Pathway Analysis - Dendríticas cells
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 células Dendríticas obtenidas en jVenn.

```
setwd("~/Desktop/ELENA_UOC/TFM")
jVenn_DendriticCells_vs_GSE22155_02 <- read_csv("DEGs/Venn_diagrams_cell_type_treatments_vs_untreated/j
head(jVenn_DendriticCells_vs_GSE22155_02) # sección de los resultados
```

```
## # A tibble: 6 x 3
##   GSE50509  GSE22155_02 `GSE50509|GSE22155_02`
##   <chr>     <chr>       <chr>
## 1 LINC01091 MYOZ3       C1orf52
## 2 CEMIP     LEFTY1      ZMYM3
## 3 CCDC51    FLNA        SYT13
## 4 NSMCE2    PASD1       CDK4
## 5 PCGF2     PKP4        HYAL3
## 6 C8orf76   CARD9       TRIM65
```

```
jVenn_DendriticCells_vs_GSE22155_47 <- read_csv("DEGs/Venn_diagrams_cell_type_treatments_vs_untreated/j
jVenn_DendriticCells_vs_GSE91061 <- read_csv("DEGs/Venn_diagrams_cell_type_treatments_vs_untreated/jVen
jVenn_DendriticCells_vs_GSE35640 <- read_csv("DEGs/Venn_diagrams_cell_type_treatments_vs_untreated/jVen
jVenn_DendriticCells_vs_GSE61992 <- read_csv("DEGs/Venn_diagrams_cell_type_treatments_vs_untreated/jVen
setwd("~/Desktop/ELENA_UOC/TFM/GOBPs") # Directorio de trabajo
jVenn_DendriticCells <- read_delim("~/Desktop/ELENA_UOC/TFM/DEGs/Venn_diagrams_cell_type_treatments_com
    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:

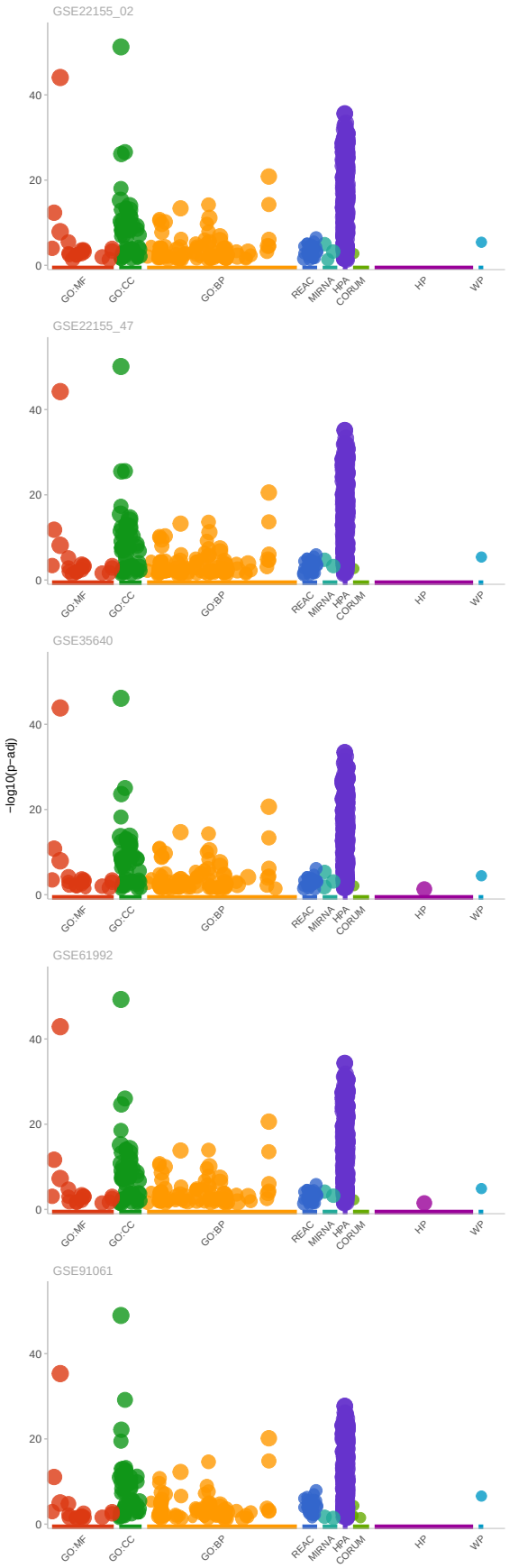
En esta sección contemplo la posibilidad que los DEGs únicos en el dataset sin tratar se traten de mecanismos que se escapan al tratamiento pero que igualmente se ven afectados por la enfermedad.

Primero realizo el pathway multianálisis:

```
multi_gostres <- gost(query = list("GSE22155_02" = unique(jVenn_DendriticCells_vs_GSE22155_02$GSE50509[
                                   "GSE22155_47" = jVenn_DendriticCells_vs_GSE22155_47$GSE50509[which(is.na(j
                                   "GSE35640" = jVenn_DendriticCells_vs_GSE35640$GSE50509[which(is.na(jVenn_D
                                   "GSE91061" = jVenn_DendriticCells_vs_GSE91061$GSE50509[which(is.na(jVenn_D
                                   "GSE61992" = jVenn_DendriticCells_vs_GSE61992$GSE50509[which(is.na(jVenn_D
                                   ),
                      evcodes = TRUE, multi_query = FALSE,
                      sources = c("GO", "REAC", "MIRNA", "CORUM", "HP", "HPA", "WP"))
head(multi_gostres$result[,1:7])
```

