Gene Enrichment Report

Ehsan Zangene

2025-06-09

Contents

Materials and Methods

Gene List Curation

Two gene sets were curated for functional enrichment analysis. The first consisted of 33 *Caenorhabditis elegans* genes, verified through a targeted RNAi screening and identified by their WormBase Gene IDs. The second was a list of 73 corresponding human orthologs provided by the HGNC symbols, curated from a cross-species homology assessment.

Functional Annotation and Enrichment Analysis

Functional enrichment analyses were performed using R (v4.3.1) with the `clusterProfiler` (v4.8.1), `org.Hs.eg.db`, `org.Ce.eg.db`, and `ReactomePA` (v1.44.0) Bioconductor packages. WormBase Gene IDs were mapped to Entrez Gene IDs using `bitr()` with the `org.Ce.eg.db` annotation. Similarly, human gene symbols were converted to Entrez IDs using `org.Hs.eg.db`.

For both species, Gene Ontology (GO) enrichment was conducted separately for the Biological Process (BP) ontology, using `enrichGO()` with Benjamini-Hochberg correction (adjusted $p < 0.05$). Human orthologs were additionally subjected to Molecular Function (MF) GO analysis and Reactome pathway enrichment via `enrichGO()` and `enrichPathway()` respectively. KEGG pathway analysis was attempted but yielded no significant enrichments for either species under the set cutoff thresholds.

Visualization

Top enriched categories (up to 20 terms) were visualized using `dotplot()` from the `enrichplot` package. All enrichment terms reported in the figures met a false discovery rate (FDR) threshold of $q < 0.2$. Y-axis text was scaled and spaced to improve legibility.

1. Load Required Libraries

```
library(readxl)
library(clusterProfiler)
library(org.Hs.eg.db)
library(org.Ce.eg.db)
library(enrichplot)
library(ReactomePA)
library(ggplot2)
library(dplyr)
```

2. Input Gene Lists

2.1 Human Orthologs

```
file_path_human <- "./30-05-25/List of human orthologs_RNAi screen.xlsx"
human_df <- read_excel(file_path_human)
symbol_col <- grep("symbol", names(human_df), ignore.case = TRUE, value = TRUE)
human_df <- human_df %>% rename(GeneSymbolColumn = all_of(symbol_col))
human_symbols <- human_df$GeneSymbolColumn
human_map <- bitr(human_symbols, fromType="SYMBOL", toType="ENTREZID", OrgDb=org.Hs.eg.db)
human_entrez <- unique(human_map$ENTREZID)
```

