Pathway Analysis - Macrophages M2
Input data: DEGs

Elena Eyre Sánchez, PhD

2025-01-04

Contents

1 Introducción y Objetivo

En este script realizo un pathways análisis con los DEGs obtenidos de los Venn Diagrams, y preparo el resultado en archivos gem para extraer un gráfico posteriormente en Cytoscape (https://cran.r-project.org/web/packages/gprofiler2/vignettes/gprofiler2.html).

2 Paquetes y datos

3 Datos

Cargo inicialmente los listados de Macrófagos tipo M2 obtenidos en jVenn.

```
setwd("~/Desktop/ELENA_UOC/TFM")
jVenn_MacrophagesM2_vs_GSE22155_02 <- read_csv("DEGs/Venn_diagrams_cell_type_treatments_vs_untreated/jV
head(jVenn_MacrophagesM2_vs_GSE22155_02) # sección de los resultados
```

```
## # A tibble: 6 x 3
##   GSE50509  GSE22155_02 `GSE50509|GSE22155_02`
##   <chr>     <chr>       <chr>
## 1 CEMIP     ADGRF5      MN1
## 2 LINC01091 TRPC6       DPT
## 3 COMP      GIMAP8      CRISPLD2
## 4 ITGBL1    GIMAP7      CDH11
## 5 PRND      TAL1        PDGFRL
## 6 GPR68     VWF         FNDC1
```

```
jVenn_MacrophagesM2_vs_GSE22155_47 <- read_csv("DEGs/Venn_diagrams_cell_type_treatments_vs_untreated/jV
jVenn_MacrophagesM2_vs_tcgaskcm <- read_csv("DEGs/Venn_diagrams_cell_type_treatments_vs_untreated/jVenn
jVenn_MacrophagesM2_vs_GSE91061 <- read_csv("DEGs/Venn_diagrams_cell_type_treatments_vs_untreated/jVenn
jVenn_MacrophagesM2_vs_GSE35640 <- read_csv("DEGs/Venn_diagrams_cell_type_treatments_vs_untreated/jVenn
jVenn_MacrophagesM2_vs_GSE61992 <- read_csv("DEGs/Venn_diagrams_cell_type_treatments_vs_untreated/jVenn
setwd("~/Desktop/ELENA_UOC/TFM/GOBPs") # Directorio de trabajo
jVenn_MacrophagesM2 <- read_delim("~/Desktop/ELENA_UOC/TFM/DEGs/Venn_diagrams_cell_type_treatments_compa
    delim = ",", escape_double = FALSE, trim_ws = TRUE)
```

4 Uncovered

En esta sección contemplo la posibilidad que los DEGs únicos en el dataset sin tratar se traten de mecanismos que se escapan al tratamiento pero que igualmente se ven afectados por la enfermedad.

Primero realizo el pathway multianálisis:

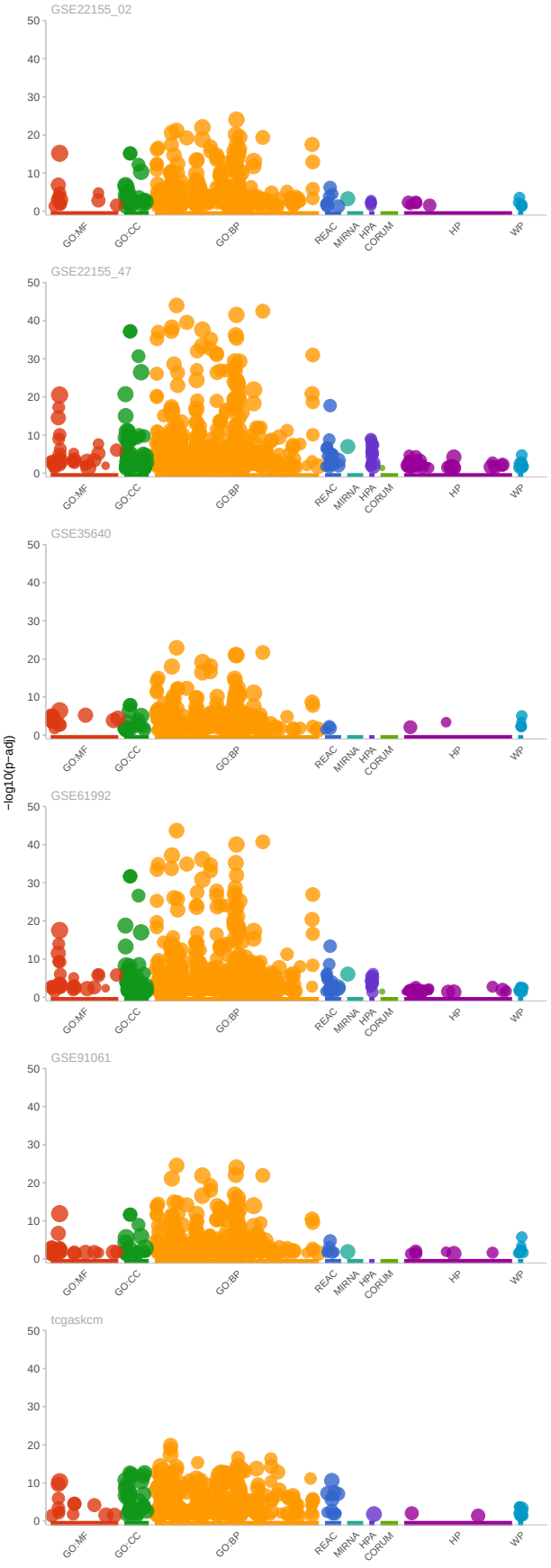
En esta sección contemplo la posibilidad que los DEGs únicos en el dataset sin tratar se traten de mecanismos que se escapan al tratamiento pero que igualmente se ven afectados por la enfermedad.

