

简介

该脚本采用hdWGCNA包实现了基于单细胞RNA-seq数据的共表达网络分析流程。共表达网络分析可以揭示基因之间的共表达模式,建立基因之间的拓扑联系,进而预测基因功能模块和找到关键的中心基因。

脚本的主要分析步骤包括:

- 数据准备:**根据指定的方法选择保留的变量基因,构建元细胞以减少细胞异质性的影响。
- 软阈值选择:**测试不同软阈值参数,筛选出能够使网络满足无尺度网络特性的最佳软阈值。
- 构建共表达网络:**基于最佳软阈值参数构建基因共表达网络,网络连接性采用TOM算法。
- 模块检测:**对共表达网络进行模块检测,识别出基因功能模块,即高度内部连接而低度互相连接的基因模块。
- 模块特征分析:**计算每个模块的特征向量和特征基因,揭示每个模块的主要功能。
- 模块关联分析:**模块之间的相关性分析,识别功能相关的模块。
- 关键基因分析:**筛选每个模块的中心或枢纽基因,这些基因在维持模块功能中的重要性最高。

通过该流程分析,可以系统地解析单细胞RNA-seq数据中的共表达模块和关键调控基因,为后续的功能富集分析奠定基础,也可以与其他组学数据整合,建立更全面的细胞功能调控网络。

参数说明

参数说明

- a/--gene_select_approach: 基因选择方法
- b/--fraction: 最小细胞表达阈值
- c/--input_gene_list: 基因列表路径
- d/--input_data: 输入的表达式矩阵路径
- e/--group_by: 元细胞分组的metadata
- f/--ident_group: 元细胞Seurat对象的Idents
- g/--k: 构建元细胞的k最近邻参数
- h/--max_shared: 元细胞最大共享细胞数
- i/--min_cells: 构建元细胞的最小细胞数
- j/--outdir: 输出目录路径
- k/--group_by_vars: 调和的分组变量

输出目录结构

outdir/
细胞群1/

```
figure/
  SoftPowers.png
  SoftPowers.pdf
  kMEs.png
  kMEs.pdf
  featureplot.png
  featureplot.pdf
  seurat_dotplot.png
  seurat_dotplot.pdf
  Correlations_plot.png
  Correlations_plot.pdf
  ModuleNetworks/
    HubGeneNetworkPlot.png
    HubGeneNetworkPlot.pdf
    ModuleUMAPPlot.png
    ModuleUMAPPlot.pdf
table/
  wgcna_modules.csv
  hMEs.csv
  MEs.csv
  modules.csv
  hub_genes.csv
TOM/
  wgcna.tom.csv
hdWGCNA_object.rds
细胞群2/
...
...
```

主要输出结果

- **figure** 目录:包含模块分析的可视化结果图,如软阈值分析图、模块特征表达图、模块网络图等
- **table** 目录:包含模块分析的结果表格,如模块-基因对应表、模块特征值矩阵、枢纽基因表等
- **TOM** 目录:包含基于TOM算法计算的基因网络连接性矩阵
- **hdWGCNA_object.rds** :保存R对象,可用于后续可视化和分析

分析意义

该分析通过构建基因共表达网络,实现了对单细胞RNA-seq数据的模块检测功能富集分析。可以系统解析复杂的基因调控网络,发现功能模块和关键基因,为研究细胞发育和功能提供重要信息。