

# main

July 10, 2021

## 1 Gene term enrichment analysis

```
[1]: suppressMessages({  
      library(clusterProfiler)  
      library(org.Hs.eg.db)  
      library(enrichplot)  
      library(pathview)})  
  
[2]: extract_geneList <- function(dft, keytype='EntrezID'){  
      dft = na.exclude(dft)  
      gg = as.vector(dft[, 'logFC'])  
      names(gg) = as.vector(dft[, keytype])  
      gg = sort(gg, decreasing=TRUE)  
      return(gg)  
}  
  
      extract_geneSet <- function(dft, keytype='EntrezID'){  
      dft = na.exclude(dft)  
      gg = subset(dft, adj.P.Val <= 0.05)  
      gene = as.vector(gg[, keytype])  
      return(gene)  
}  
  
[3]: run_pathview <- function(geneList, pathwayID){  
      pathview(gene.data = geneList, pathway.id = pathwayID,  
              species = 'hsa', limit = list(gene=0.5, cpd=0.5))  
}
```

### 1.1 Genes

```
[4]: inputfile = "../_m/genes/diffExpr_maleVfemale_full.txt"  
  
      zz = read.delim(inputfile)  
      geneList1 = extract_geneList(zz)
```

```
[5]: aa_uniprot <- merge(bitr(names(geneList1), fromType="ENTREZID",
                             toType=c("UNIPROT", "ENSEMBL"), OrgDb="org.Hs.eg.db"),
                        zz, by.x='ENSEMBL', by.y='ensemblID')
aa_kegg <- merge(bitr_kegg(aa_uniprot[['UNIPROT']], fromType='uniprot',
                             toType='kegg', organism='hsa'),
                 aa_uniprot, by.x='uniprot', by.y='UNIPROT')
```

'select()' returned 1:many mapping between keys and columns

Warning message in bitr(names(geneList1), fromType = "ENTREZID", toType =  
c("UNIPROT", :  
"10.52% of input gene IDs are fail to map..."  
Reading KEGG annotation online:

Warning message in bitr\_kegg(aa\_uniprot[["UNIPROT"]], fromType = "uniprot",  
toType = "kegg", :  
"24.38% of input gene IDs are fail to map..."

```
[6]: geneList2 = extract_geneList(aa_kegg, 'kegg')
gene2 = extract_geneSet(aa_kegg, 'kegg')
```

### 1.1.1 Pathview plotting

```
[7]: run_pathview(geneList2, "hsa04020")
```

Info: Downloading xml files for hsa04020, 1/1 pathways..

Info: Downloading png files for hsa04020, 1/1 pathways..

'select()' returned 1:1 mapping between keys and columns

Info: Working in directory /ceph/users/jbenja13/github\_projects/sex\_differences\_  
sz/differential\_expression/caudate/dopamine\_system/\_m

Info: Writing image file hsa04020.pathview.png

```
[8]: #run_pathview(geneList2, "hsa04724")
#run_pathview(geneList2, "hsa04725")
#run_pathview(geneList2, "hsa04726")
#run_pathview(geneList2, "hsa04727")
run_pathview(geneList2, "hsa04728")
```

Info: Downloading xml files for hsa04728, 1/1 pathways..

Info: Downloading png files for hsa04728, 1/1 pathways..

'select()' returned 1:1 mapping between keys and columns

Info: Working in directory /ceph/users/jbenja13/github\_projects/sex\_differences\_sz/differential\_expression/caudate/dopamine\_system/\_m

Info: Writing image file hsa04728.pathview.png

```
[9]: #run_pathview(geneList2, "hsa04010")
run_pathview(geneList2, "hsa04713")
#run_pathview(geneList2, "hsa04927")
#run_pathview(geneList2, "hsa04710")
```

Info: Downloading xml files for hsa04713, 1/1 pathways..

Info: Downloading png files for hsa04713, 1/1 pathways..

