                          1.460              8.301            2.636
     2:                          0.000              0.000            0.000
     3:                          6.203              0.938            0.000
     4:                          0.000              0.000            0.000
     5:                          6.894              3.753            0.000
  ---
 17259:                          1.584              1.011            0.000
 17260:                         21.824             17.400           41.867
 17261:                        351.405            857.603          742.793
 17262:                         12.374             13.511           93.429
 17263:                         31.748             12.787            9.187
        Esophageal Epithelial Fibroblast Aortic Adventitial Fibroblast Cardiac
     1:                 8.286                          9.701               24.256
     2:                 0.000                          0.000                0.000
     3:                 0.436                          5.699               10.457
     4:                17.008                          0.000                0.000
     5:                18.752                         19.151               15.473
  ---
 17259:                 0.872                          0.988                0.465
 17260:                 7.414                         20.302               21.306
 17261:               593.088                        515.496              597.537
 17262:                 3.489                          8.258               10.805
 17263:                 4.361                         20.657               22.829
        Fibroblast Choroid Plexus Fibroblast Conjunctival Fibroblast Dermal
     1:                   35.255                      20.855                22.007
     2:                    0.000                       0.000                 0.000
     3:                    7.497                       0.234                 1.313
     4:                    0.000                       0.000                 0.095
     5:                   44.240                       7.264                 6.952
  ---
 17259:                    0.378                       1.875                 0.965
 17260:                   14.575                       9.139                14.446
 17261:                  452.713                     698.512               493.593
 17262:                   11.345                       6.561                 8.834
 17263:                   19.612                      19.917                29.152
        Fibroblast Gingival Fibroblast Lymphatic Fibroblast Periodontal Ligament
     1:                 26.186                34.565                            20.547
     2:                  0.000                 0.000                             0.000
     3:                  0.465                 0.000                             1.168
     4:                  0.000                 0.000                             0.000
     5:                 21.772                28.743                            13.828
  ---
 17259:                  0.414                 0.000                             0.613
```

```
17260:              26.905              5.509                   25.036
17261:             458.303           1757.244                  560.575
17262:              10.650              4.431                    9.916
17263:              36.798              8.407                   29.337
        Fibroblast Pulmonary Artery Fibroblast Skin Normal Gingival Epithelial
    1:                    15.881                 32.488                    5.443
    2:                     0.000                  0.000                    0.000
    3:                     0.340                 23.182                    0.064
    4:                     0.000                  0.000                   26.966
    5:                    12.364                 11.766                   11.276
   ---
17259:                     0.227                  0.235                    1.261
17260:                    14.860                 16.632                   17.275
17261:                   468.361                479.812                  359.155
17262:                     5.785                  8.612                    7.890
17263:                    12.931                 29.297                   17.675
        Granulocyte Macrophage progenitor Hair Follicle Dermal Papilla
    1:                            34.492                         35.204
    2:                             0.000                          0.000
    3:                             0.000                          9.592
    4:                             0.000                          0.000
    5:                             0.000                          6.254
   ---
17259:                             0.000                          0.460
17260:                            28.104                         13.679
17261:                           226.111                        941.256
17262:                            77.925                         13.769
17263:                            22.994                         26.221
        Hair Follicle Outer Root Sheath Hepatic Sinusoidal Endothelial
    1:                            50.829                         11.219
    2:                             0.000                          0.000
    3:                            69.227                         36.612
    4:                             0.000                          0.000
    5:                            26.718                         13.360
   ---
17259:                             0.000                          1.070
17260:                            18.619                         10.445
17261:                           492.240                        894.871
17262:                            10.575                          9.780
17263:                            21.046                         13.323
        Hepatic Stellate Hepatocyte Immature Langerhans
    1:           26.137   2377.989                 46.017
    2:            0.000    153.024                  0.000
```

```
   3:            1.235    1515.493               9.836
   4:            0.000       0.000              10.474
   5:            8.997       0.000               0.000
   ---
17259:           0.000       4.341               0.000
17260:          12.112      33.607               5.106
17261:        1114.881      94.638             161.451
17262:           8.467      16.168              20.296
17263:           7.909      42.797              27.800
     Intestinal Epithelial (polarized) Iris Pigment Epithelial
   1:                          0.984                    7.206
   2:                         53.254                    0.000
   3:                          9.593                    0.172
   4:                          0.000                    0.000
   5:                          6.026                    4.975
   ---
17259:                         1.599                    0.515
17260:                        16.111                   15.098
17261:                       118.807                  471.983
17262:                         7.379                    5.319
17263:                        15.988                   19.730
     Keratinocyte Epidermal Keratinocyte Oral Keratocytes Lens Epithelial
   1:                 4.995             1.852      31.620          12.195
   2:                 0.000             0.000       0.000           0.000
   3:                 0.034             0.000       0.000           0.663
   4:                20.528            13.148       0.000           0.000
   5:                 2.383            11.574       9.466           6.038
   ---
17259:                0.488             1.204       1.497           0.676
17260:                9.766            21.388       6.804          19.175
17261:              473.740           258.600     535.837         469.660
17262:                9.513             5.185      10.281           8.082
17263:               16.939            29.165      12.492          20.688
     Macrophage Monocyte derived Mallassez derived Mammary Epithelial
   1:                    17.923             9.523               4.266
   2:                     0.000             0.000               0.000
   3:                   208.304             0.000               2.494
   4:                     0.000             5.063               1.235
   5:                     0.000            14.173               1.850
   ---
17259:                    0.000             0.794               0.765
17260:                   29.538            11.063              11.799
17261:                  544.167           391.202             360.113
```