Primero realizo el pathway multianálisis:

```
multi_gostres <- gost(query = list("GSE22155_02" = unique(jVenn_MacrophagesM2_vs_GSE22155_02$GSE50509[wh
                                   "GSE22155_47" = jVenn_MacrophagesM2_vs_GSE22155_47$GSE50509[which(is.na(jV
                                   "GSE35640" = jVenn_MacrophagesM2_vs_GSE35640$GSE50509[which(is.na(jVenn_Ma
                                   "GSE91061" = jVenn_MacrophagesM2_vs_GSE91061$GSE50509[which(is.na(jVenn_Ma
                                   "GSE61992" = jVenn_MacrophagesM2_vs_GSE61992$GSE50509[which(is.na(jVenn_Ma
                                   "tcgaskcm" = jVenn_MacrophagesM2_vs_tcgaskcm$GSE50509[which(is.na(jVenn_Ma
                                   ),
                      evcodes = TRUE, multi_query = FALSE,
                      sources = c("GO", "REAC", "MIRNA", "CORUM", "HP", "HPA", "WP"))
head(multi_gostres$result[,1:7])
```

```
##         query significant      p_value term_size query_size intersection_size
## 1 GSE22155_02        TRUE 9.002839e-25      5899        846              384
## 2 GSE22155_02        TRUE 9.758562e-23      6453        846              403
## 3 GSE22155_02        TRUE 6.181380e-22      2684        846              217
## 4 GSE22155_02        TRUE 2.574064e-21      4643        846              315
## 5 GSE22155_02        TRUE 7.024756e-21      3973        846              281
## 6 GSE22155_02        TRUE 3.416424e-20      2955        846              227
##   precision
## 1 0.4539007
## 2 0.4763593
## 3 0.2565012
## 4 0.3723404
## 5 0.3321513
## 6 0.2683215
```

```
p <- gostplot(multi_gostres, capped = FALSE, interactive = F)
p
```


GSE22155_02
GSE22155_47
GSE35640
GSE61992
GSE91061
tcgaskcm
−log10(p-adj)
0
10
20
30
40
50
GO:MF
GO:CC
GO:BP
REAC
MIRNA
HPA
CORUM
HP
WP

Estos resultados son los que utilizo para generar los archivos gem que archivo para usar más tarde en Cytoscape.

```
# format to GEM
gem <- multi_gostres$result[,c("query", "term_id", "term_name", "p_value", "intersection_size")]
colnames(gem) <- c("query", "GO.ID", "Description", "p.Val", "Genes")
gem$FDR <- gem$p.Val
gem$Phenotype = "+1"

gem %>% group_by(query) %>%
  group_walk(~
    write.table(data.frame(.x[,c("GO.ID", "Description", "p.Val", "FDR", "Phenotype", "Genes")]),
                file = paste0("./Venn_diagrams_cell_type_treatments_vs_untreated/gProfiler_MacrophagesM
                sep = "\t", quote = F, row.names = F))
```

Proporciones de los pathways para cada estudio:

```
df <- as.data.frame(multi_gostres$result[,1:13])
table(df$query)
```

```
##
## GSE22155_02 GSE22155_47    GSE35640    GSE61992    GSE91061    tcgaskcm
##         360         650         295         591         363         422
```

