

单细胞转录组学研究分析报告

项目名称：单细胞转录组学研究分析报告

项目编号：xxx

分析人员：冯刚毅

审核人员：

报告日期：2025年04月23日

报告单位：西安浩瑞基因技术有限公司

单细胞转录组学研究分析报告

1 技术背景与简介

1.1 技术原理

1.2 技术流程

1.3 分析流程

2 分析结果

2.1 质控分析

2.1.1 下机数据质控

2.1.2 基因组比对

2.1.3 基因表达定量分析

2.2 数据标准化

2.2.1 双细胞过滤

2.2.2 数据质控

2.3 高变基因选择和PCA降维

2.3.1 高变特征筛选

2.3.2 主成分分析

2.4 细胞聚类分群

2.4.1 细胞聚类

2.4.2 细胞分群

2.4.3 细胞周期分析

2.5 细胞类型注释以及marker基因的鉴定

2.5.1 细胞类型注释

2.5.2 注释后marker基因的鉴定

2.5.3 marker基因富集分析

2.6 轨迹分析

2.6.1 拟时序分析

2.6.2 基因拟时序

2.7 PAGA分析

2.8 RNA velocity分析

2.9 细胞通讯分析

2.9.1 细胞间受体配体分析

2.9.2 多个配体-受体介导的细胞互作关系分析

2.10 SCENIC单细胞转录因子分析

3 分析软件说明

4 参考文献

5 联系我们

联系方式

1 技术背景与简介

1.1 技术原理

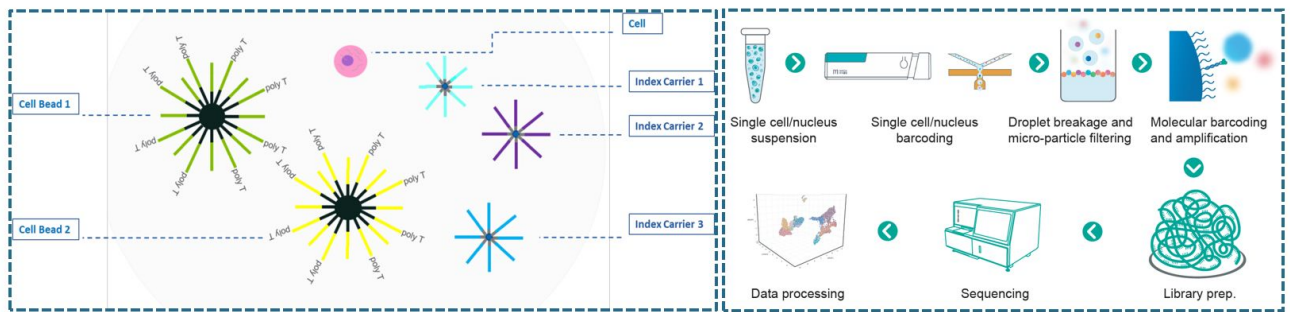
单细胞转录组 (Single-Cell Transcriptomics) 是研究单个细胞中基因表达的技术, 旨在揭示不同细胞在转录水平上的异质性。与传统的转录组测序技术 (通常在多个细胞的混合样本中进行) 不同, 单细胞转录组通过对单个细胞进行高通量RNA测序, 可以提供细胞层面的基因表达信息。单细胞转录组测序 (single cell RNA sequencing, scRNA-seq) 对组织中的细胞进行独立分析, 在细胞图谱构建、细胞亚群细化及稀有细胞类型鉴定、疾病标志性致病基因及干细胞发育分化等过程中发挥重要作用。单细胞转录组学在人类健康方面主要涉及发育研究、免疫系统研究和肿瘤的研究。华大单细胞液滴生成系统, 是基于负压的液滴微流控系统, 通过引入自主专利的液滴标签技术 (Disc-seq: Droplet-indexed high-throughput single-cell sequencing), 将带有标签的捕获磁珠与单个细胞或者细胞核包裹在液滴中, 采用Droplet Index的技术实现磁珠的超泊松分布, 在液滴中完成细胞裂解和捕获 mRNA 或 DNA 分子及用于识别来自同一液滴磁珠的标签序列, 对 cDNA 和 Droplet Index 进行文库构建和测序, 即可一次性获得大量细胞的基因表达或染色质开放区基因信息。此技术与常规液滴单细胞测序技术如 Drop-seq 平台 1%-3% 的细胞回收率相比, 回收率提高至30%-60%, 显著提高了细胞的利用率, 降低了单次细胞投入量。



