

GO富集分析（R studio）

1.安装与加载

- 1.1 安装R studio软件
- 1.2 加载R包（DESeq2, ClusterProfiler, org.Hs.eg.db等）



2.数据准备

- 1.1 输入差异基因DEGs列表
- 1.2 提取显著差异基因（ $FDR < 0.05$ ）
- 1.3 基因ID转换（SYMBOL→Entrez）
- 1.4 GO注释



3.GO富集分析

enrichGO函数分析，包含“BP”、“MF”和“CC”



4.结果可视化

条形图、点图和富集网络图等可视化结果



5.结果保存

差异数据存为.csv文件，可视化结果存为.pdf文件

常见问题分析

1. R Studio分析读取不到基因

可能为基因ID类型问题，读取差异基因文件时基因是symbol格式，需通过bitr函数进行转换。

2. 数据分析过程中Count matrix中含有非整数

DESeq2只接受整数表达计数，应检查数据是否为TPM或FPKM。

3. 没有显著富集的结果

可能是样本量小或者标准设定过严，可调整p值阈值 <0.1 ，或放宽 \log_2FC 阈值,另外检查基因ID是否正确，考虑使用更大的背景基因集。

4. 富集分析的结果太多/太少

调整pvalueCutoff和qvalueCutoff参数，使用filter函数筛选结果。

5. 富集与预期生物过程不符合

可能原因是背景基因设置不正确，设置universe = all_genes_entrez参数以指定背景基因集。