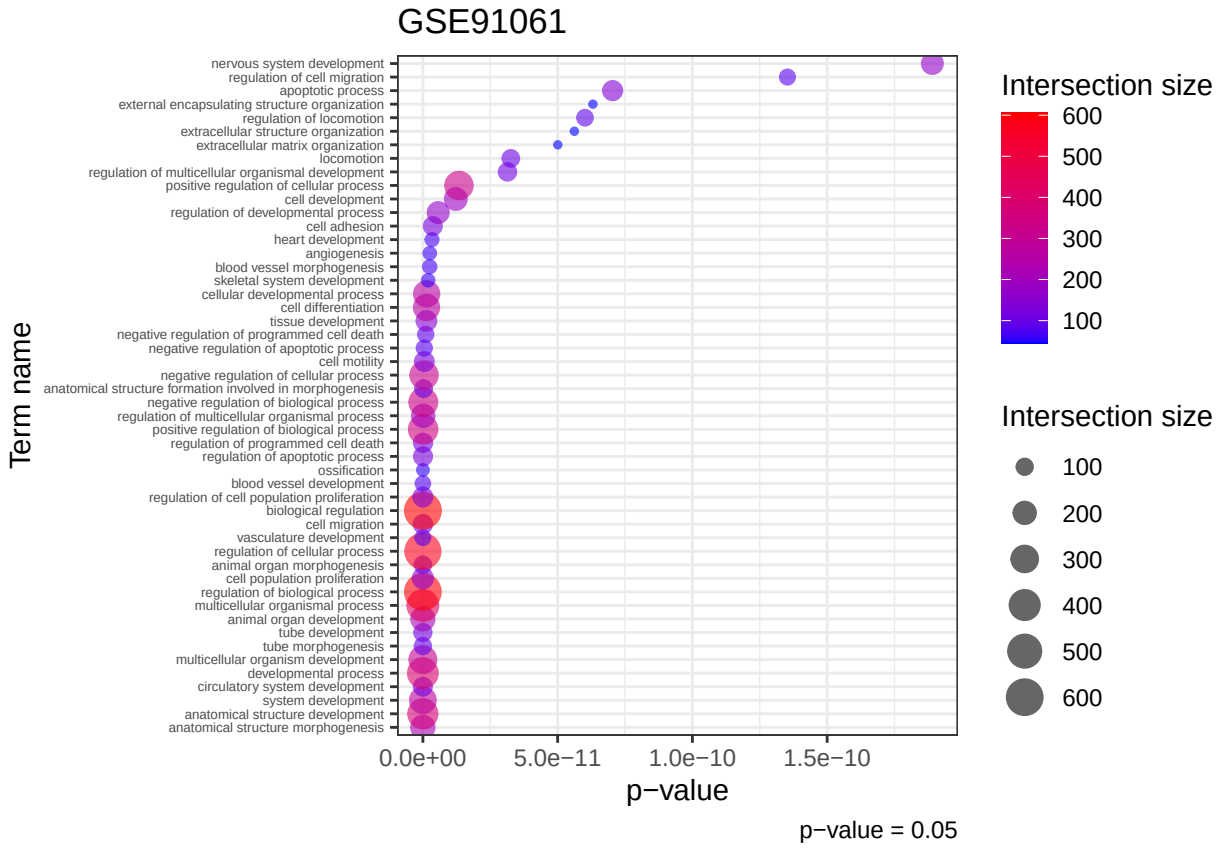
```
prop.table(table(df$query))*100
```

```
##
## GSE22155_02 GSE22155_47    GSE35640    GSE61992    GSE91061    tcgaskcm
##    13.42783    24.24468    11.00336    22.04401    13.53972    15.74040
```

```
#table(df$term_name)
#prop.table(table(df$term_name))*100
df2 <- as.data.frame(table(df$term_name, df$query))
#length(unique(df2$Var1))
write.table(df, file = "./Venn_diagrams_cell_type_treatments_vs_untreated/MacrM2_GOBPs.txt", sep = "\t"
write.table(df2, file = "./Venn_diagrams_cell_type_treatments_vs_untreated/MacrM2_GOBPs_freq.txt", sep =
#rm(jVenn_MacrophagesM2_vs_GSE22155_02, jVenn_MacrophagesM2_vs_GSE22155_47, jVenn_MacrophagesM2_vs_GSE3
```

Barplot of the top GO-BPs:

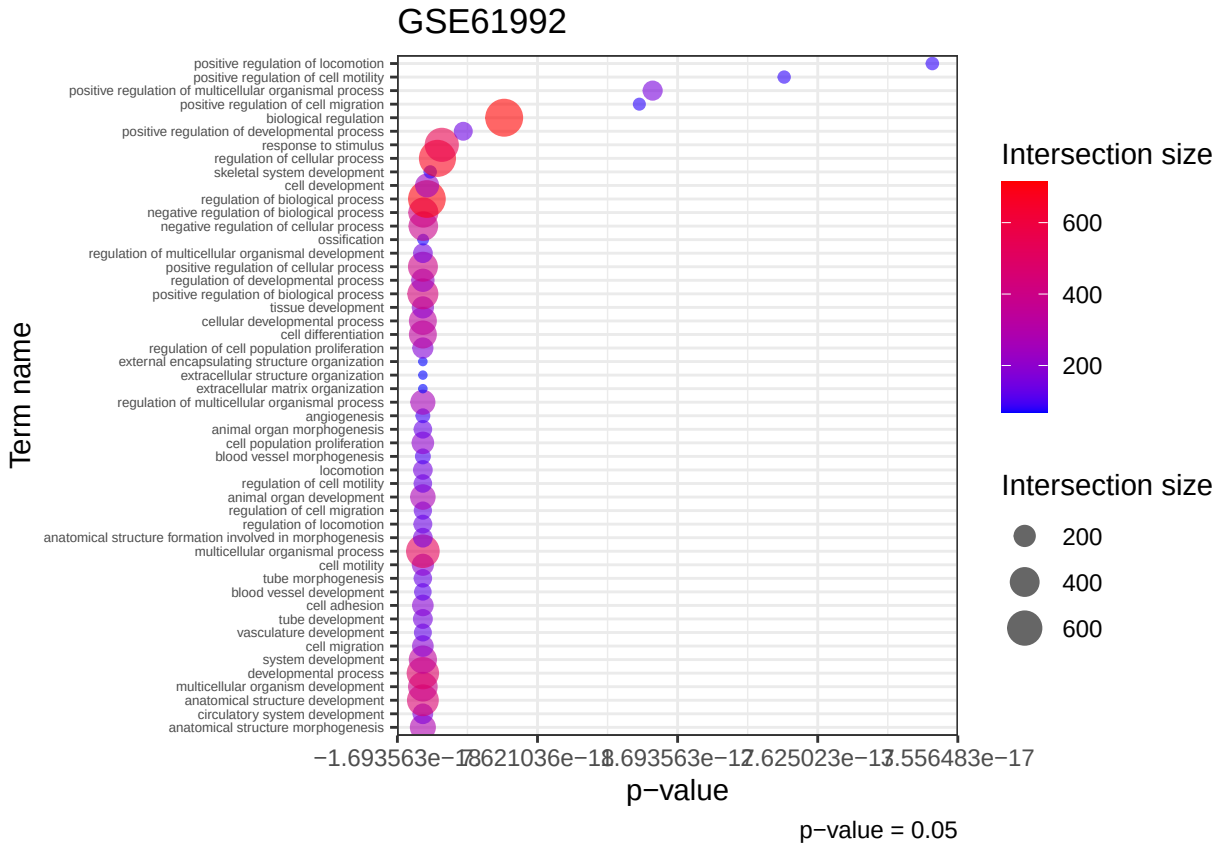
```
plot_gobps <- function(study, n = 50){
df2 <- df[df$source == "GO:BP",]
df2_GSE91061 <- df2[df2$query == study,]
df2_GSE91061 <- df2_GSE91061[1:n,]
ggplot(df2_GSE91061, aes(p_value, reorder(term_name,p_value))) + ggtitle(study) +
  geom_point(aes(p_value, reorder(term_name,p_value), colour = intersection_size, size = intersection_si
  theme_bw() + theme(axis.text.y.left = element_text(size = 5)) +
  labs(x= "p-value", y="Term name", colour = "Intersection size", size = "Intersection size", caption =
  scale_color_gradient(low="blue", high="red")
}

plot_gobps(study = "GSE91061")
```




```
print("\n")
```

```
## [1] "\n"
```