```
##         query significant      p_value term_size query_size intersection_size
## 1 GSE22155_02        TRUE 1.713402e-03        74        400                23
## 2 GSE22155_02        TRUE 1.440922e-21      5986       1567               628
## 3 GSE22155_02        TRUE 5.337948e-15      1761       1567               231
## 4 GSE22155_02        TRUE 5.978821e-15      1194       1567               174
## 5 GSE22155_02        TRUE 4.207302e-14      4912       1567               507
## 6 GSE22155_02        TRUE 7.206663e-12      4836       1567               490
##   precision
## 1 0.0575000
## 2 0.4007658
## 3 0.1474154
## 4 0.1110402
## 5 0.3235482
## 6 0.3126994
```

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


GSE22155_02
GSE22155_47
GSE35640
GSE61992
GSE91061
−log10(p−adj)
0
20
40
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_DendriticCel
                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)
```

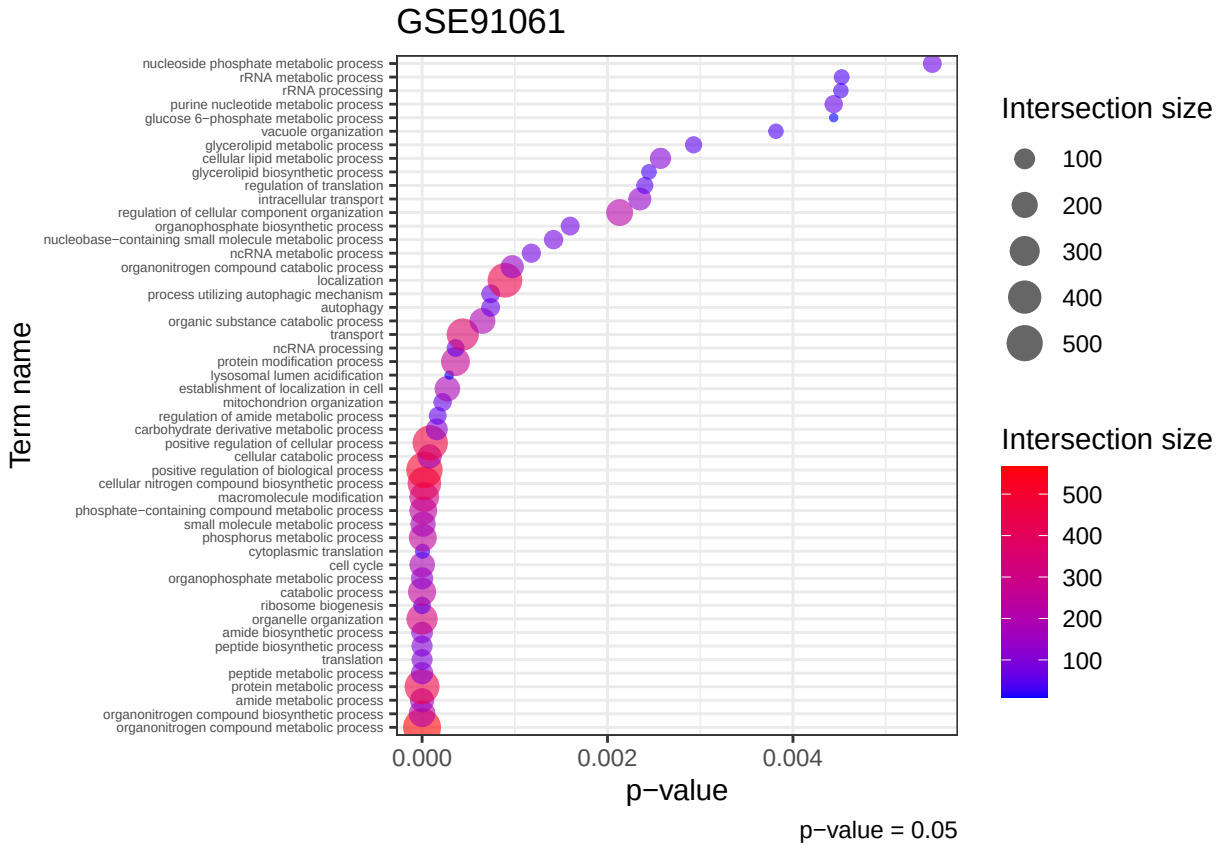
```
##
## GSE22155_02 GSE22155_47    GSE35640    GSE61992    GSE91061
##         439         438         409         419         413
```

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

```
##
## GSE22155_02 GSE22155_47    GSE35640    GSE61992    GSE91061
##    20.72710    20.67989    19.31067    19.78281    19.49953
```

```
#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/DendriticCells_GOBPs.txt", sep
write.table(df2, file = "./Venn_diagrams_cell_type_treatments_vs_untreated/DendriticCells_GOBPs_freq.tx
rm(jVenn_DendriticCells_vs_GSE22155_02, jVenn_DendriticCells_vs_GSE22155_47, jVenn_DendriticCells_vs_GS
```

