library(clusterProfiler)

library(tidyverse)

library(readxl)

library(ggh4x)

# --- 第1步: 准备数据 ---

go\_annotation\_path <- "/home/data/t210549/liuhuacheng/wrky注释/全基因组/extracted\_gene\_go\_id.csv"

go\_annotation <- read.csv(go\_annotation\_path, header = FALSE)

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

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2步: 进行富集分析 ---

ego <- enricher(gene = gene\_list,

TERM2GENE = go\_annotation,

pvalueCutoff = 0.05,

qvalueCutoff = 0.05)

# --- 第3步: 数据处理与筛选 ---

if (nrow(ego) == 0) {

stop("富集分析没有找到任何显著结果，请检查您的输入基因或pvalue/qvalue阈值。")

}

go\_annotation\_full\_path <- "/home/data/t210549/liuhuacheng/wrky注释/全基因组/all\_go\_long.csv"

go\_annotation\_full <- read.csv(go\_annotation\_full\_path, header = TRUE)

ego\_result\_with\_ontology <- as.data.frame(ego) %>%

left\_join(go\_annotation\_full, by = c("ID" = "go\_id"))

if (any(is.na(ego\_result\_with\_ontology$term) | is.na(ego\_result\_with\_ontology$ontology))) {

warning("部分富集到的GO ID在注释文件中未找到描述或ontology信息，将被移除。")

ego\_result\_with\_ontology <- ego\_result\_with\_ontology %>%

filter(!is.na(term) & !is.na(ontology))

}

# 确保每个GO术语只出现一次

Go\_count <- ego\_result\_with\_ontology %>%

distinct(term, ontology, .keep\_all = TRUE) %>%

select(GO\_term = term, Aspect = ontology, Count) %>%

mutate(Aspect = factor(Aspect, levels = c("BP", "MF", "CC"))) %>%

group\_by(Aspect) %>%

arrange(desc(Count)) %>%

slice\_head(n = 10) %>%

ungroup() %>%

arrange(Aspect, desc(Count))

# 检查数据

cat("最终用于绘图的GO术语数量:", nrow(Go\_count), "\n")

print(table(Go\_count$Aspect))

# 计算Count轴的最大值，避免太空旷

# 根据数据动态计算合适的x轴上限

max\_count <- max(Go\_count$Count)

# 设置x轴上限为最大值的1倍，但不超过最大值的下一个整数十位数

x\_max <- min(max\_count \* 1, ceiling(max\_count/10)\*10 + 10)

cat("Count最大值:", max\_count, "，X轴上限设置为:", x\_max, "\n")

# 准备绘图数据 - 根据例图调整

Go\_count\_plot <- Go\_count %>%

mutate(

# 对GO术语进行换行处理，宽度根据例图调整

GO\_term\_wrapped = str\_wrap(GO\_term, width = 40),

# 创建显示标签，只包含GO术语

display\_label = GO\_term\_wrapped

) %>%

# 设置因子水平以确保正确的排序

mutate(display\_label = factor(display\_label, levels = rev(unique(display\_label))))

# --- 第4步: 优化绘图，自动调整Count轴范围 ---

plot\_final <- Go\_count\_plot %>%

ggplot(aes(x = Count, y = display\_label, fill = Aspect)) +

geom\_col(width = 0.8) +

geom\_text(aes(label = Count), hjust = -0.2, size = 3.5, fontface = "bold") +

# 动态调整x轴范围，避免太空旷

scale\_x\_continuous(

expand = expansion(mult = c(0, 0.15)), # 减少右侧扩展

limits = c(0, x\_max), # 使用动态计算的x轴上限

breaks = function(x) pretty(x, n = min(5, length(unique(Go\_count$Count)))) # 动态设置刻度

) +

scale\_fill\_manual(

values = c("BP" = "#74c476", "MF" = "#41b6c4", "CC" = "#9e9ac8"),

name = "Ontology"

) +

# 使用分面，将标签放在右侧

facet\_grid(Aspect ~ ., scales = "free\_y", space = "free\_y") +

ggtitle(label = "GO annotation") + # 根据例图修改标题

labs(x = "Count", y = NULL) + # 根据例图修改x轴标签

theme\_bw() +

theme(

# 分面标签设置 - 根据例图调整

strip.background = element\_rect(fill = "white", color = "black"),

strip.text.y = element\_text(

size = 12,

face = "bold",

color = "black",

angle = 0,

hjust = 0.5,

margin = margin(3, 3, 3, 3)

),

strip.placement = "outside",

# 标题设置 - 根据例图调整

plot.title = element\_text(

size = 16,

hjust = 0.5,

face = "bold",

margin = margin(b = 10)

),

# 坐标轴设置 - 根据例图调整

axis.text.x = element\_text(size = 10, face = "bold"),

axis.text.y = element\_text(

size = 9,

hjust = 1,

color = "black",

margin = margin(r = 5)

),

axis.title.x = element\_text(

size = 12,

face = "bold",

margin = margin(t = 5)

),

# 隐藏图例（因为已经有分面标签）

legend.position = "none",

# 图表边框和网格 - 根据例图调整

panel.border = element\_rect(linewidth = 0.8, color = "black"),

panel.spacing = unit(0.2, "lines"), # 减少间距

panel.grid.major.x = element\_line(color = "grey90", linewidth = 0.3),

panel.grid.minor = element\_blank(),

panel.grid.major.y = element\_blank(),

# 调整边距，确保右侧有足够空间显示分面标签

plot.margin = margin(t = 0.5, r = 1, b = 0.5, l = 1.5, unit = "cm")

)

# 显示图形

print(plot\_final)

# 动态调整保存高度

plot\_height <- max(8, nrow(Go\_count\_plot) \* 0.4)

ggsave("/home/data/t210549/liuhuacheng/wrky注释/全基因组/200vsCKGO\_Enrichment\_Final.pdf", plot\_final,

width = 12, height = plot\_height, dpi = 300)