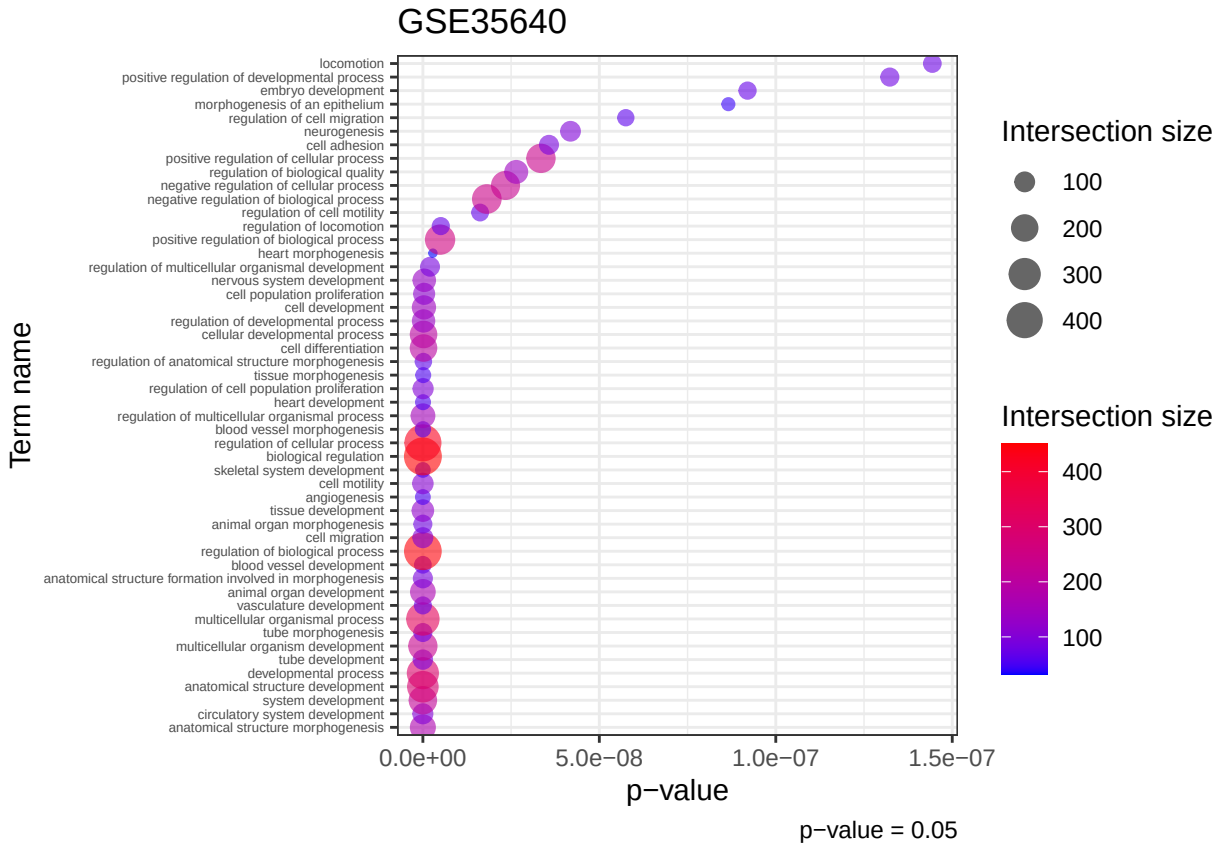
```
plot_gobps(study = "GSE61992")
```




```
print("\n")
```

```
## [1] "\n"
```