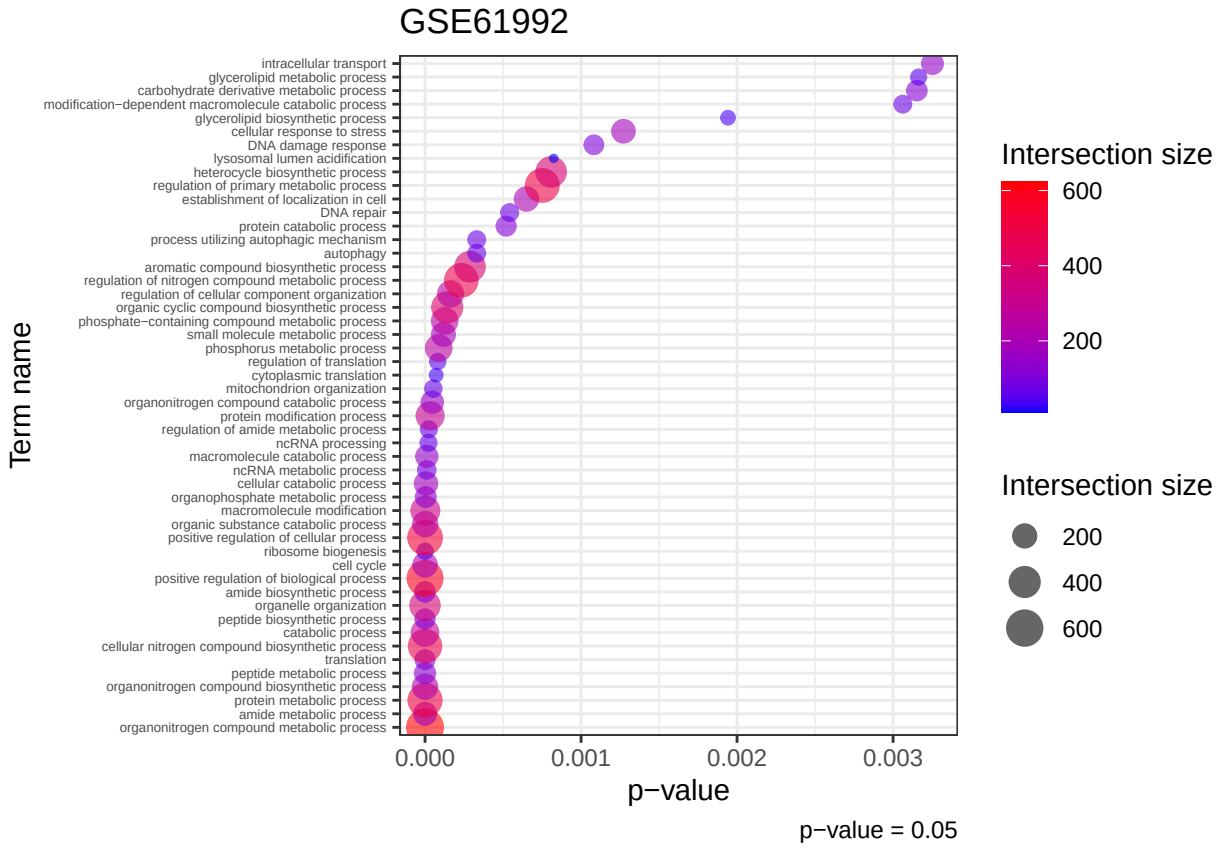
```
plot_gobps <- function(study, n = 50){
df2 <- df[df$source == "GO:BP",]
df2_GSE91061 <- df2[df2$query == study,]
df2_GSE91061 <- df2_GSE91061[1:n,]
ggplot(df2_GSE91061, aes(p_value, reorder(term_name,p_value))) + ggtitle(study) +
  geom_point(aes(p_value, reorder(term_name,p_value), colour = intersection_size, 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"
```