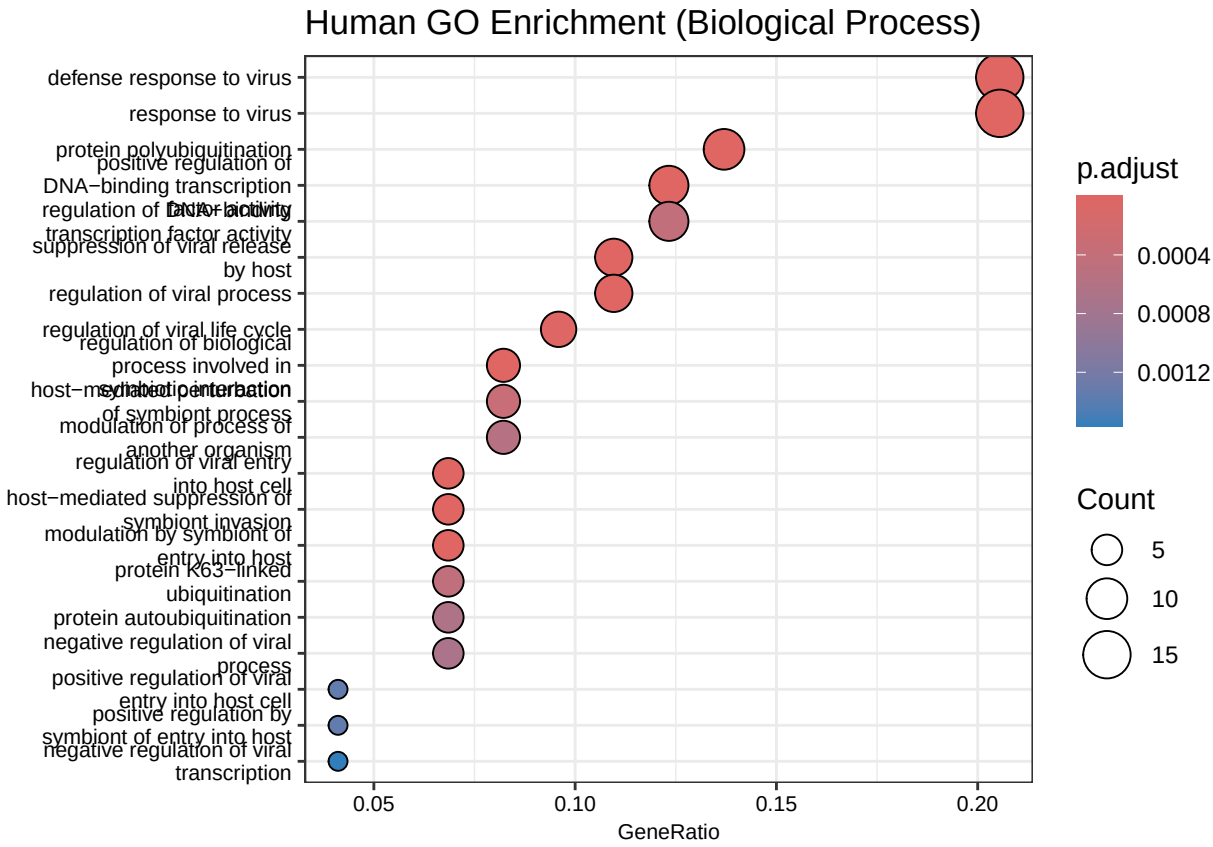
2.2 Worm Genes

```
file_path_worm <- "./06-2024/List of verified genes.xlsx"
worm_df <- read_excel(file_path_worm)
wb_col <- grep("WormBase Gene", colnames(worm_df), ignore.case = TRUE, value = TRUE)
worm_df <- worm_df %>% rename(GeneID = all_of(wb_col))
worm_map <- bitr(worm_df$GeneID, fromType = "WORMBASE", toType = "ENTREZID", OrgDb = org.Ce.eg.db)
worm_entrez <- unique(worm_map$ENTREZID)
```