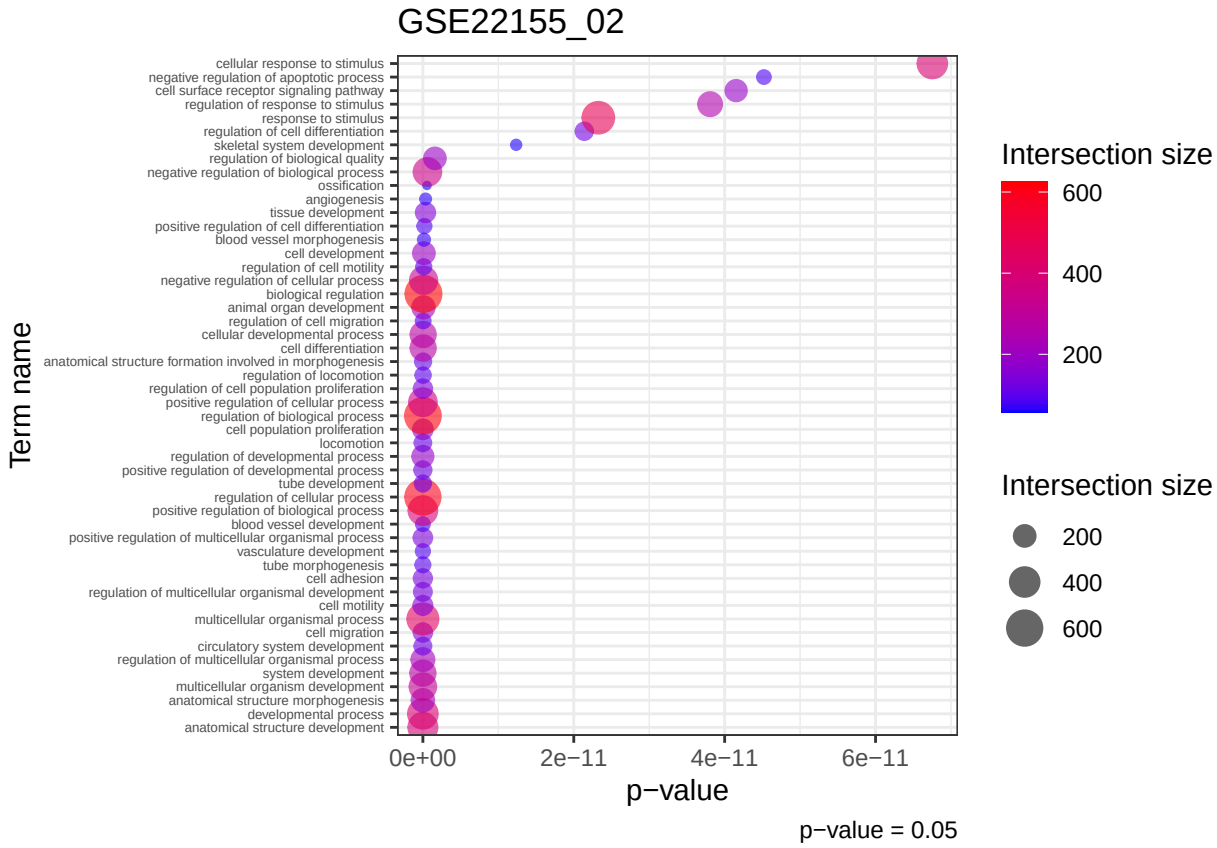
```
plot_gobps(study = "GSE35640")
```




```
print("\n")
```

```
## [1] "\n"
```

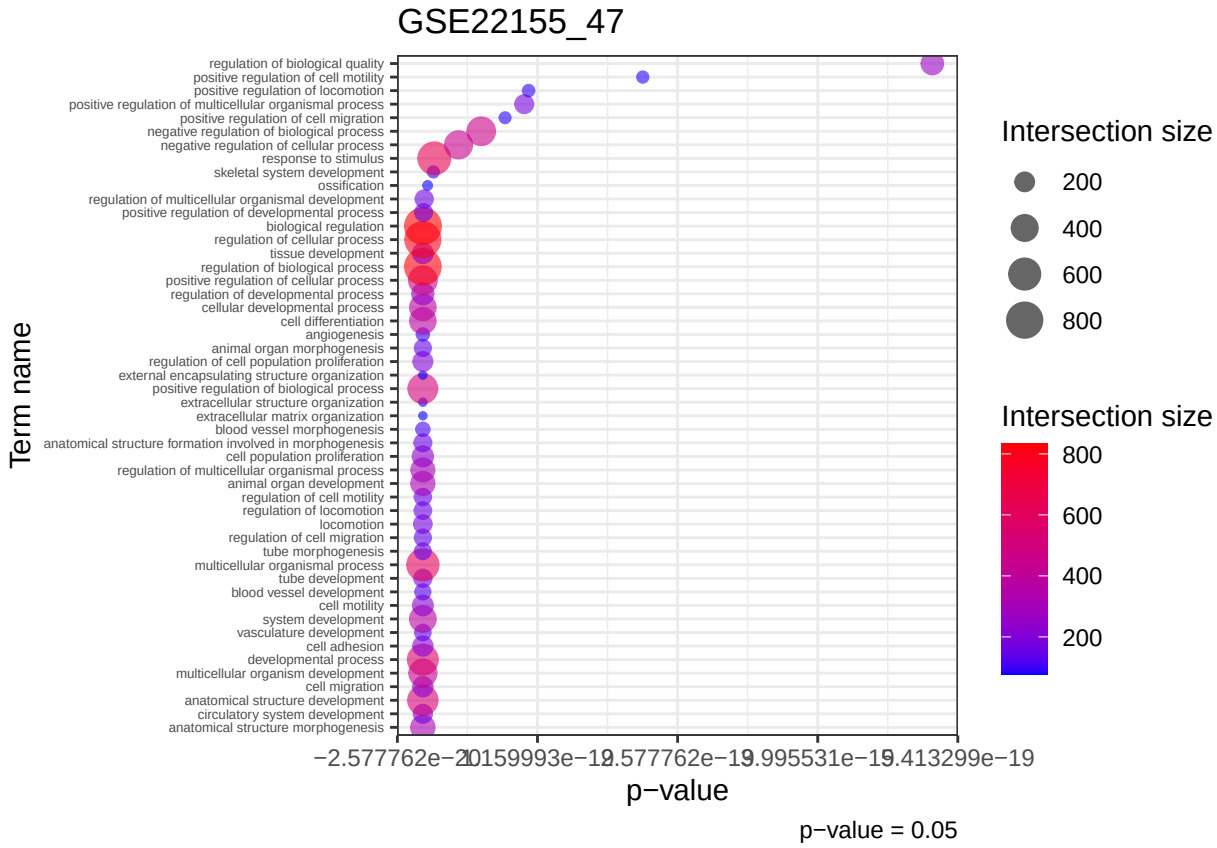
```
plot_gobps(study = "GSE22155_02")
```




```
print("\n")
```

```
## [1] "\n"
```

```
plot_gobps(study = "GSE22155_47")
```