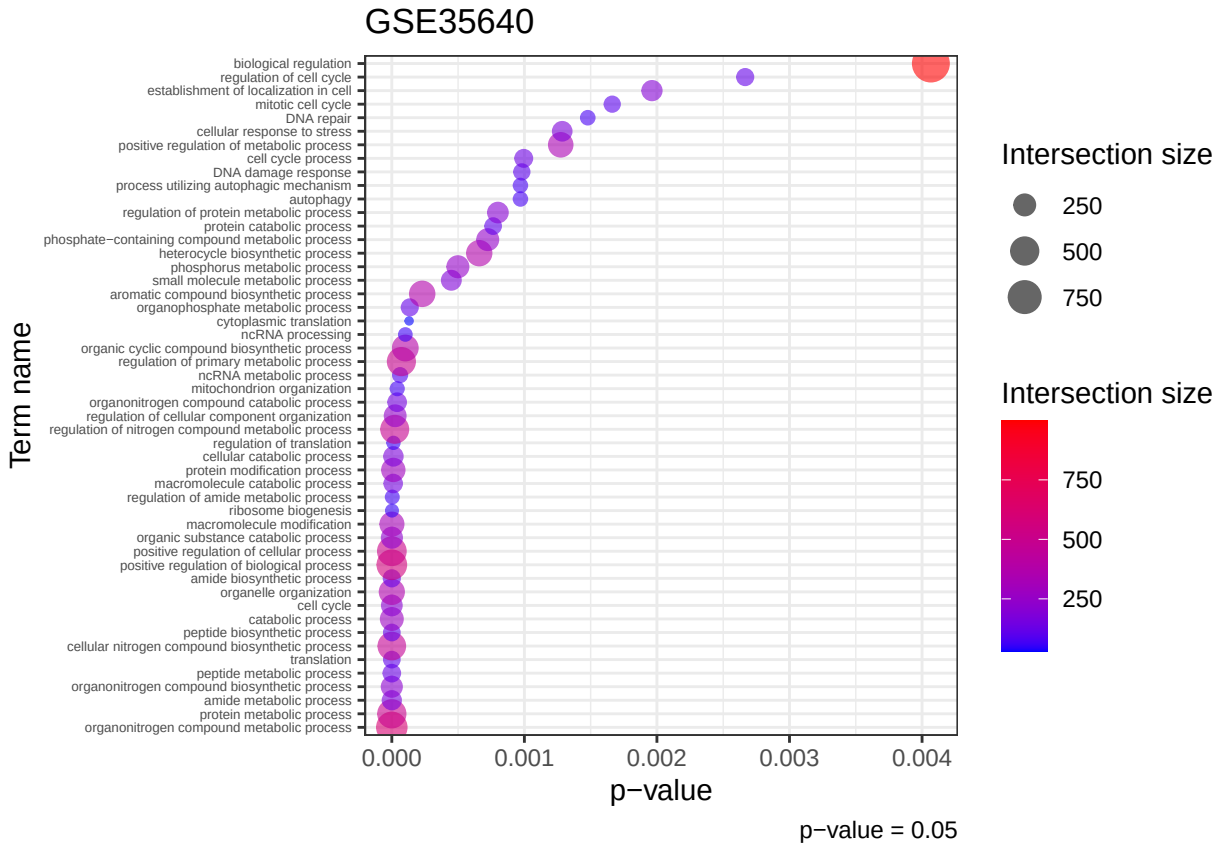
```
plot_gobps(study = "GSE61992")
```




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

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

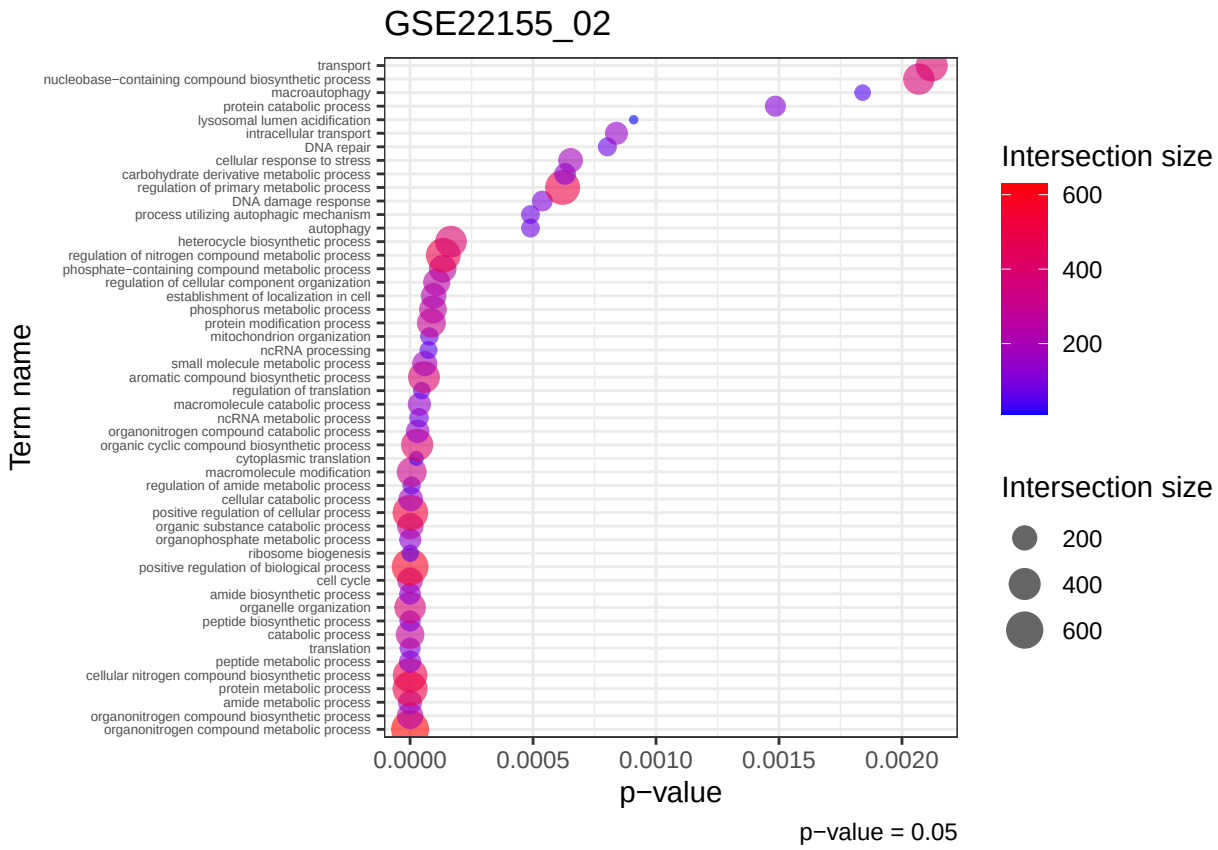
```
plot_gobps(study = "GSE35640")
```




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

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

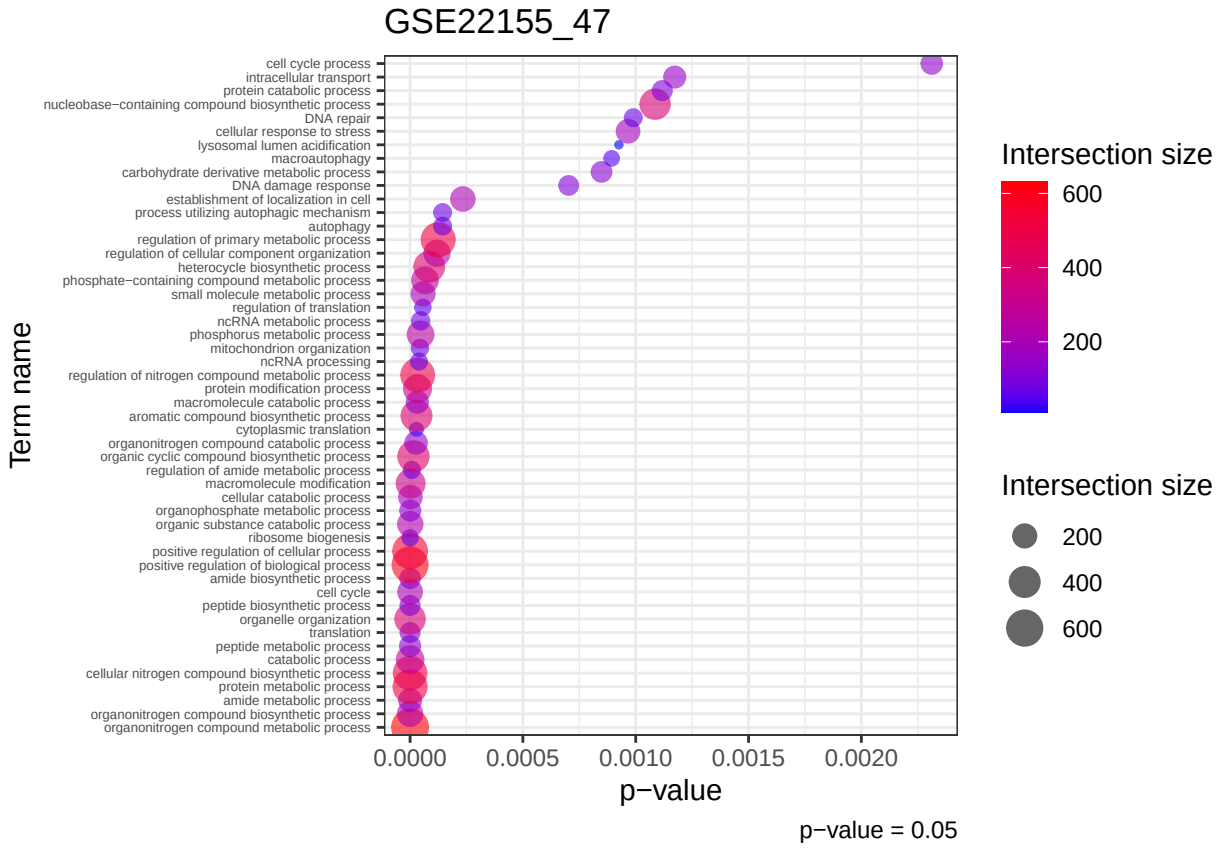
```
plot_gobps(study = "GSE22155_02")
```




```
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_DendriticCells$GSE61992[which(is.na(jVenn_Dendriti
                                   "GSE91061" = jVenn_DendriticCells$GSE91061[which(is.na(jVenn_DendriticCells
                                   ),
                      evcodes = TRUE, multi_query = FALSE,
                      sources = c("GO", "REAC", "MIRNA", "CORUM", "HP", "HPA", "WP"))
head(multi_gostres$result[,1:7])
```