3. Enrichment Analysis

3.1 Human GO: Biological Process

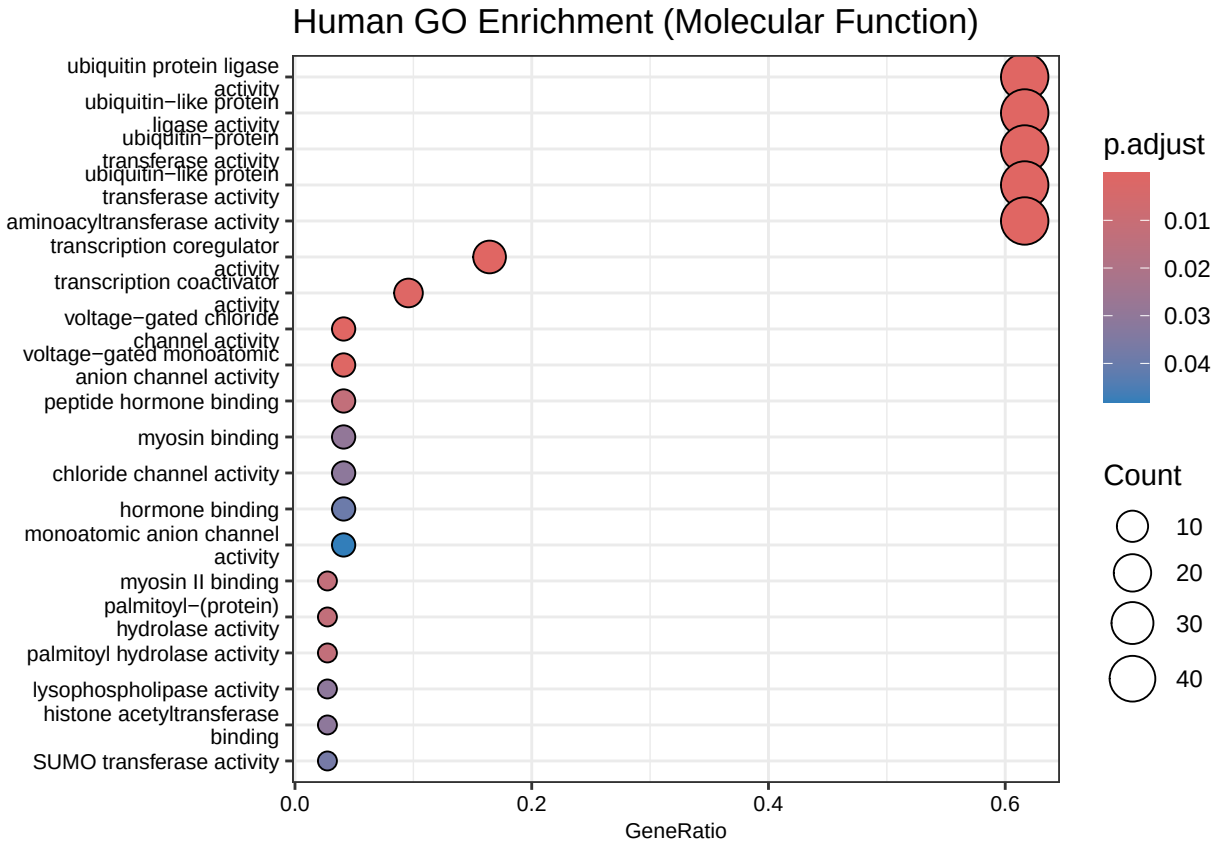
```
ego_bp <- enrichGO(gene = human_entrez, OrgDb = org.Hs.eg.db, keyType = "ENTREZID",
                   ont = "BP", pAdjustMethod = "BH", pvalueCutoff = 0.05,
                   qvalueCutoff = 0.2, readable = TRUE)
dotplot(ego_bp, showCategory = 20, font.size = 8) + ggtitle("Human GO Enrichment (Biological Process)")
```