```
17262:                         21.500            6.146             7.575
17263:                         24.682           19.728            27.103
       Mast cells Mast cells stimulated Mature Adipocyte Melanocyte Meningeal
   1:     21.826                 11.824            1.629     24.663     6.175
   2:      0.000                  0.000            0.000      0.000     0.000
   3:      1.178                  0.000           85.879    423.667     2.797
   4:      0.496                  0.000            0.000      0.209     0.000
   5:      0.883                  0.000            5.599      0.575     7.811
  ---
17259:      0.119                  0.000            0.300      0.554     1.516
17260:     23.254                  8.868           38.571     15.233    14.621
17261:    130.957                205.441           75.078    176.569   725.339
17262:     24.955                  4.434           17.865     11.846     8.827
17263:     26.307                 38.920           14.656     16.781    16.446
       Mesenchymal precursor Adipose Mesenchymal precursor Bone Marrow
   1:                    19.417                              15.766
   2:                     0.000                               0.000
   3:                     3.791                               4.242
   4:                     0.000                               0.000
   5:                    10.586                              23.057
  ---
17259:                     0.412                               0.063
17260:                    10.694                              15.663
17261:                   404.482                             398.606
17262:                     6.075                               7.430
17263:                    26.862                              22.748
       Mesenchymal precursor Cardiac Mesenchymal stem cell Adipose
   1:                    16.772                             123.496
   2:                     0.000                               0.000
   3:                     2.287                              42.790
   4:                     0.000                               0.000
   5:                    22.477                              80.009
  ---
17259:                     0.687                               0.109
17260:                    12.000                              14.489
17261:                   281.751                            1794.597
17262:                     6.228                               8.004
17263:                    23.183                              16.787
       Mesenchymal stem cell Amniotic Membrane
   1:                              21.221
   2:                               0.000
   3:                              64.543
   4:                               0.000
```

```
   5:                                           6.722
   ---
17259:                                          0.653
17260:                                         21.727
17261:                                        805.240
17262:                                          6.710
17263:                                         25.101
        Mesenchymal stem cell Bone Marrow Mesenchymal stem cell Hepatic
   1:                           10.101                            25.629
   2:                            0.000                             0.000
   3:                            1.121                             5.059
   4:                            0.000                             0.000
   5:                            4.531                             4.491
   ---
17259:                           1.463                             1.818
17260:                          12.460                            28.307
17261:                         739.854                           348.455
17262:                           8.035                             8.266
17263:                          10.272                            27.484
        Mesenchymal stem cell Umbilical Mesenchymal stem cell Wharton's Jelly
   1:                           50.172                            28.828
   2:                            0.000                             0.000
   3:                            0.000                             0.129
   4:                            0.000                             0.000
   5:                           13.804                             4.395
   ---
17259:                           0.264                             0.259
17260:                          10.866                            18.486
17261:                         485.315                           305.986
17262:                           7.237                             7.756
17263:                          10.922                            19.520
        Mesothelial Migratory Langerhans
   1:      34.815           22.289
   2:       0.000            0.000
   3:      91.831          100.280
   4:       0.000            0.065
   5:       9.413            2.969
   ---
17259:     0.067            0.000
17260:    17.012            8.426
17261:   471.516          641.794
17262:     8.412           16.995
17263:    13.762           26.937
```

	Multipot. Cord Blood Unrestrict. Somatic stem cell	Myoblast	Myotube
1:	15.398	24.219	24.929
2:	0.000	0.000	0.000
3:	0.000	1.243	29.737
4:	0.000	0.000	0.000
5:	4.718	12.731	8.903

17259:	0.694	0.176	0.267
17260:	22.414	22.962	24.039
17261:	541.254	368.368	462.343
17262:	7.854	9.323	8.992
17263:	26.099	24.100	22.347

	NK cells	Nasal Epithelial	Neural stem cell	Neurons	Neutrophils
1:	10.954	10.309	6.114	6.177	0.802
2:	0.000	0.000	0.000	0.000	0.000
3:	0.066	0.311	96.416	22.745	0.000
4:	0.000	21.389	0.254	0.000	0.000
5:	0.171	9.087	0.127	0.223	0.000

17259:	0.229	0.140	0.905	0.000	0.000
17260:	16.003	14.102	45.513	32.986	65.872
17261:	150.932	224.852	348.359	176.113	1826.202
17262:	43.102	12.315	13.429	12.432	78.486
17263:	39.510	23.803	32.353	28.234	11.961

	Nucleus Pulposus	Olfactory Epithelial	Osteoblast	Osteocyte
1:	28.643	28.786	2.784	27.879
2:	0.000	0.000	0.000	0.000
3:	1.212	2.805	0.849	16.605
4:	0.000	0.000	0.000	0.000
5:	34.574	3.817	1.863	27.933

17259:	0.000	0.717	0.184	0.000
17260:	21.465	29.622	22.979	15.396
17261:	807.148	589.053	598.789	447.318
17262:	7.636	11.644	10.015	12.868
17263:	28.063	26.305	22.024	18.103

	Pancreatic Stromal	Pericytes	Placental Epithelial	Preadipocyte Breast
1:	25.804	81.360	2.573	26.514
2:	0.000	0.000	0.000	0.000
3:	6.343	1.740	0.101	17.948
4:	0.000	0.000	18.928	0.000
5:	9.818	13.108	8.783	23.957