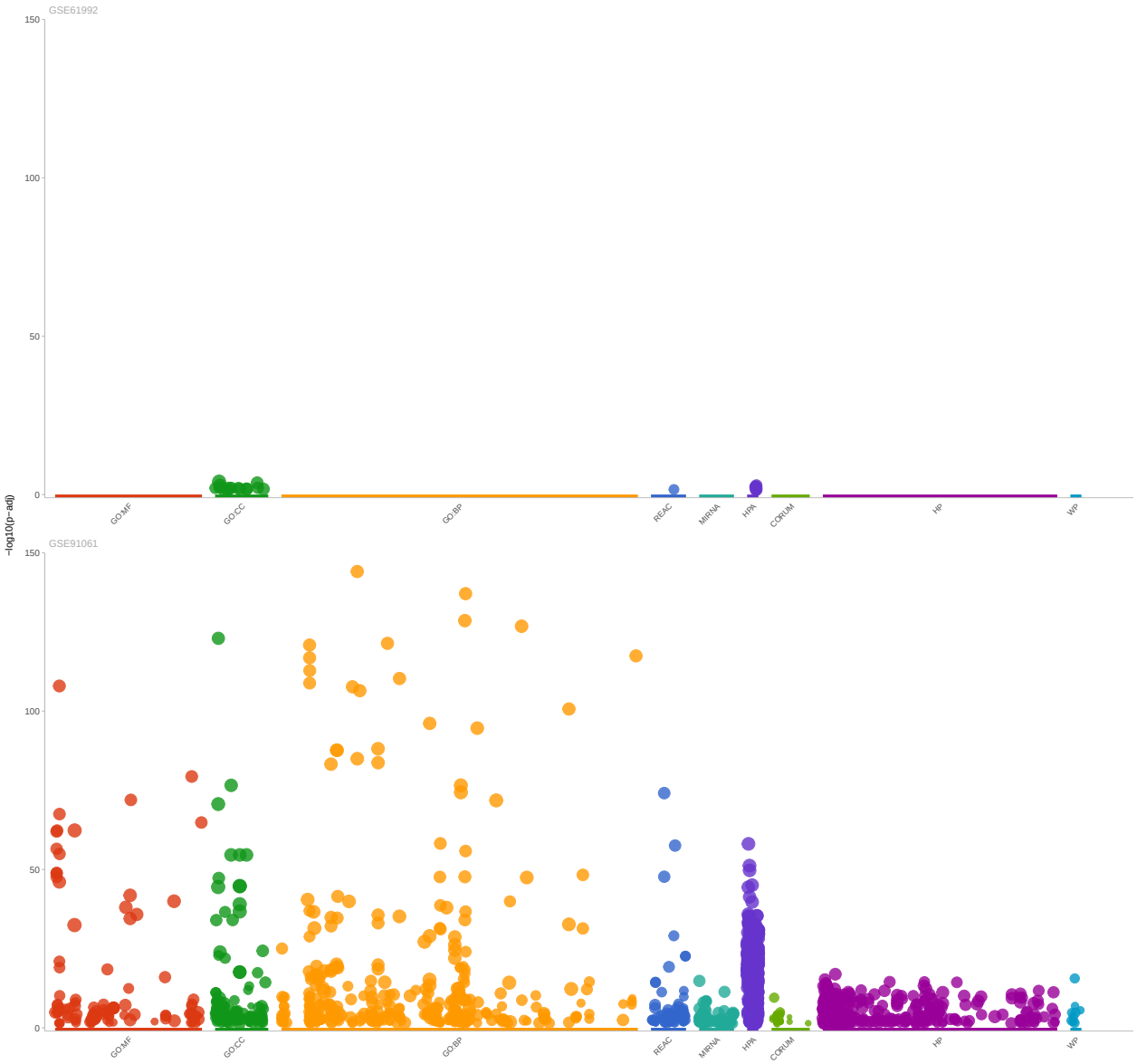
```
##      query significant      p_value term_size query_size intersection_size
## 1 GSE61992        TRUE 7.923024e-05     12345         81               67
## 2 GSE61992        TRUE 1.870319e-04      2203         81               24
## 3 GSE61992        TRUE 1.286308e-03      5487         81               39
## 4 GSE61992        TRUE 7.012270e-03      1194         81               15
## 5 GSE61992        TRUE 7.432560e-03      1200         81               15
## 6 GSE61992        TRUE 8.370056e-03      1521         81               17
##   precision
## 1 0.8271605
## 2 0.2962963
## 3 0.4814815
## 4 0.1851852
## 5 0.1851852
```

```
## 6 0.2098765
```