5 Exclusive

En esta sección analizo los DEGs exclusivos para cada estudio, y realizo un multi-análisis para cada estudio. Los resultados los guardo como archivos gem específicos para usar en Cytoscape.

Primero realizo el pathway multianálisis:

```
multi_gostres <- gost(query = list("GSE61992" = jVenn_MacrophagesM2$GSE61992[which(is.na(jVenn_Macropha
                                   "GSE91061" = jVenn_MacrophagesM2$GSE91061[which(is.na(jVenn_MacrophagesM2$
                                   "tcgaskcm" = jVenn_MacrophagesM2$`TCGA-SKCM`[which(is.na(jVenn_MacrophagesM
                                   ),
                      evcodes = TRUE, multi_query = FALSE,
                      sources = c("GO", "REAC", "MIRNA", "CORUM", "HP", "HPA", "WP"))
head(multi_gostres$result[,1:7])
```

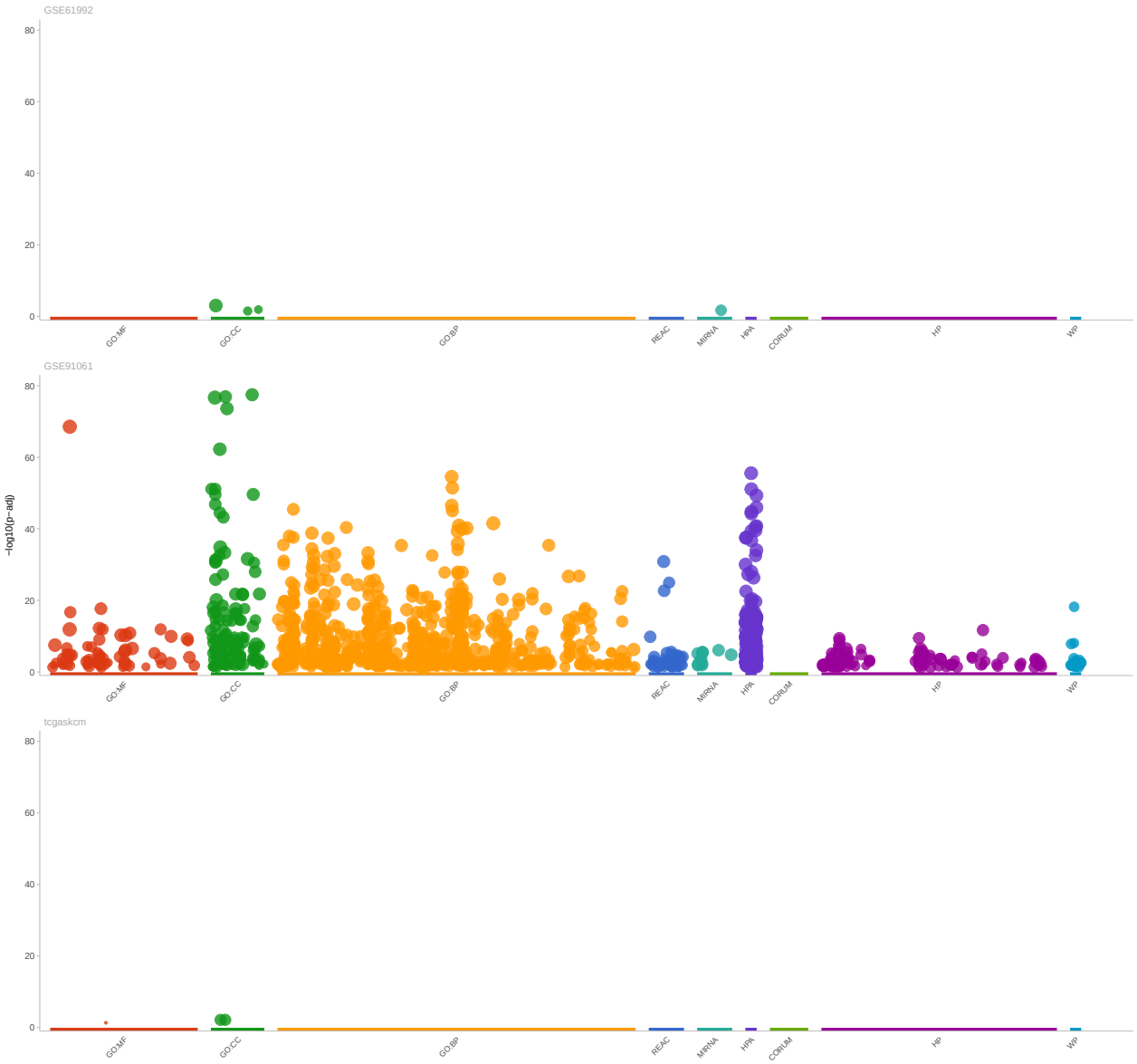
```
##      query significant      p_value term_size query_size intersection_size
## 1 GSE61992        TRUE 9.168699e-04      5487         29               19
## 2 GSE61992        TRUE 1.261225e-02         9         29                2
## 3 GSE61992        TRUE 3.175138e-02        14         29                2
## 4 GSE61992        TRUE 1.968881e-02       222         27                5
## 5 GSE91061        TRUE 2.713593e-55      6250       2330             1039
## 6 GSE91061        TRUE 3.476259e-52      3939       2330              732
##    precision
## 1 0.65517241
## 2 0.06896552
## 3 0.06896552
## 4 0.18518519
```

```
## 5 0.44592275
## 6 0.31416309
```