```
17259:            0.695     0.098            0.202            0.140
17260:           27.281     4.923           13.618           23.652
17261:          504.272   790.225          427.509          494.841
17262:            9.210     8.964            8.024            6.965
17263:           29.106     9.846           11.451           24.937
      Preadipocyte Omental Preadipocyte Perirenal Preadipocyte Subcutaneous
   1:             31.362                 57.883                     31.618
   2:              0.000                  0.000                      0.000
   3:             20.348                  7.421                      3.173
   4:              0.000                  0.000                      0.000
   5:              8.280                 17.810                      6.012
  ---
17259:             0.428                  0.000                      0.397
17260:            28.414                 14.842                     20.294
17261:           485.452                552.115                    647.132
17262:             7.652                  4.453                      6.832
17263:            21.306                 11.873                     16.432
      Preadipocyte Visceral Prostate Epithelial Prostate Stromal
   1:             17.867                  3.093           16.122
   2:              0.000                  7.816            0.000
   3:              0.622                  0.138           13.502
   4:              0.000                 15.235            0.000
   5:             17.783                 13.987           10.854
  ---
17259:             0.570                  0.242            1.390
17260:            20.144                 16.644           19.337
17261:           504.228                313.816          457.992
17262:             8.435                  7.171            8.223
17263:            25.769                 19.604           20.297
      Renal Cortical Epithelial Renal Epithelial Renal Glomerular Endothelial
   1:                     3.642            8.393                        4.962
   2:                     0.000            0.000                        0.000
   3:                     0.000            0.000                       47.414
   4:                     0.039            0.000                        0.000
   5:                    12.396           31.502                       36.860
  ---
17259:                     1.133            2.336                        1.272
17260:                    11.790           12.459                        9.821
17261:                   339.622          581.510                      826.222
17262:                     8.704            4.548                       13.692
17263:                    16.972           16.062                       14.051
      Renal Mesangial Renal Proximal Tubular Epithelial Reticulocyte
   1:             19.607                          4.602            1.113
```

```
    2:         0.000                            0.000        0.000
    3:         0.251                            0.000        0.131
    4:         0.000                            0.000        0.000
    5:        36.755                           26.727       41.660
   ---
17259:         0.522                            2.278        0.000
17260:         7.441                           10.504        9.483
17261:       925.429                          314.538       20.467
17262:         7.887                            6.491       14.248
17263:         8.118                           11.819       15.576
       Retinal Pigment Epithelial Salivary Acinar Sebocyte Sertoli
    1:                      4.445              8.125    9.870   3.214
    2:                      0.000              0.000    0.000   0.000
    3:                      0.194              2.552    0.000  35.350
    4:                      0.000             35.931   24.871   0.000
    5:                      4.433             21.075    5.033  27.316
   ---
17259:                      1.083              0.069    1.912   0.000
17260:                     16.906             12.555    9.280  25.709
17261:                    369.460            268.938 1059.838 393.676
17262:                      8.508             18.132   11.290  19.282
17263:                     22.610             33.129    8.857  17.675
       Skeletal Muscle Skeletal Muscle Satellite Small Airway Epithelial
    1:        12.024                 14.713                        3.268
    2:         0.000                  0.000                        0.000
    3:         5.353                  4.979                        0.132
    4:         0.000                  0.000                        4.941
    5:         9.673                 13.759                       14.181
   ---
17259:         0.487                  0.227                        0.055
17260:        15.615                 24.160                       12.988
17261:       444.258                468.990                      406.527
17262:         8.857                  9.533                        6.950
17263:        28.755                 34.142                       27.946
       Smooth Muscle Aortic Smooth Muscle Brachiocephalic
    1:               40.748                         45.693
    2:                0.000                          0.000
    3:                0.174                          0.093
    4:                0.000                          0.000
    5:               16.562                         17.797
   ---
17259:                1.135                          0.000
17260:               11.313                         14.311
```