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_DendriticCells_Only_geneset_", unique(
                sep = "\t", quote = F, row.names = F))
```

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



Proporciones de los pathways para cada estudio:

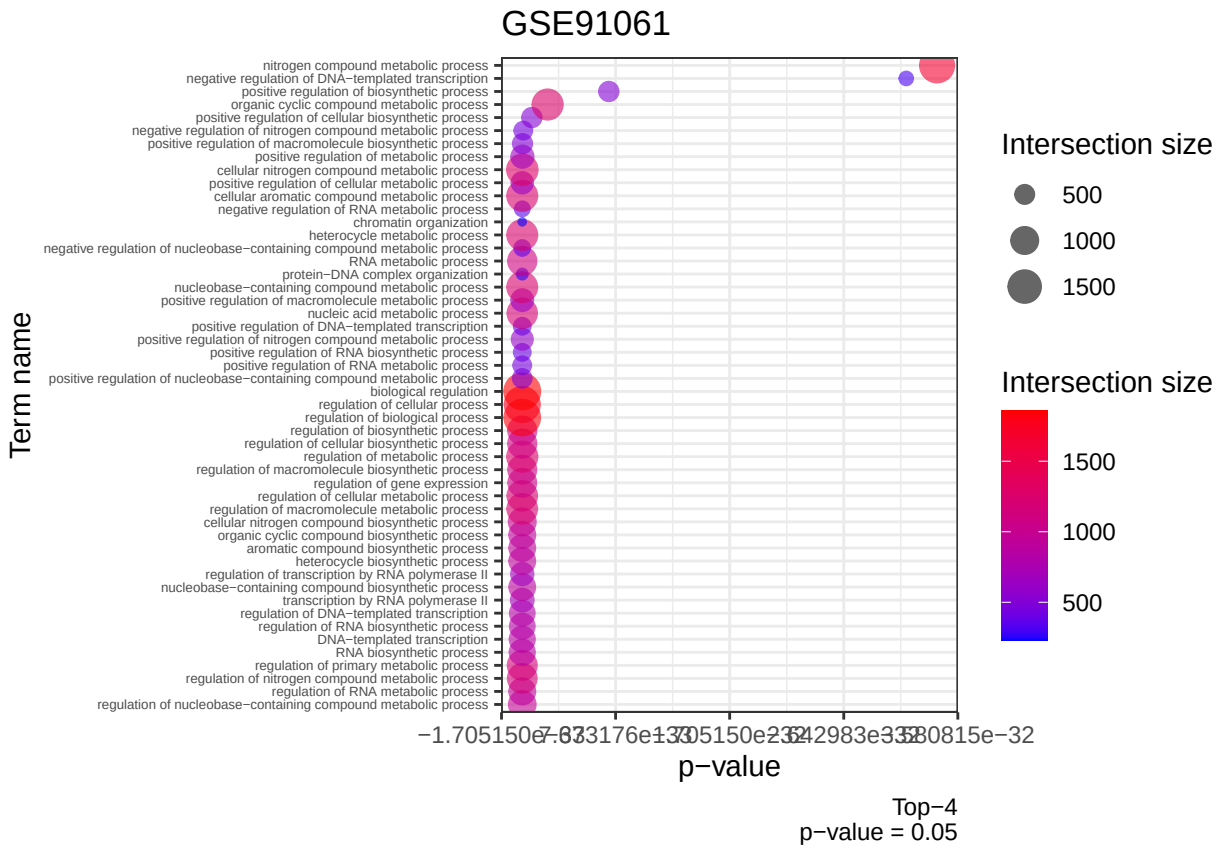
```
df <- as.data.frame(multi_gostres$result[,1:13])
table(df$query)
```

```
##
## GSE61992 GSE91061
##       25     1409
```

```
write.table(gem, file = "./Treatment_comparisons/gProfiler_DendriticCells_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")
```

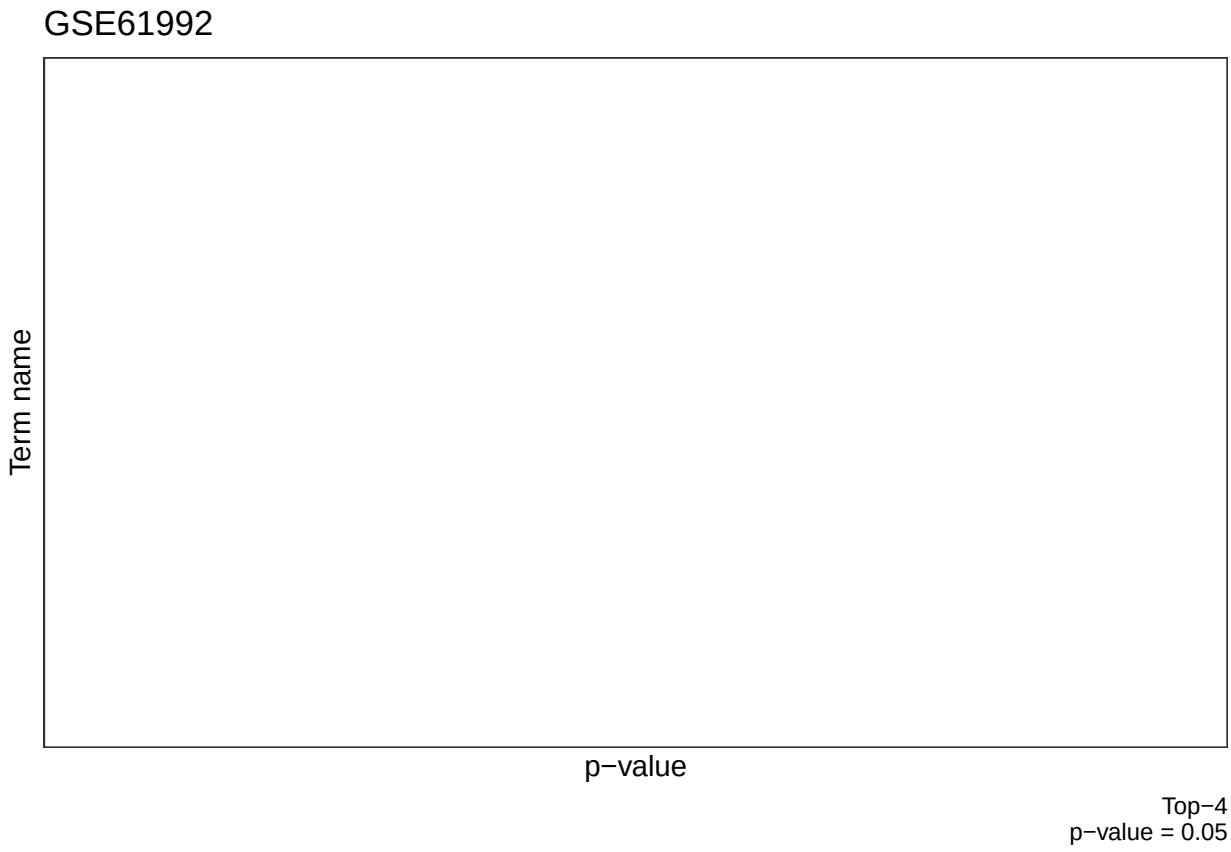
```
}
df2 <- df[df$source == "GO:BP",]
df2 <- df2[1:50,]
plot_gobps(study = "GSE91061")
```




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

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

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



6 Máximo solapamiento

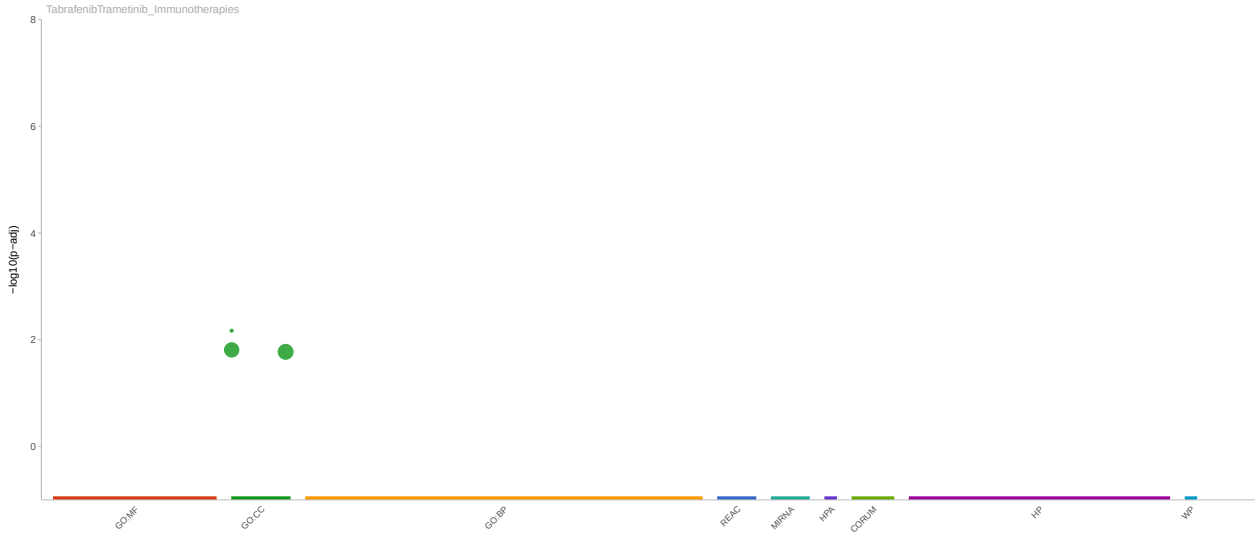
En esta sección analizo los DEGs que más se comparten entre los estudios para observar los mecanismos que pudieran haber en común entre los tratamientos.

Primero realizo el pathway multianálisis:

```
multi_gostres <- gost(query = list("TabrafenibTrametinib_Immunotherapies" = jVenn_DendriticCells$`GSE91
),
                      evcodes = TRUE, multi_query = FALSE,
                      sources = c("GO", "REAC", "MIRNA", "CORUM", "HP", "HPA", "WP"))
head(multi_gostres$result[,1:7])
```

```
##                                query significant     p_value term_size
## 1 TabrafenibTrametinib_Immunotherapies        TRUE 0.006753903        16
## 2 TabrafenibTrametinib_Immunotherapies        TRUE 0.015485312        24
## 3 TabrafenibTrametinib_Immunotherapies        TRUE 0.016825261        25
##   query_size intersection_size precision
## 1         15                2 0.1333333
## 2         15                2 0.1333333
## 3         15                2 0.1333333
```

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



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", "G
                file = paste0("./Treatment_comparisons/gProfiler_DendriticCells_maxOverlap_geneset_", u
                sep = "\t", quote = F, row.names = F))
```

Al mirar los procesos por cada comparación salieron demasiados para poder visualizarlos bien, por lo que decidí restringir a los procesos GO-BPs únicamente y guardarlos en pdf.

```
#unique(gem$query)

gem2 <- gem[grep("GO",gem$GO.ID),]
plot_gobps <- function(query, cutoff = 0.05) {
gem2 <- gem2[gem2$FDR <= cutoff,]
gem2_1 <- gem2[gem2$query == query, ]
ggplot(gem2_1, aes(FDR, Description)) + ggtitle(query) +
  geom_point(aes(FDR, Description, colour = Genes, size = Genes), alpha = 0.6)+
  theme_bw() + theme(axis.text.y.left = element_text(size = 5)) +
  labs(caption = paste0("FDR = ",cutoff))+
  scale_color_gradient(low="blue", high="red")
}

pdf(file = paste0("./Treatment_comparisons/gProfiler_DendriticCells_maxOverlap_TabrafenibTrametinib_Immu
plot_gobps("TabrafenibTrametinib_Immunotherapies")
dev.off()
```

```
## pdf
##   2
```

```
df <- as.data.frame(multi_gostres$result[,1:13])
df2 <- df[df$source == "GO:BP",]
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_s
  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 = "TabrafenibTrametinib_Immunotherapies")
```