'select()' returned 1:1 mapping between keys and columns

Info: Working in directory /ceph/users/jbenja13/github\_projects/sex\_differences\_sz/differential\_expression/caudate/dopamine\_system/\_m

Info: Writing image file hsa04713.pathview.png

## 1.2 Session Info

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

```
[1] "2021-07-10 10:27:29 EDT"
```

```
   user  system elapsed
18.891   1.688   27.501
```

Session info

| setting | value |
|---------|-------|
|---------|-------|

|         |                              |
|---------|------------------------------|
| version | R version 4.0.3 (2020-10-10) |
|---------|------------------------------|

|    |            |
|----|------------|
| os | Arch Linux |
|----|------------|

|        |                   |
|--------|-------------------|
| system | x86_64, linux-gnu |
|--------|-------------------|

|    |     |
|----|-----|
| ui | X11 |
|----|-----|

|          |      |
|----------|------|
| language | (EN) |
|----------|------|

|         |             |
|---------|-------------|
| collate | en_US.UTF-8 |
|---------|-------------|

|       |             |
|-------|-------------|
| ctype | en_US.UTF-8 |
|-------|-------------|

|    |                  |
|----|------------------|
| tz | America/New_York |
|----|------------------|

|      |            |
|------|------------|
| date | 2021-07-10 |
|------|------------|

Packages

| package | * version | date | lib | source |
|---------|-----------|------|-----|--------|
|---------|-----------|------|-----|--------|

|                 |          |            |     |                |
|-----------------|----------|------------|-----|----------------|
| AnnotationDbi   | * 1.52.0 | 2020-10-27 | [1] | Bioconductor   |
| assertthat      | 0.2.1    | 2019-03-21 | [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   |
| BiocManager     | 1.30.16  | 2021-06-15 | [1] | CRAN (R 4.0.3) |
| BiocParallel    | 1.24.1   | 2020-11-06 | [1] | Bioconductor   |
| Biostrings      | 2.58.0   | 2020-10-27 | [1] | Bioconductor   |
| bit             | 4.0.4    | 2020-08-04 | [1] | CRAN (R 4.0.2) |
| bit64           | 4.0.5    | 2020-08-30 | [1] | CRAN (R 4.0.2) |
| bitops          | 1.0-7    | 2021-04-24 | [1] | CRAN (R 4.0.3) |
| blob            | 1.2.1    | 2020-01-20 | [1] | CRAN (R 4.0.2) |
| cachem          | 1.0.5    | 2021-05-15 | [1] | CRAN (R 4.0.3) |
| cli             | 3.0.0    | 2021-06-30 | [1] | CRAN (R 4.0.3) |
| clusterProfiler | * 3.18.1 | 2021-02-09 | [1] | Bioconductor   |
| colorspace      | 2.0-2    | 2021-06-24 | [1] | CRAN (R 4.0.3) |
| cowplot         | 1.1.1    | 2020-12-30 | [1] | CRAN (R 4.0.2) |
| crayon          | 1.4.1    | 2021-02-08 | [1] | CRAN (R 4.0.3) |
| data.table      | 1.14.0   | 2021-02-21 | [1] | CRAN (R 4.0.3) |
| DBI             | 1.1.1    | 2021-01-15 | [1] | CRAN (R 4.0.2) |
| digest          | 0.6.27   | 2020-10-24 | [1] | CRAN (R 4.0.2) |
| DO.db           | 2.9      | 2021-04-08 | [1] | Bioconductor   |
| DOSE            | 3.16.0   | 2020-10-27 | [1] | Bioconductor   |
| downloader      | 0.4      | 2015-07-09 | [1] | CRAN (R 4.0.2) |
| dplyr           | 1.0.7    | 2021-06-18 | [1] | CRAN (R 4.0.3) |
| ellipsis        | 0.3.2    | 2021-04-29 | [1] | CRAN (R 4.0.3) |
| enrichplot      | * 1.10.2 | 2021-01-28 | [1] | Bioconductor   |
| 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) |
| fastmatch       | 1.1-0    | 2017-01-28 | [1] | CRAN (R 4.0.3) |
| fgsea           | 1.16.0   | 2020-10-27 | [1] | Bioconductor   |
| generics        | 0.1.0    | 2020-10-31 | [1] | CRAN (R 4.0.2) |
| ggforce         | 0.3.3    | 2021-03-05 | [1] | CRAN (R 4.0.3) |
| ggplot2         | 3.3.5    | 2021-06-25 | [1] | CRAN (R 4.0.3) |
| ggraph          | 2.0.5    | 2021-02-23 | [1] | CRAN (R 4.0.3) |
| ggrepel         | 0.9.1    | 2021-01-15 | [1] | CRAN (R 4.0.2) |
| glue            | 1.4.2    | 2020-08-27 | [1] | CRAN (R 4.0.2) |
| GO.db           | 3.12.1   | 2021-04-08 | [1] | Bioconductor   |
| GOSemSim        | 2.16.1   | 2020-10-29 | [1] | Bioconductor   |
| graph           | 1.68.0   | 2020-10-27 | [1] | Bioconductor   |
| graphlayouts    | 0.7.1    | 2020-10-26 | [1] | CRAN (R 4.0.3) |
| gridExtra       | 2.3      | 2017-09-09 | [1] | CRAN (R 4.0.2) |
| gtable          | 0.3.0    | 2019-03-25 | [1] | CRAN (R 4.0.2) |
| htmltools       | 0.5.1.1  | 2021-01-22 | [1] | CRAN (R 4.0.2) |
| httr            | 1.4.2    | 2020-07-20 | [1] | CRAN (R 4.0.2) |
| igraph          | 1.2.6    | 2020-10-06 | [1] | CRAN (R 4.0.2) |