```
17261:                    744.286                        547.700
17262:                     10.759                          8.519
17263:                      9.526                         17.754
        Smooth Muscle Brain Vascular Smooth Muscle Bronchial
    1:                     25.877                         33.998
    2:                      0.000                          0.000
    3:                      3.687                         22.458
    4:                      0.000                          0.000
    5:                      4.468                         13.933
   ---
17259:                      0.890                          0.726
17260:                     24.562                         11.288
17261:                    414.742                       1125.933
17262:                     10.198                          8.764
17263:                     27.923                         18.051
        Smooth Muscle Carotid Smooth Muscle Colonic
    1:                     39.639                         15.752
    2:                      0.000                          0.000
    3:                      0.159                         10.153
    4:                      0.000                          0.000
    5:                     19.114                          7.930
   ---
17259:                      0.000                          0.331
17260:                     12.293                         21.923
17261:                    970.678                        358.749
17262:                      9.186                          9.284
17263:                     13.333                         28.665
        Smooth Muscle Coronary Artery Smooth Muscle Esophageal
    1:                    100.375                         21.456
    2:                      0.000                          0.000
    3:                      0.669                          0.975
    4:                      0.000                          0.000
    5:                     19.340                         15.604
   ---
17259:                      0.202                          2.926
17260:                      8.915                          4.389
17261:                    732.392                       1411.700
17262:                     14.216                          6.339
17263:                      8.280                          8.777
        Smooth Muscle Internal Thoracic Artery Smooth Muscle Prostate
    1:                              43.965                        21.860
    2:                               0.000                         0.000
    3:                               0.122                         6.740
```

```
    4:                                        0.000                     0.000
    5:                                       31.124                     9.725
    ---
17259:                                        0.244                     1.311
17260:                                        8.531                    17.391
17261:                                      534.323                   374.423
17262:                                       10.441                     9.558
17263:                                       13.609                    24.070
       Smooth Muscle Pulmonary Artery Smooth Muscle Subclavian Artery
    1:                               48.593                           18.313
    2:                                0.000                            0.000
    3:                                2.260                            0.146
    4:                                0.000                            0.000
    5:                               30.512                           25.372
    ---
17259:                                0.000                            0.146
17260:                               16.951                           11.053
17261:                              679.179                          688.957
17262:                                6.780                           10.125
17263:                                6.780                           11.370
       Smooth Muscle Tracheal Smooth Muscle Umbilical Artery
    1:                         26.873                          34.897
    2:                          0.000                           0.000
    3:                         37.598                           2.324
    4:                          0.000                           0.000
    5:                         24.153                           8.914
    ---
17259:                          0.328                           0.237
17260:                         20.024                          10.054
17261:                        740.403                         977.311
17262:                          6.578                           9.201
17263:                         15.163                           9.848
       Smooth Muscle Umbilical Vein Smooth Muscle Uterine Synoviocyte Tenocyte
    1:                           13.021                       18.802      9.504   18.775
    2:                            0.000                        0.000      0.000    0.000
    3:                            0.143                        3.193      2.208    5.266
    4:                            0.000                        0.000      0.000    0.000
    5:                            1.855                        2.129     65.447   61.794
    ---
17259:                            0.430                        1.064      0.481    0.000
17260:                           17.677                       18.448     21.608   24.898
17261:                          510.591                      433.518    467.281  489.985
17262:                            5.382                        9.933      7.342    6.676
```

```
17263:                        23.406              24.833      25.040   22.721
        Trabecular Meshwork Tracheal Epithelial Urothelial n
    1:              30.337              10.808      2.750 1
    2:               0.000               0.000      0.000 1
    3:               0.000               0.223      0.000 1
    4:               0.000               4.885     10.780 1
    5:               6.097              11.199     26.715 1
   ---
17259:               2.393               0.096      0.000 1
17260:              10.372               7.905     11.614 1
17261:            1313.519             616.760    540.465 1
17262:               8.898               6.903      6.794 1
17263:              19.108              11.410     10.543 1
```

References