3.2 Human GO: Molecular Function

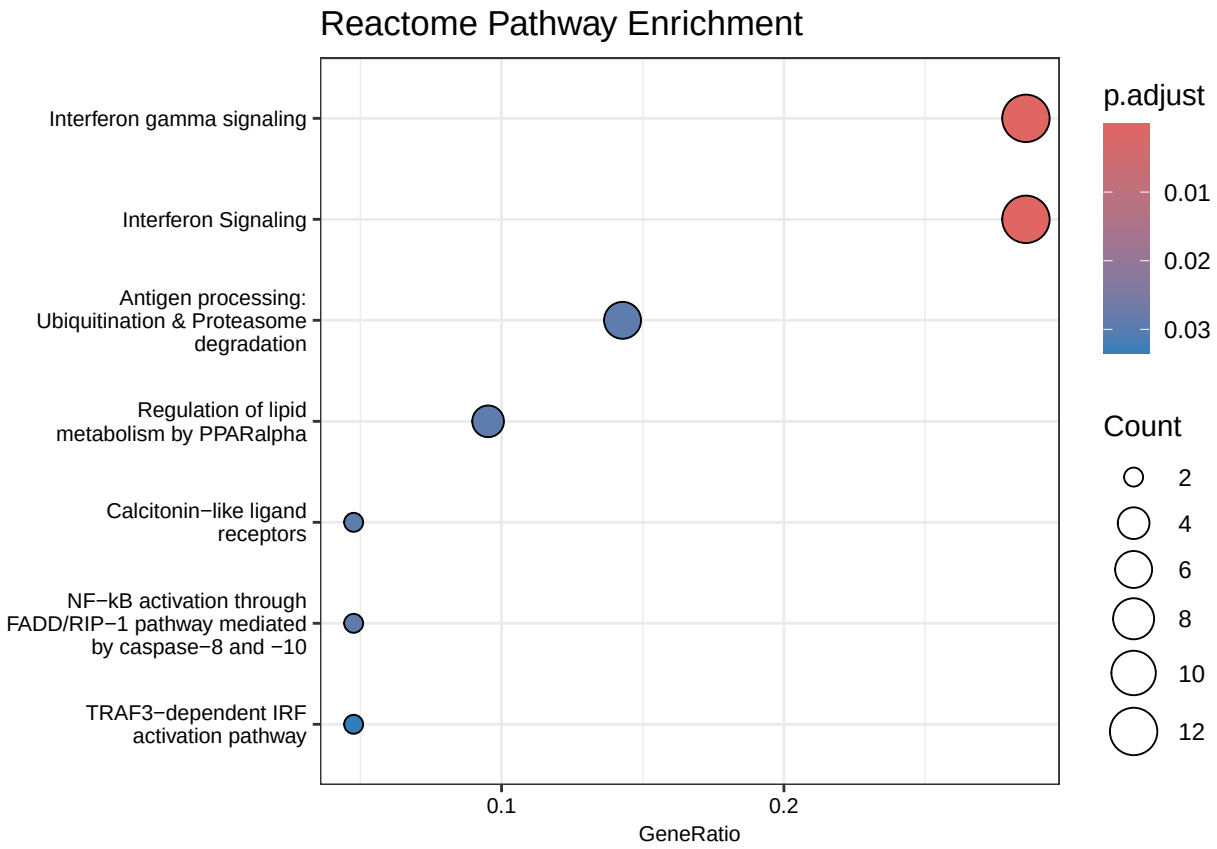
```
ego_mf <- enrichGO(gene = human_entrez, OrgDb = org.Hs.eg.db, keyType = "ENTREZID",
                   ont = "MF", pAdjustMethod = "BH", pvalueCutoff = 0.05,
```

```
                    qvalueCutoff = 0.2, readable = TRUE)
dotplot(ego_mf, showCategory = 20, font.size = 8) + ggtitle("Human GO Enrichment (Molecular Function)")
```