图1-1 单细胞测序仪器

1.2 技术流程

取到组织后, 进行组织解离获得质量合格的单细胞悬液, 再使用DNBelab C系列高通量单细胞RNA文库制备试剂盒套装V3.0, 通过DNBelab C-TaiM 4单细胞液滴生成仪, 将制备好的细胞 (核) 悬液与逆转录反应的相关试剂混匀作为细胞相, 将含有标签信息的大磁珠Cell Beads和小磁珠Index Carrier作为磁珠相, 通过液滴微流控技术, 将细胞相和磁珠相用油滴包裹, 进而生成液滴。然后, 在每个液滴中, 细胞裂解释放mRNA, 大磁珠Cell Beads释放Cell Barcode, 小磁珠Index Carrier释放Droplet Barcode, 完成液滴内逆转录反应。随后破乳回收cDNA中间产物和 Oligo产物, 分别进行cDNA文库和Oligo文库的构建, 最后上机测序, 进行数据分析。



大、小双微珠捕获更精确

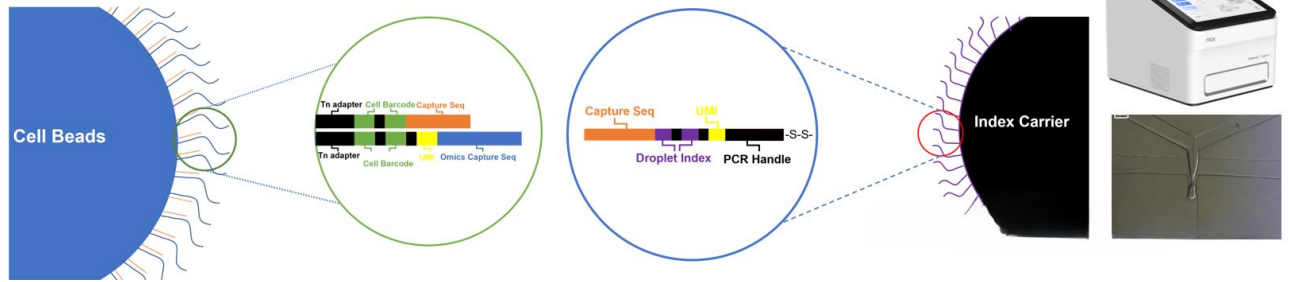


图1-2 单细胞技术简介

1.3 分析流程

dnbc4tools¹是一款为生物信息学研究人员设计的数据处理工具，主要用于处理大规模的 Next Generation Sequencing (NGS) 数据。这款工具集成了质量控制、数据过滤、比对、转录组分析等多种功能，能够为研究者提供一站式的解决方案。具体分析过程如下图所示：

