

# GSEA output

Abinet Joan

2023-09-29 13:43:20 +0200

## Contents

Loading Package	1
Modify dataframe format	1
Creating Dotplot	1
BP results . . . . .	1
CC results . . . . .	2
MF results . . . . .	3
KEGG results . . . . .	4

## Loading Package

```
suppressMessages(library(ggplot2))
suppressMessages(library("forcats"))
```

## Modify dataframe format

```
modify_GSEA_data <- function(GSEA_results){
  GSEA_results<- GSEA_results[order(GSEA_results$NES, decreasing = T),]
  GSEA_results <- GSEA_results[1:20,]

  # define the column has a factor so that the order of Ontology is maintain in the plot.
  GSEA_results$NAME <- factor(GSEA_results$NAME, levels = GSEA_results$NAME)
  GSEA_results$NAME <- fct_rev(GSEA_results$NAME)

  GSEA_results
}
```

## Creating Dotplot

### BP results

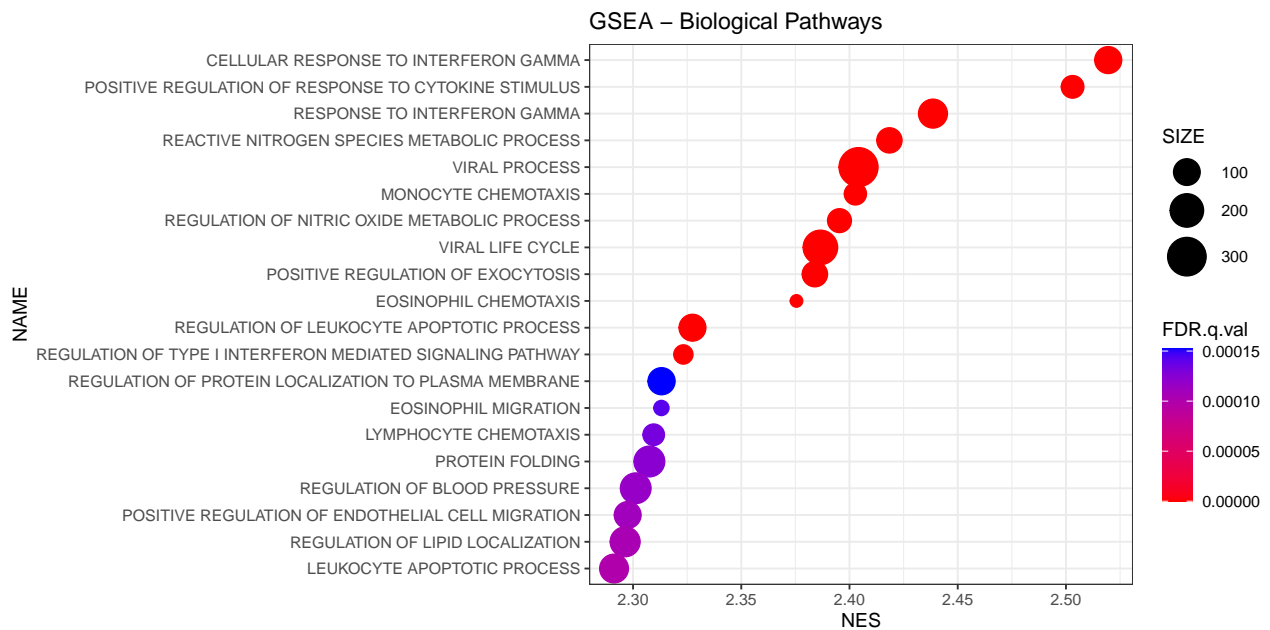
```
dir <- "GSEA_Output/"
GSEA_BP_res <-
read.table(paste0(dir,"my_analysis.Gsea.1675247466167_BP/gsea_report_for_2_1675247466167.tsv"),
header = T, sep = "\t")
```

```
GSEA_BP_res <- modify_GSEA_data(GSEA_BP_res)

GSEA_BP_res$NAME <- substring(GSEA_BP_res$NAME, 6)
GSEA_BP_res$NAME <- gsub("_", " ", GSEA_BP_res$NAME)

GSEA_BP_res$NAME <- fct_rev(factor(GSEA_BP_res$NAME, levels = GSEA_BP_res$NAME))

ggplot(GSEA_BP_res, aes(x = NES, y = NAME, size = SIZE, color = FDR.q.val)) + geom_point()
+ ggtitle("GSEA - Biological Pathways") + scale_color_gradient(low = "red", high =
"blue") + theme_bw() + scale_size(range = c(3,10))
```



```
#ggsave("plot_result/GSEA_results_BP_céci2.pdf", width = 12, height = 8)
```

