# --- 第1步: 加载必要的包 ---

# 如果没有安装，请先运行: BiocManager::install("clusterProfiler")

library(clusterProfiler)

# 如果没有安装，请先运行: install.packages("tidyverse")

library(tidyverse) # 或者只加载 library(ggplot2)

library(readxl)

# --- 第2步: 准备数据 ---

# 1. 准备 GO 注释文件

# 确保你的 "your\_go\_annotation\_file.csv" 文件在正确的工作目录下

# 文件格式: 第一列是GO ID, 第二列是基因ID

go\_annotation <- read.csv("/home/data/t210549/liuhuacheng/wrky注释/全基因组/extracted\_gene\_go\_id.csv", header = FALSE)

colnames(go\_annotation) <- c("GO\_ID", "Entrez\_ID")

# --- 第2步: 读取并准备差异基因列表 ---

deg\_file\_path <- "/home/data/t210549/liuhuacheng/wrky注释/全基因组/各组差异基因/200vsCK.xlsx"

deg\_df <- read\_excel(deg\_file\_path, col\_names = FALSE)

deg\_genes\_vector <- deg\_df[[1]]

gene\_list <- deg\_genes\_vector[!is.na(deg\_genes\_vector)]

cat("成功提取差异基因列表，共", length(gene\_list), "个基因。\n")

# --- 第3步: 进行富集分析 ---

# 使用 enricher 函数进行自定义富集分析

# 对于巨菌草，readable=TRUE 可能无效，可以设为 FALSE

ego <- enricher(gene = gene\_list,

TERM2GENE = go\_annotation,

pvalueCutoff = 0.05,

qvalueCutoff = 0.05,

)

# --- 第4步: 可视化 (在富集分析完成后) ---

# 1. 读取包含GO ID和描述的注释文件

go\_annotation\_full <- read.csv("/home/data/t210549/liuhuacheng/wrky注释/全基因组/all\_go\_long.csv", header = TRUE)

# 2. 创建一个从 GO ID 到 GO 术语描述的映射

# 选择 go\_id 和 term 两列，并去除重复项以确保唯一性

goid\_to\_term\_map <- go\_annotation\_full %>%

select(go\_id, term) %>%

distinct(go\_id, .keep\_all = TRUE)

# 3. 将这个映射应用到富集分析结果中

# ego@result $ID 列包含富集到的GO ID

ego@result <- ego@result %>%

left\_join(goid\_to\_term\_map, by = c("ID" = "go\_id"))

# 4. 【关键步骤】将原来的 ID 列（GO ID）替换为 term 列（GO 描述）

# 这样 dotplot 就会自动使用描述作为Y轴标签

ego@result$Description <- ego@result$term

pdf(file = "/home/data/t210549/liuhuacheng/wrky注释/GO\_Enrichment\_Bubble\_Plot.pdf", width = 10, height = 8)

# 6. 现在使用 dotplot 绘图，Y轴将显示GO描述

p <- dotplot(ego, showCategory = 20)

# 显示图片

print(p)

dev.off()

# 保存图片

#ggsave("/home/data/t210549/liuhuacheng/wrky注释/200vsCK\_GO\_Enrichment\_Bubble\_Plot.png", plot = p, width = 10, height = 8, dpi = 300)

print("富集分析和绘图完成！图片已保存为 GO\_Enrichment\_Bubble\_Plot.pdf")

====================================条形图==================================

# --- 第1步: 加载必要的包 ---

# 如果没有安装，请先运行: BiocManager::install("clusterProfiler")

library(clusterProfiler)

# 如果没有安装，请先运行: install.packages("tidyverse")

library(tidyverse) # 或者只加载 library(ggplot2)

library(readxl)

# --- 第2步: 准备数据 ---

# 1. 准备 GO 注释文件

# 确保你的 "your\_go\_annotation\_file.csv" 文件在正确的工作目录下

# 文件格式: 第一列是GO ID, 第二列是基因ID

go\_annotation <- read.csv("/home/data/t210549/liuhuacheng/wrky注释/全基因组/extracted\_gene\_go\_id.csv", header = FALSE)

colnames(go\_annotation) <- c("GO\_ID", "Entrez\_ID")

