############################################################

# Title: Gene Set Enrichment with Enrichr

# Description:

# This script reads a list of genes from an Excel file and

# performs enrichment analysis using Enrichr.

# Results are saved as CSV files (one per database).

# ---- Install required packages if missing ----

if (!requireNamespace("enrichR", quietly = TRUE)) install.packages("enrichR")

if (!requireNamespace("readxl", quietly = TRUE)) install.packages("readxl")

library(enrichR) # For connecting to the Enrichr API

library(readxl) # For reading Excel files

# ---- Step 1: Read gene list from Excel file ----

# Replace "path/to/List\_of\_genes.xlsx" with your own file path.

# The Excel file should contain a single column of gene symbols, no header.

file\_path <- "path/to/List\_of\_genes.xlsx"

genes <- read\_excel(file\_path, col\_names = FALSE)

gene\_list <- as.character(genes[[1]])

# ---- Step 2: Define databases of interest ----

# These names must match EXACTLY what Enrichr supports (check with listEnrichrDbs()).

requested\_dbs <- c(

"WikiPathways\_2024\_Human",

"GO\_Cellular\_Component\_2023",

"GO\_Biological\_Process\_2023",

"GO\_Molecular\_Function\_2023",

"ChEA\_2022",

"DrugMatrix",

"Jensen\_TISSUES",

"Jensen\_COMPARTMENTS",

"HMDB\_Metabolites"

)

# ---- Step 3: Validate requested databases ----

# Compare requested names against available Enrichr libraries.

available\_dbs <- listEnrichrDbs()$libraryName

not\_found <- setdiff(requested\_dbs, available\_dbs)

if (length(not\_found) > 0) {

message("⚠️ The following requested databases are NOT available in Enrichr right now:")

print(not\_found)

}

usable\_dbs <- intersect(requested\_dbs, available\_dbs)

if (!length(usable\_dbs)) stop("❌ None of the requested databases matched. Please update requested\_dbs.")

# ---- Step 4: Run enrichment analysis ----

# This sends the gene list to Enrichr and retrieves results for the selected databases.

enrich\_results <- enrichr(gene\_list, usable\_dbs)

# ---- Step 5: Save results ----

# Each database result will be saved as a separate CSV file.

# Files will be created in the current working directory by default.

for (db in names(enrich\_results)) {

safe\_db <- gsub("[^A-Za-z0-9\_]", "\_", db) # sanitize database name for filenames

out\_file <- paste0("Enrichr\_", safe\_db, ".csv")

write.csv(enrich\_results[[db]], out\_file, row.names = FALSE)

message("Saved: ", out\_file)

}

message("Enrichment finished. Results are saved as CSV files in the working directory.")