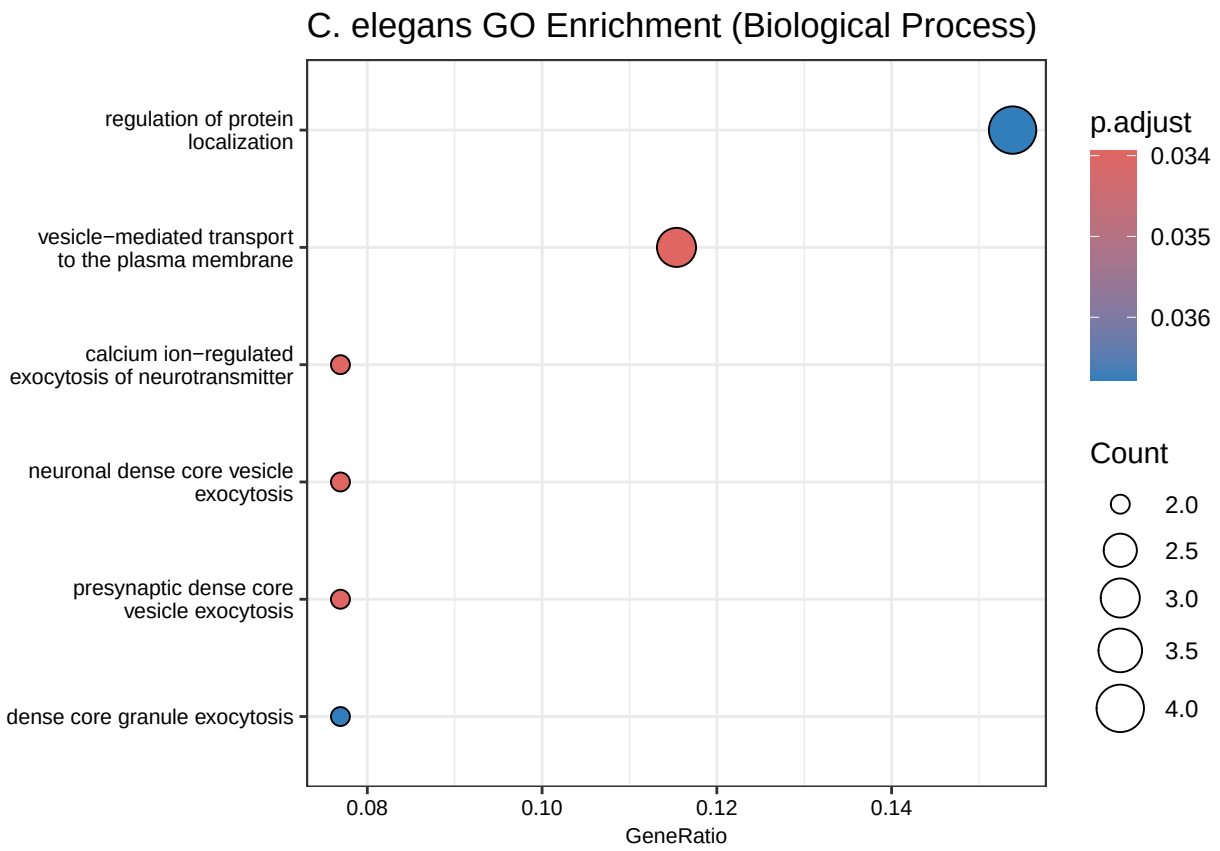
3.3 Human Reactome Pathways

```
ereactome <- enrichPathway(gene = human_entrez, organism = "human",
                           pvalueCutoff = 0.05, readable = TRUE)
dotplot(ereactome, showCategory = 20, font.size = 8) + ggtitle("Reactome Pathway Enrichment")
```



3.4 C. elegans GO: Biological Process

```
ego_ce_bp <- enrichGO(gene = worm_entrez, OrgDb = org.Ce.eg.db, keyType = "ENTREZID",
                      ont = "BP", pAdjustMethod = "BH", pvalueCutoff = 0.05,
                      qvalueCutoff = 0.2, readable = TRUE)
dotplot(ego_ce_bp, showCategory = 20, font.size = 8) + ggtitle("C. elegans GO Enrichment (Biological Pr
```



4. Export Results (Optional)

```
write.csv(as.data.frame(ego_bp), "Human_GO_BP.csv", row.names = FALSE)
write.csv(as.data.frame(ego_mf), "Human_GO_MF.csv", row.names = FALSE)
write.csv(as.data.frame(ereactome), "Human_Reactome.csv", row.names = FALSE)
write.csv(as.data.frame(ego_ce_bp), "Worm_GO_BP.csv", row.names = FALSE)
```

Conclusions

Our functional enrichment analysis revealed that the human orthologs of *C. elegans* genes identified in the RNAi screen were significantly associated with antiviral defense mechanisms and protein ubiquitination pathways. The most enriched biological processes included "defense response to virus" and "regulation of viral entry into host cells", primarily driven by TRIM family genes. Molecular function analysis underscored ubiquitin ligase activity as a key functional hallmark.

Reactome pathway analysis further supported these findings, highlighting "Interferon gamma signaling" and "Ubiquitination & Proteasome degradation" as the most prominent enriched pathways. In contrast, KEGG enrichment yielded no significant terms, suggesting a potential bias in pathway representation or limited coverage of TRIM-centric antiviral pathways in KEGG.

In the *C. elegans* dataset, GO enrichment for biological processes indicated roles in "dense core vesicle exocytosis" and "calcium ion-regulated neurotransmitter release", suggesting involvement in neurosecretory

or synaptic functions. These findings provide mechanistic insight into the biological roles of candidate genes from the screen and their conserved relevance across species.