# --- 第2步: 读取并准备差异基因列表 ---

deg\_file\_path <- "/home/data/t210549/liuhuacheng/wrky注释/全基因组/各组差异基因/200vsCK.xlsx"

deg\_df <- read\_excel(deg\_file\_path, col\_names = FALSE)

deg\_genes\_vector <- deg\_df[[1]]

gene\_list <- deg\_genes\_vector[!is.na(deg\_genes\_vector)]

cat("成功提取差异基因列表，共", length(gene\_list), "个基因。\n")

# --- 第3步: 进行富集分析 ---

# 使用 enricher 函数进行自定义富集分析

# 对于巨菌草，readable=TRUE 可能无效，可以设为 FALSE

ego <- enricher(gene = gene\_list,

TERM2GENE = go\_annotation,

pvalueCutoff = 0.05,

qvalueCutoff = 0.05,

)

# --- 第4步: 可视化 (在富集分析完成后) ---

# 1. 读取包含GO ID和描述的注释文件

go\_annotation\_full <- read.csv("/home/data/t210549/liuhuacheng/wrky注释/全基因组/all\_go\_long.csv", header = TRUE)

# 2. 创建一个从 GO ID 到 GO 术语描述的映射

# 选择 go\_id 和 term 两列，并去除重复项以确保唯一性

goid\_to\_term\_map <- go\_annotation\_full %>%

select(go\_id, term) %>%

distinct(go\_id, .keep\_all = TRUE)

# 3. 将这个映射应用到富集分析结果中

# ego@result $ID 列包含富集到的GO ID

ego@result <- ego@result %>%

left\_join(goid\_to\_term\_map, by = c("ID" = "go\_id"))

# 4. 【关键步骤】将原来的 ID 列（GO ID）替换为 term 列（GO 描述）

# 这样 barplot 就会自动使用描述作为Y轴标签

ego@result$Description <- ego@result$term

# 5. 【修改点】使用 barplot 函数绘制条形图

pdf(file = "/home/data/t210549/liuhuacheng/wrky注释/GO\_Enrichment\_Bar\_Plot.pdf", width = 10, height = 8)

# 【核心修改】只传递 ego 对象，不传递任何参数

p <- barplot(ego)

# 显示图片

print(p)

# 关闭PDF设备，完成保存

dev.off()

print("富集分析和条形图绘制完成！图片已保存为 GO\_Enrichment\_Bar\_Plot.pdf")

# --- 第4步: 可视化 (在富集分析完成后) ---

# 1. 读取包含GO ID和描述的注释文件

go\_annotation\_full <- read.csv("/home/data/t210549/liuhuacheng/wrky注释/全基因组/all\_go\_long.csv", header = TRUE)

# 2. 创建一个从 GO ID 到 GO 术语描述的映射

# 选择 go\_id 和 term 两列，并去除重复项以确保唯一性

goid\_to\_term\_map <- go\_annotation\_full %>%

select(go\_id, term) %>%

distinct(go\_id, .keep\_all = TRUE)

# 3. 将这个映射应用到富集分析结果中

# ego@result $ID 列包含富集到的GO ID

ego@result <- ego@result %>%

left\_join(goid\_to\_term\_map, by = c("ID" = "go\_id"))

# 4. 【关键步骤】将原来的 ID 列（GO ID）替换为 term 列（GO 描述）

# 这样 barplot 就会自动使用描述作为Y轴标签

ego@result$Description <- ego@result$term

# 5. 【修改点】使用 barplot 函数绘制条形图

pdf(file = "/home/data/t210549/liuhuacheng/wrky注释/GO\_Enrichment\_Bar\_Plot.pdf", width = 10, height = 8)

# 【核心修改】只传递 go\_res 对象，不传递任何参数

# 注意：由于 barplot 函数不支持直接传递 x 和 color 参数，我们将在后续使用 ggplot2 进行自定义

p <- barplot(Go\_count,

 x ="GeneRatio",

color ="p.adjust",

showCategory =10,

split="ontology")+

facet\_grid(ONTOLOGY~., scale='free')

# 显示图片

print(p)

# 关闭PDF设备，完成保存

dev.off()

print("富集分析和条形图绘制完成！图片已保存为 GO\_Enrichment\_Bar\_Plot.pdf")

plot <- barplot(go\_res, x ="GeneRatio", color ="p.adjust",#默认参数（x和color可以根据eG里面的内容更改）        showCategory =10,#显示前10        split="ONTOLOGY")+  facet\_grid(ONTOLOGY~., scale='free')