## CC results

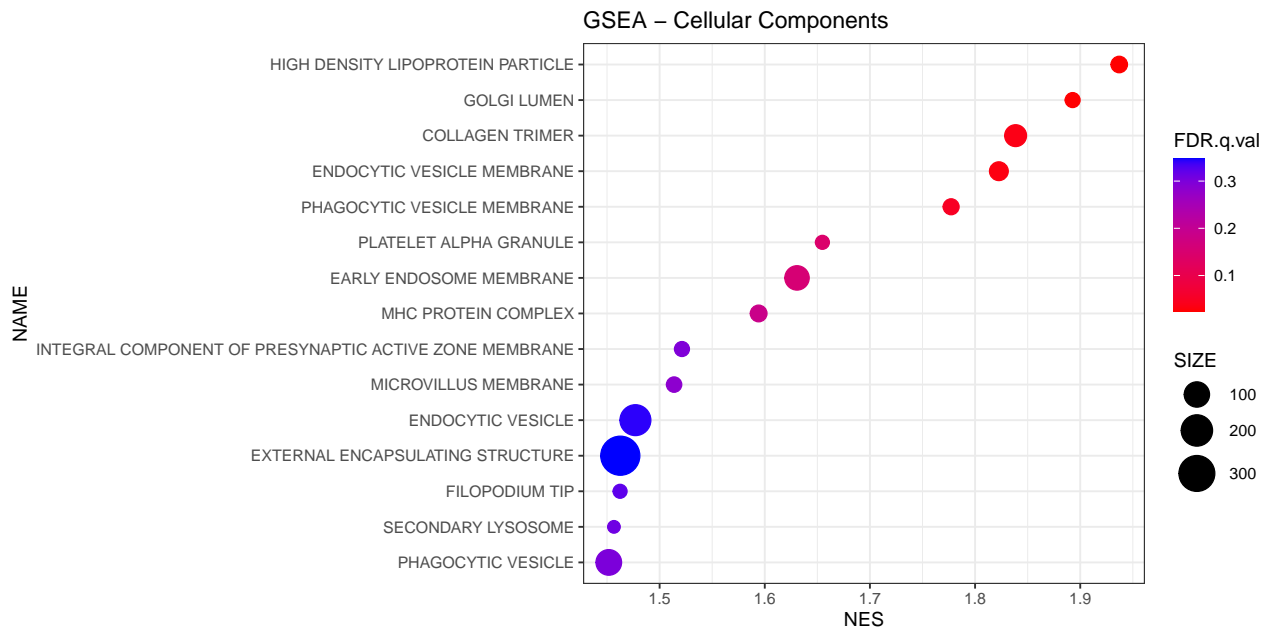
```
GSEA_CC_res <-
read.table(paste0(dir,"my_analysis.Gsea.1675339652585_CC/gsea_report_for_2_1675339652585.tsv"),
header = T, sep = "\t")

GSEA_CC_res <- modify_GSEA_data(GSEA_CC_res)

GSEA_CC_res <- GSEA_CC_res[1:15,]

GSEA_CC_res$NAME <- substring(GSEA_CC_res$NAME, 6)
GSEA_CC_res$NAME <- gsub("_", " ", GSEA_CC_res$NAME)
GSEA_CC_res$NAME <- fct_rev(factor(GSEA_CC_res$NAME, levels = GSEA_CC_res$NAME))

ggplot(GSEA_CC_res, aes(x = NES, y = NAME, size = SIZE, color = FDR.q.val)) + geom_point()
+ ggtitle("GSEA - Cellular Components") + scale_color_gradient(low = "red", high =
"blue") + theme_bw() + scale_size(range = c(3,10))
```



```
#ggsave("plot_result/GSEA_results_CC_céci2.pdf", width = 12, height = 8)
```