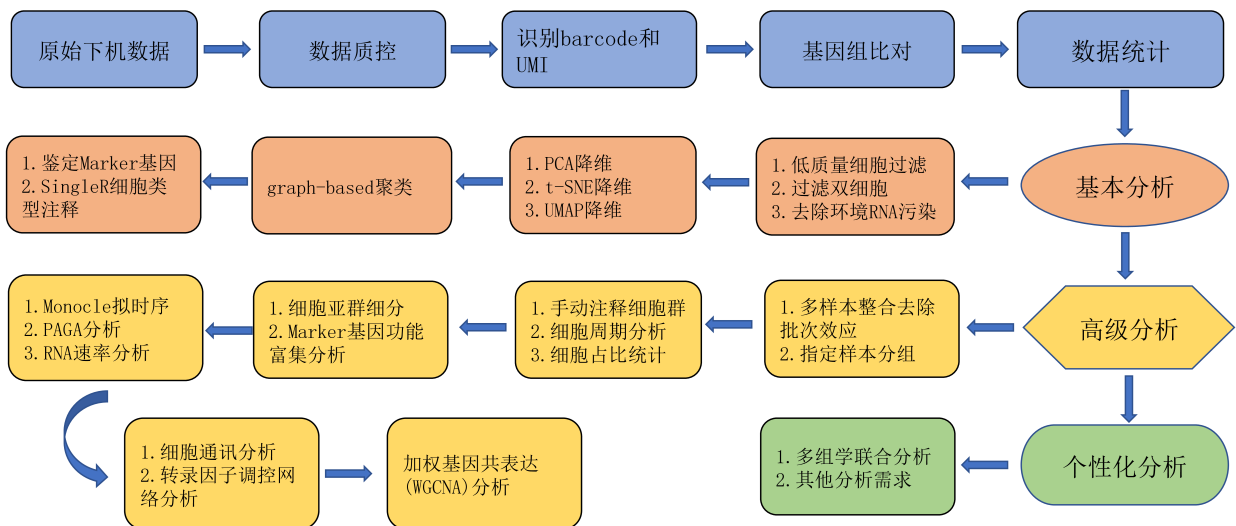


图1-3 分析流程

2 分析结果

2.1 质控分析

2.1.1 下机数据质控

测序数据下机后，会对read1和read2进行数据质控。read1包含20bp cell barcode序列，10 bpUMI序列以及cDNA序列，read2仅包含cDNA序列。接下来使用dnbc4tools工具对read1和read2进行质控过滤。质控统计结果见表1。

表1 下机数据质控统计结果

SampleName	Species	cDNA Number of reads	Fraction Reads in cell	Sequencing saturation	cDNA Q30 bases in reads
G1-SC-1	Mus_musculus	627,507,641	56.45%	37.02%	91.70%

SampleName: 样本名称。

Species: 物种名称。

cDNA Number of reads: 通过测序程序获得的原始cDNA片段的数量。

Fraction Reads in cell: 在所有测序数据中，能够被分配到单个细胞的reads所占的比例。

Sequencing saturation: 测序饱和度。

cDNA Q30 bases in reads: cDNA序列片段中质量分数大于等于30的碱基所占的百分比。

2.1.2 基因组比对

使用dnbc4tools软件将clean reads比对到参考基因组上，并根据基因注释文件，分别统计比对到基因组上以及基因组上外显子区、内含子区和基因间区等区域的reads比例。比对结果见表2:

表2 比对结果统计表

SampleName	Species	Reads mapped to genome	Reads mapped to exonic regions	Reads mapped to intronic regions	Reads mapped to intergenic regions
G1-SC-1	Mus_musculus	94.64%	86.5%	5.6%	7.9%

Reads mapped to genome: 在单细胞测序数据分析中，比对到参考基因组上的Reads在总Reads中所占的比例。

Reads mapped to exonic regions: 在单细胞测序数据分析中，比对到基因组外显子区域的Reads在总Reads中所占的比例。

Reads mapped to intronic regions: 在单细胞测序数据分析中，比对到基因组内含子区域的Reads在总Reads中所占的比例。

Reads mapped to intergenic regions: 在单细胞测序数据分析中，比对到基因组中非编码区域（即基因间区域）的Reads在总Reads中所占的比例。

2.1.3 基因表达定量分析

使用dnbc4tools软件，将比对到参考基因组唯一位置的reads（Uniquely Mapping Reads）与基因的对应关系以及基因表达量（UMI数量）进行统计，得到每个细胞基因捕获情况，具体结果见表3：

表3 基因捕获情况统计

SampleName	Species	Estimated number of cell	Median UMI counts per cell	Median genes per cell	Total genes detected
G1-SC-1	Mus_musculus	13,314	5,563	1,631	24,651

Estimated number of cell: 捕获的细胞数。

Median UMI counts per cell: 每个细胞中UMI（Unique Molecular Identifiers）的数量的中位数。

Median genes per cell: 每个细胞中检测到的基因数的中位数。

Total genes detected: 在单细胞测序分析中，至少检测到一条UMI（Unique Molecular Identifiers）的基因总数。

官方报告详见src/summary/BGI_result/1.raw_data/2.report

2.2 数据标准化

2.2.1 双细胞过滤

双细胞是指由于实验原因，在单细胞微流控等流程中某个液滴中含有2个及以上的细胞，而在同一个液滴中的细胞在后续分析中带有相同的Cell Barcode，从而被视为一个细胞的伪细胞。这类伪细胞的主要特点是检测到的UMI数和基因数往往比正常细胞要多一倍及以上，另外可能会带有不同细胞类型的经典marker基因，会给细胞类型鉴定带来阻碍。使用R包DoubletFinder²对双细胞进行鉴定并过滤。