Estos resultados son los que utilizo para generar los archivos gem que archivo para usar más tarde en Cytoscape.

```
# format to GEM
gem <- multi_gostres$result[,c("query", "term_id", "term_name", "p_value", "intersection_size")]
colnames(gem) <- c("query", "GO.ID", "Description", "p.Val", "Genes")
gem$FDR <- gem$p.Val
gem$Phenotype = "+1"

gem %>% group_by(query) %>%
  group_walk(~
    write.table(data.frame(.x[,c("GO.ID", "Description", "p.Val", "FDR", "Phenotype", "Genes")]),
                file = paste0("./Treatment_comparisons/gProfiler_MacrophagesM2_Only_geneset_", unique(.y
                sep = "\t", quote = F, row.names = F))
```

```
p <- gostplot(multi_gostres, capped = FALSE, interactive = F)
p
```



Proporciones de los pathways para cada estudio:

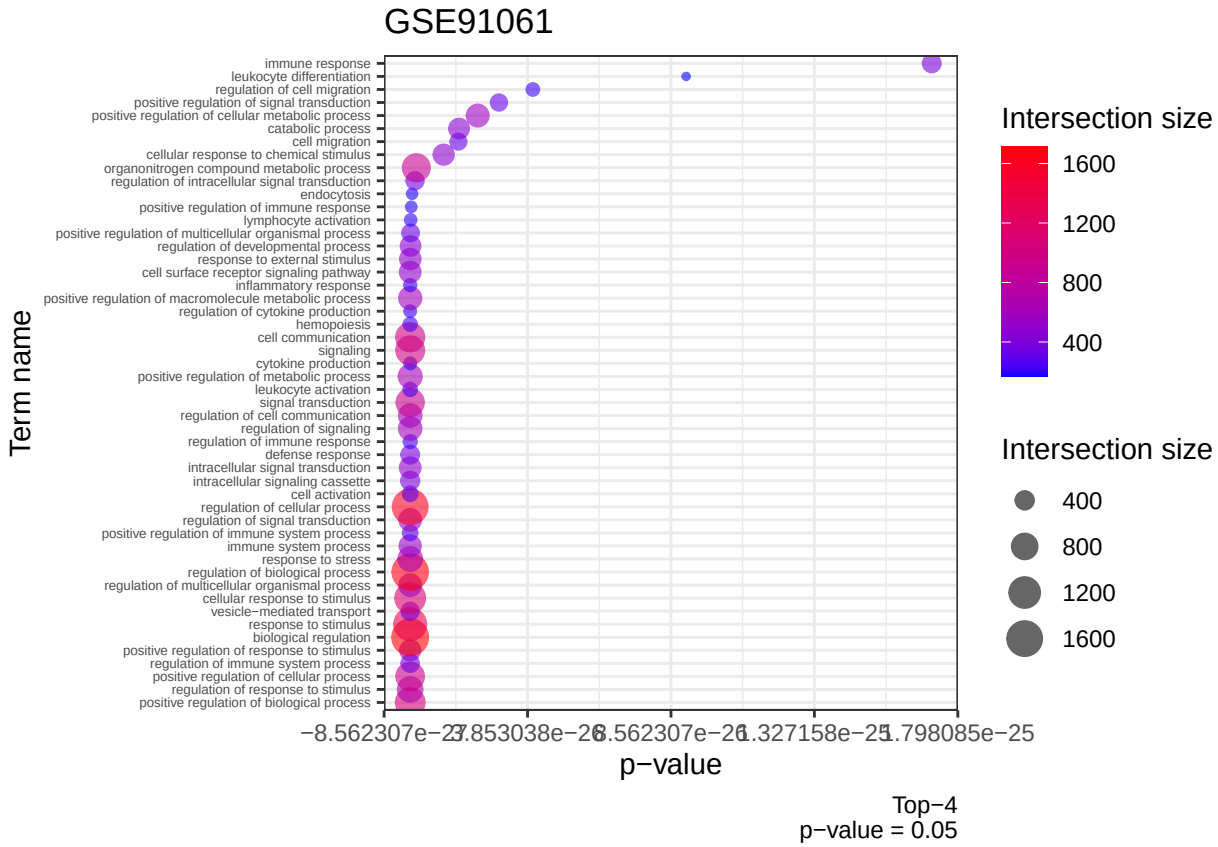
```
df <- as.data.frame(multi_gostres$result[,1:13])
table(df$query)
```

```
##
## GSE61992 GSE91061 tcgaskcm
##        4     1361        3
```

```
write.table(gem, file = "./Treatment_comparisons/gProfiler_MacrophagesM2_Only_genesets.txt", sep = "\t"
```

```
plot_gobps <- function(study){
df2_GSE91061 <- df2[df2$query == study,]
ggplot(df2_GSE91061, aes(p_value, reorder(term_name,p_value))) + ggtitle(study) +
  geom_point(aes(p_value, reorder(term_name,p_value), colour = intersection_size, size = intersection_si
  theme_bw() + theme(axis.text.y.left = element_text(size = 5)) +
  labs(x= "p-value", y="Term name", colour = "Intersection size", size = "Intersection size", caption =
  scale_color_gradient(low="blue", high="red")
```