|              |          |            |     |                |
|--------------|----------|------------|-----|----------------|
| IRanges      | * 2.24.1 | 2020-12-12 | [1] | Bioconductor   |
| IRdisplay    | 1.0      | 2021-01-20 | [1] | CRAN (R 4.0.2) |
| IRkernel     | 1.2      | 2021-05-11 | [1] | CRAN (R 4.0.3) |
| jsonlite     | 1.7.2    | 2020-12-09 | [1] | CRAN (R 4.0.2) |
| KEGGgraph    | 1.50.0   | 2020-10-27 | [1] | Bioconductor   |
| KEGGREST     | 1.30.1   | 2020-11-23 | [1] | Bioconductor   |
| lattice      | 0.20-41  | 2020-04-02 | [2] | CRAN (R 4.0.3) |
| lifecycle    | 1.0.0    | 2021-02-15 | [1] | CRAN (R 4.0.3) |
| magrittr     | 2.0.1    | 2020-11-17 | [1] | CRAN (R 4.0.2) |
| MASS         | 7.3-53   | 2020-09-09 | [2] | CRAN (R 4.0.3) |
| Matrix       | 1.3-4    | 2021-06-01 | [1] | CRAN (R 4.0.3) |
| memoise      | 2.0.0    | 2021-01-26 | [1] | CRAN (R 4.0.2) |
| munsell      | 0.5.0    | 2018-06-12 | [1] | CRAN (R 4.0.2) |
| org.Hs.eg.db | * 3.12.0 | 2021-02-24 | [1] | Bioconductor   |
| pathview     | * 1.30.1 | 2020-12-10 | [1] | Bioconductor   |
| pbdZMQ       | 0.3-5    | 2021-02-10 | [1] | CRAN (R 4.0.3) |
| pillar       | 1.6.1    | 2021-05-16 | [1] | CRAN (R 4.0.3) |
| pkgconfig    | 2.0.3    | 2019-09-22 | [1] | CRAN (R 4.0.2) |
| plyr         | 1.8.6    | 2020-03-03 | [1] | CRAN (R 4.0.2) |
| png          | 0.1-7    | 2013-12-03 | [1] | CRAN (R 4.0.2) |
| polyclip     | 1.10-0   | 2019-03-14 | [1] | CRAN (R 4.0.2) |
| purrr        | 0.3.4    | 2020-04-17 | [1] | CRAN (R 4.0.2) |
| qvalue       | 2.22.0   | 2020-10-27 | [1] | Bioconductor   |
| R6           | 2.5.0    | 2020-10-28 | [1] | CRAN (R 4.0.2) |
| RColorBrewer | 1.1-2    | 2014-12-07 | [1] | CRAN (R 4.0.2) |
| Rcpp         | 1.0.7    | 2021-07-07 | [1] | CRAN (R 4.0.3) |
| RCurl        | 1.98-1.3 | 2021-03-16 | [1] | CRAN (R 4.0.3) |
| repr         | 1.1.3    | 2021-01-21 | [1] | CRAN (R 4.0.2) |