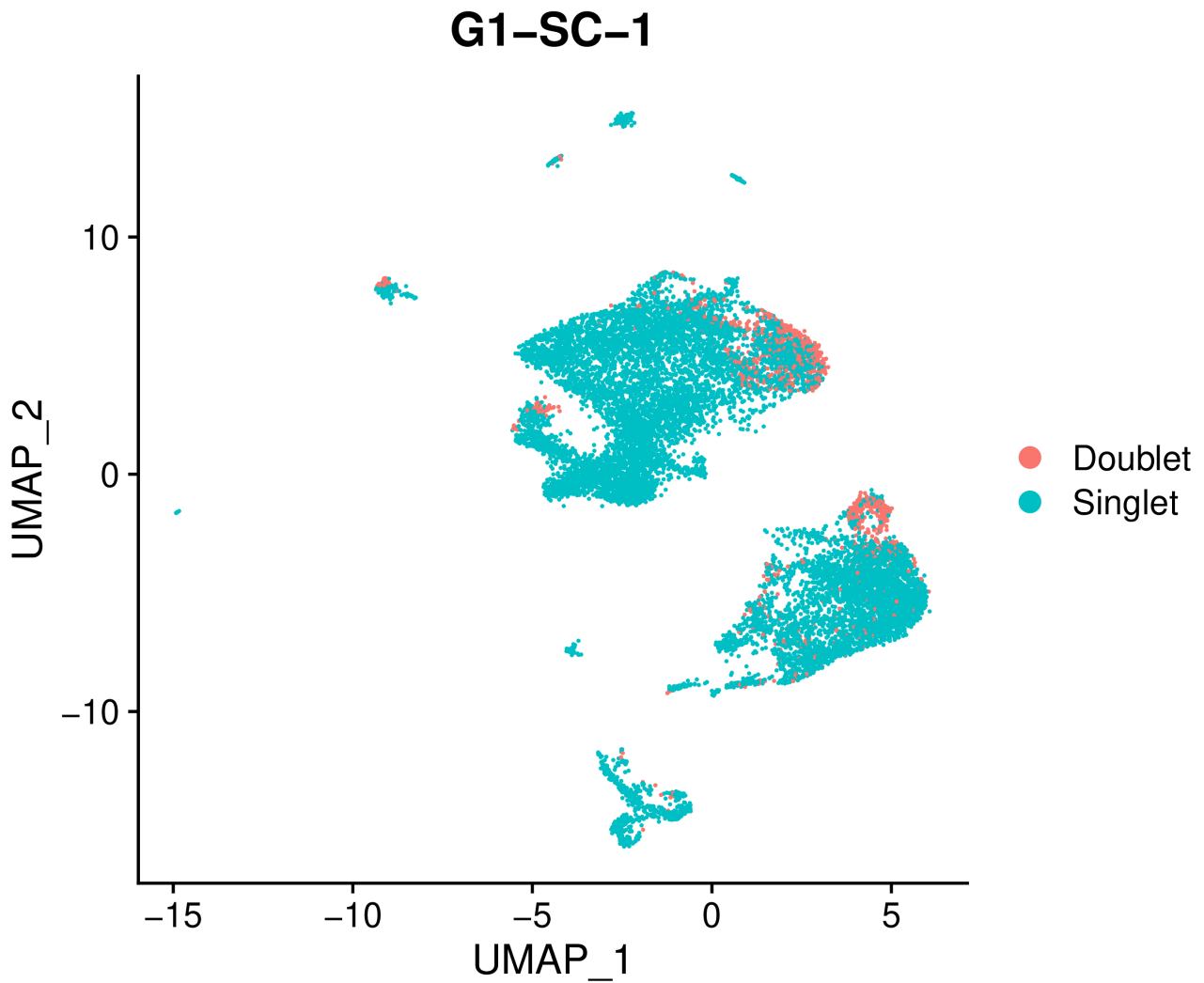


图2.2.1-1 双细胞鉴定UMAP图

Doublet: 双细胞, Singlet: 单个细胞

进行双细胞过滤之后, 使用R包SoupX³去除基因组污染, 然后进行基础质控。

2.2.2 数据质控

使用R包Seurat⁴对单个样品的不同文库分别进行质控, 满足质控标准后合并用于后续分析。通过可视化细胞的基因数量分布图和UMI分布图, 可以评估每个样品的细胞活性及基因表达情况。线粒体RNA比例, RNA数量和表达量在一定程度上反映出细胞活性和质量。一般来说, 线粒体RNA比例越低、RNA数量越多、RNA表达量越高的细胞, 其活性和质量越好。