## MF results

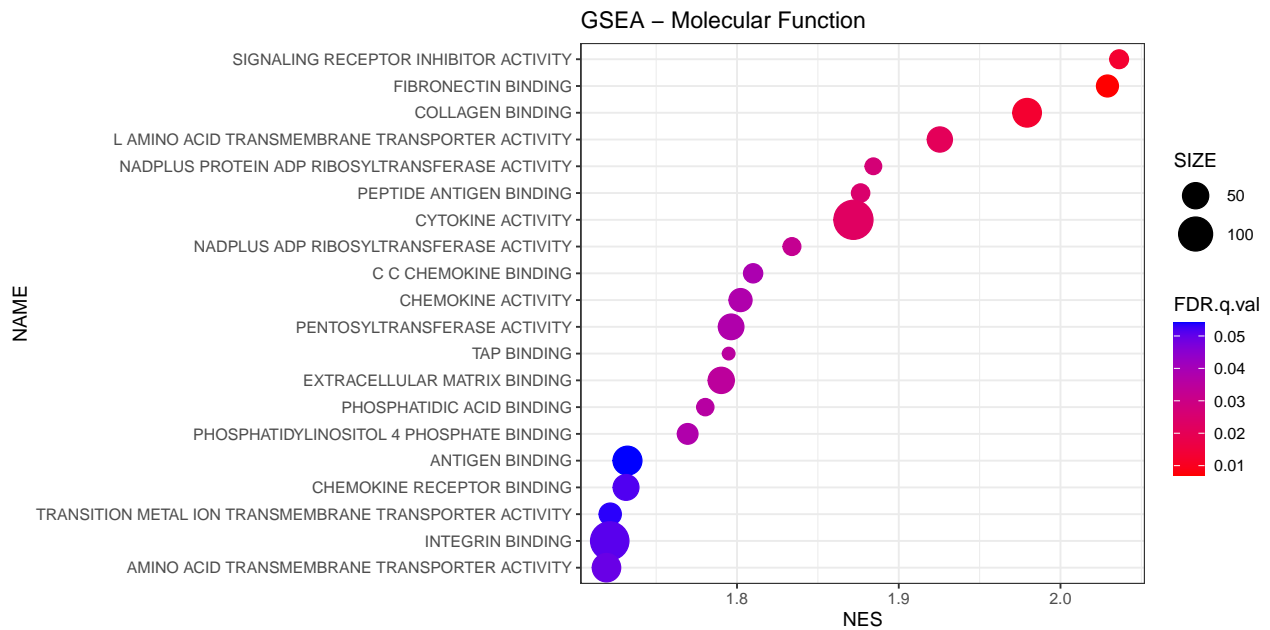
```
GSEA_MF_res <-
read.table(paste0(dir,"my_analysis.Gsea.1675341806364_MF/gsea_report_for_2_1675341806364.tsv"),
header = T, sep = "\t")

GSEA_MF_res <- modify_GSEA_data(GSEA_MF_res)

GSEA_MF_res$NAME <- substring(GSEA_MF_res$NAME, 6)
GSEA_MF_res$NAME <- gsub("_", " ", GSEA_MF_res$NAME)

GSEA_MF_res$NAME <- fct_rev(factor(GSEA_MF_res$NAME, levels = GSEA_MF_res$NAME))

ggplot(GSEA_MF_res, aes(x = NES,y =NAME, size = SIZE, color = FDR.q.val)) + geom_point()
+ ggtitle("GSEA – Molecular Function") + scale_color_gradient(low = "red", high = "blue")
+ theme_bw() + scale_size(range = c(3,10))
```



```
#ggsave("plot_result/GSEA_results_MF_céci2.pdf", width = 12, height = 8)
```