| reshape2     | 1.4.4    | 2020-04-09 | [1] | CRAN (R 4.0.2) |
| Rgraphviz    | 2.34.0   | 2020-10-27 | [1] | Bioconductor   |
| rlang        | 0.4.11   | 2021-04-30 | [1] | CRAN (R 4.0.3) |
| RSQLite      | 2.2.7    | 2021-04-22 | [1] | CRAN (R 4.0.3) |
| rvcheck      | 0.1.8    | 2020-03-01 | [1] | CRAN (R 4.0.3) |
| S4Vectors    | * 0.28.1 | 2020-12-09 | [1] | Bioconductor   |
| scales       | 1.1.1    | 2020-05-11 | [1] | CRAN (R 4.0.2) |
| scatterpie   | 0.1.6    | 2021-04-23 | [1] | CRAN (R 4.0.3) |
| sessioninfo  | 1.1.1    | 2018-11-05 | [1] | CRAN (R 4.0.2) |
| shadowtext   | 0.0.8    | 2021-04-23 | [1] | CRAN (R 4.0.3) |
| stringi      | 1.6.2    | 2021-05-17 | [1] | CRAN (R 4.0.3) |
| stringr      | 1.4.0    | 2019-02-10 | [1] | CRAN (R 4.0.2) |
| tibble       | 3.1.2    | 2021-05-16 | [1] | CRAN (R 4.0.3) |
| tidygraph    | 1.2.0    | 2020-05-12 | [1] | CRAN (R 4.0.3) |
| tidyr        | 1.1.3    | 2021-03-03 | [1] | CRAN (R 4.0.3) |
| tidyselect   | 1.1.1    | 2021-04-30 | [1] | CRAN (R 4.0.3) |
| tweenr       | 1.0.2    | 2021-03-23 | [1] | CRAN (R 4.0.3) |
| utf8         | 1.2.1    | 2021-03-12 | [1] | CRAN (R 4.0.3) |
| uuid         | 0.1-4    | 2020-02-26 | [1] | CRAN (R 4.0.2) |
| vctrs        | 0.3.8    | 2021-04-29 | [1] | CRAN (R 4.0.3) |

|             |          |            |     |                |
|-------------|----------|------------|-----|----------------|
| viridis     | 0.6.1    | 2021-05-11 | [1] | CRAN (R 4.0.3) |
| viridisLite | 0.4.0    | 2021-04-13 | [1] | CRAN (R 4.0.3) |
| withr       | 2.4.2    | 2021-04-18 | [1] | CRAN (R 4.0.3) |
| XML         | 3.99-0.6 | 2021-03-16 | [1] | CRAN (R 4.0.3) |
| XVector     | 0.30.0   | 2020-10-27 | [1] | Bioconductor   |
| zlibbioc    | 1.36.0   | 2020-10-27 | [1] | Bioconductor   |

[1] /home/jbenja13/R/x86\_64-pc-linux-gnu-library/4.0  
 [2] /usr/lib/R/library