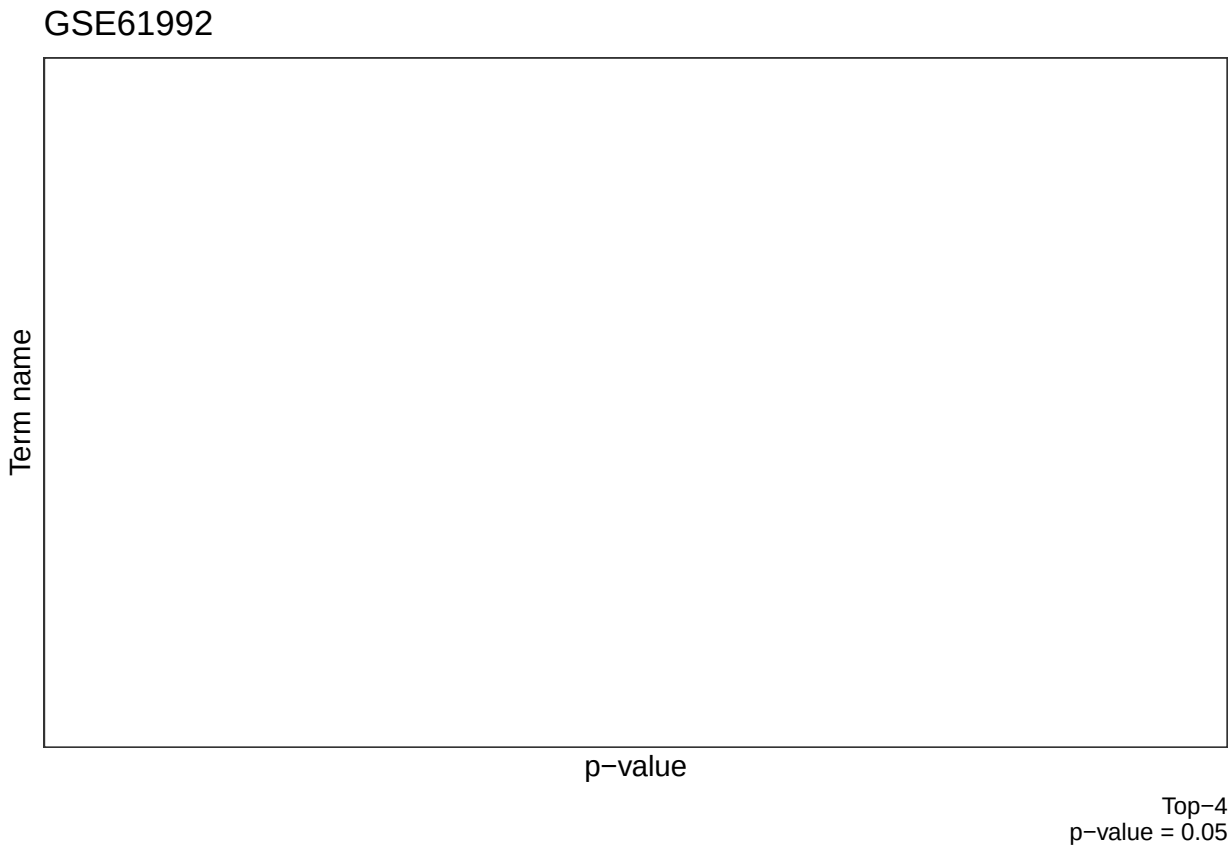
```
}
df2 <- df[df$source == "GO:BP",]
df2 <- df2[1:50,]
plot_gobps(study = "GSE91061")
```




```
print("\n")
```

```
## [1] "\n"
```

```
df2 <- df[df$source == "GO:BP",]
plot_gobps(study = "GSE61992")
```



6 Máximo solapamiento

En esta sección analizo los DEGs que más se comparten entre los estudios para observar los mecanismos que pudieran haber en común entre los tratamientos.

Primero realizo el pathway multianálisis, pero no se obtienen resultados. Por lo tanto, no se procede con el siguiente paso.

```
#multi_gostres <- gost(query = list("antiCTLA4PD1_TabrafenibTrametinib" = jVenn_MacrophagesM2$`GSE91061
#                      ),
#                      evcodes = TRUE, multi_query = FALSE,
#                      sources = c("GO", "REAC", "MIRNA", "CORUM", "HP", "HPA", "WP"))
#head(multi_gostres$result[,1:7])
```