## KEGG results

```
GSEA_KEGG_res <-
read.table(paste0(dir,"my_analysis.Gsea.1675776471604_KEGG/gsea_report_for_Motro_1675776471604.tsv"),
header = T, sep = "\t")

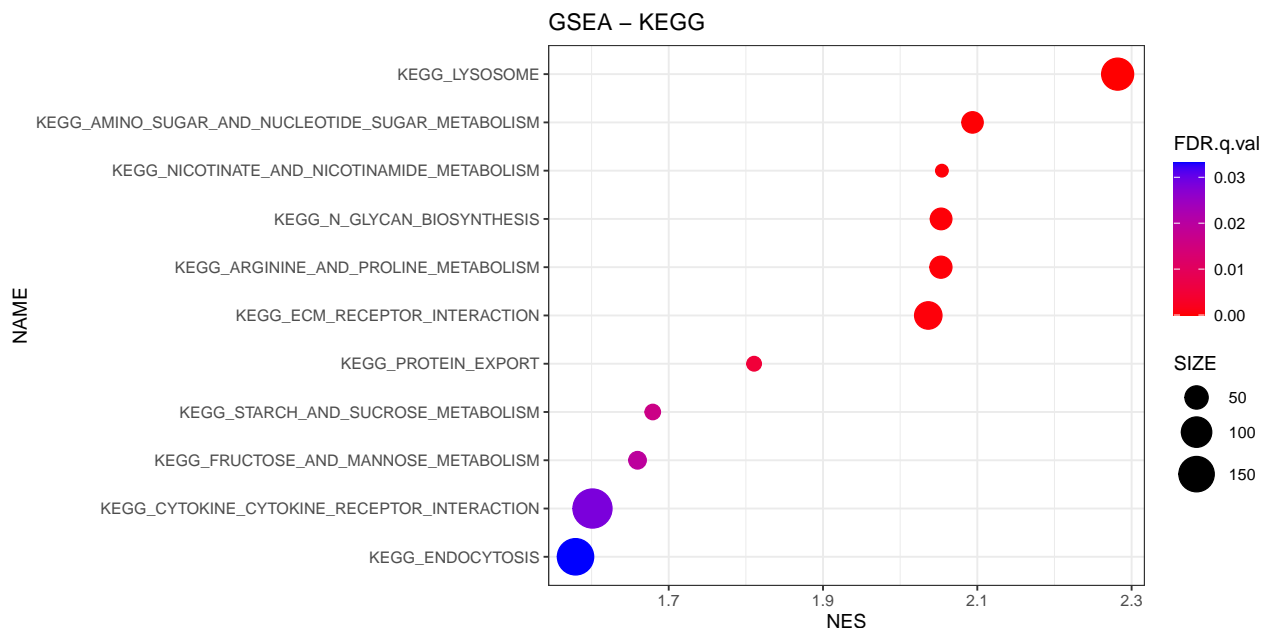
GSEA_KEGG_res<- GSEA_KEGG_res[order(GSEA_KEGG_res$NES, decreasing = T),]

#GSEA_KEGG_res <- modify_GSEA_data(GSEA_KEGG_res)

GSEA_KEGG_res <- GSEA_KEGG_res[c(4,8,9,10,11,12,26,37,38,45,47),]

GSEA_KEGG_res$NAME <- fct_rev(factor(GSEA_KEGG_res$NAME, levels = GSEA_KEGG_res$NAME))

ggplot(GSEA_KEGG_res, aes(x = NES,y =NAME, size = SIZE, color = FDR.q.val)) +
geom_point() + ggtitle("GSEA - KEGG") + scale_color_gradient(low = "red", high = "blue")
+ theme_bw()+ scale_size(range = c(3,10))
```



```
#ggsave("plot_result/GSEA_results_MF_céci2.pdf", width = 12, height = 8)
```

```
sessionInfo()
```

```
## R version 4.3.1 (2023-06-16)
## Platform: x86_64-pc-linux-gnu (64-bit)
## Running under: Ubuntu 22.04.3 LTS
##
## Matrix products: default
## BLAS: /usr/lib/x86_64-linux-gnu/openblas-pthread/libblas.so.3
## LAPACK: /usr/lib/x86_64-linux-gnu/openblas-pthread/libopenblas-p-r0.3.20.so; LAPACK version 3.10.0
##
## locale:
##  [1] LC_CTYPE=en_US.UTF-8 LC_NUMERIC=C
##  [3] LC_TIME=fr_BE.UTF-8 LC_COLLATE=en_US.UTF-8
##  [5] LC_MONETARY=fr_BE.UTF-8 LC_MESSAGES=en_US.UTF-8
##  [7] LC_PAPER=fr_BE.UTF-8 LC_NAME=C
##  [9] LC_ADDRESS=C LC_TELEPHONE=C
## [11] LC_MEASUREMENT=fr_BE.UTF-8 LC_IDENTIFICATION=C
##
## time zone: Europe/Brussels
## tzcode source: system (glibc)
##
## attached base packages:
## [1] stats graphics grDevices utils datasets methods base
##
## other attached packages:
## [1] forcats_1.0.0 ggplot2_3.4.2
##
## loaded via a namespace (and not attached):
##  [1] vctrs_0.6.3 cli_3.6.1 knitr_1.43 rlang_1.1.1
##  [5] xfun_0.39 highr_0.10 generics_0.1.3 labeling_0.4.2
##  [9] glue_1.6.2 colorspace_2.1-0 htmltools_0.5.5 scales_1.2.1
## [13] fansi_1.0.4 rmarkdown_2.23 grid_4.3.1 evaluate_0.21
```

```
## [17] munsell_0.5.0    tibble_3.2.1    fastmap_1.1.1    yaml_2.3.7
## [21] lifecycle_1.0.3  compiler_4.3.1  dplyr_1.1.2      pkgconfig_2.0.3
## [25] rstudioapi_0.14  farver_2.1.1    digest_0.6.33    R6_2.5.1
## [29] tidyselect_1.2.0 utf8_1.2.3      pillar_1.9.0     magrittr_2.0.3
## [33] withr_2.5.0      tools_4.3.1     gtable_0.3.3
```