# ============================================================

# 【统计每个比较组中 WRKY 差异表达情况（自动补齐所有组）】

# ============================================================

library(dplyr)

library(readxl)

library(tidyr)

# ---------- 1. 文件路径 ----------

wrky\_file <- "/home/data/t210549/liuhuacheng/wrky注释/全基因组/火山图/wrky分组.xlsx"

deg\_file <- "/home/data/t210549/liuhuacheng/wrky注释/全基因组/DESeq2\_all\_pairwise\_results.csv"

# ---------- 2. 读取文件 ----------

wrky <- read\_excel(wrky\_file)

deg <- read.csv(deg\_file, header = TRUE)

# ---------- 3. 检查必要列 ----------

required\_deg\_cols <- c("gene\_id", "log2FoldChange", "padj", "comparison")

missing\_cols <- setdiff(required\_deg\_cols, colnames(deg))

if (length(missing\_cols) > 0) {

stop("❌ DEG 文件缺少以下列：", paste(missing\_cols, collapse = ", "))

}

if (!"gene\_id" %in% colnames(wrky)) {

stop("❌ WRKY 表中未找到 'gene\_id' 列。")

}

# ---------- 4. 筛选显著差异基因 ----------

deg\_sig <- deg %>%

filter(abs(log2FoldChange) >= 1, padj < 0.05)

# ---------- 5. 与 WRKY 基因交叉 ----------

wrky\_deg <- inner\_join(deg\_sig, wrky, by = "gene\_id")

# ---------- 6. 提取所有比较组 ----------

all\_groups <- deg %>%

distinct(comparison)

# ---------- 7. 按 comparison 分组统计，并补齐 ----------

wrky\_summary <- wrky\_deg %>%

group\_by(comparison) %>%

summarise(

total\_WRKY = n(),

up\_WRKY = sum(log2FoldChange > 0),

down\_WRKY = sum(log2FoldChange < 0)

) %>%

right\_join(all\_groups, by = "comparison") %>%

mutate(

total\_WRKY = replace\_na(total\_WRKY, 0),

up\_WRKY = replace\_na(up\_WRKY, 0),

down\_WRKY = replace\_na(down\_WRKY, 0)

) %>%

arrange(comparison)

# ---------- 8. 输出结果 ----------

cat("✅ 每个比较组的 WRKY 差异表达统计（包括无差异组）：\n")

print(wrky\_summary)

# 输出目录

output\_dir <- "/home/data/t210549/liuhuacheng/wrky注释/全基因组/wrky差异基因"

# 汇总表

write.csv(

wrky\_summary,

file.path(output\_dir, "WRKY\_DEG\_summary\_by\_comparison.csv"),

row.names = FALSE

)

cat("✅ 统计表已保存为 WRKY\_DEG\_summary\_by\_comparison.csv\n")

# 明细表

write.csv(

wrky\_deg,

file.path(output\_dir, "WRKY\_DEG\_detail\_all.csv"),

row.names = FALSE

)

cat("✅ 所有差异 WRKY 基因明细已保存为 WRKY\_DEG\_detail\_all.csv\n")