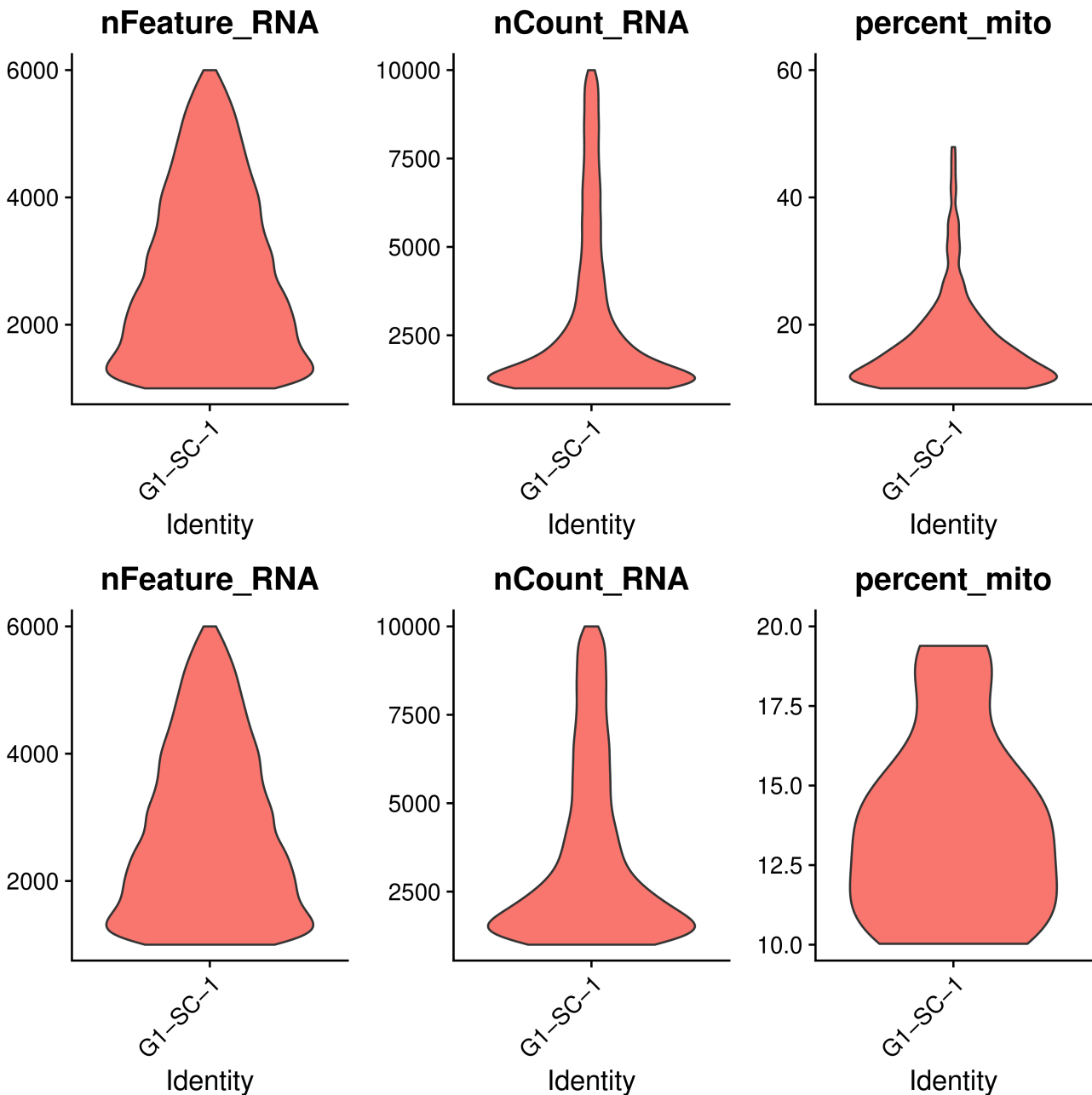


图2.2.2-1 质控结果小提琴图

左图：mRNA数；中图：mRNA表达量；右图：线粒体RNA比例。第一排为质控之前的数据情况，第二排为质控之后的数据情况。根据定量结果进行数据过滤，过滤掉mRNA表达量过低或过高，以及线粒体RNA比例过高的细胞（过滤标准：保留基因数大于500以及线粒体基因占比小于20%的细胞）。

2.3 高变基因选择和PCA降维

2.3.1 高变特征筛选

基于每个样品中所有基因表达量平均值和分散度（均值和方差），使用R包-Seurat筛选出每个样品中呈现高变异度的基因，用于下游的PCA分析。默认挑选变异程度最高的2000个基因（由于之前进行细胞与基因的过滤，会导致少部分基因被过滤掉，实际基因数可能会略低于2000）。

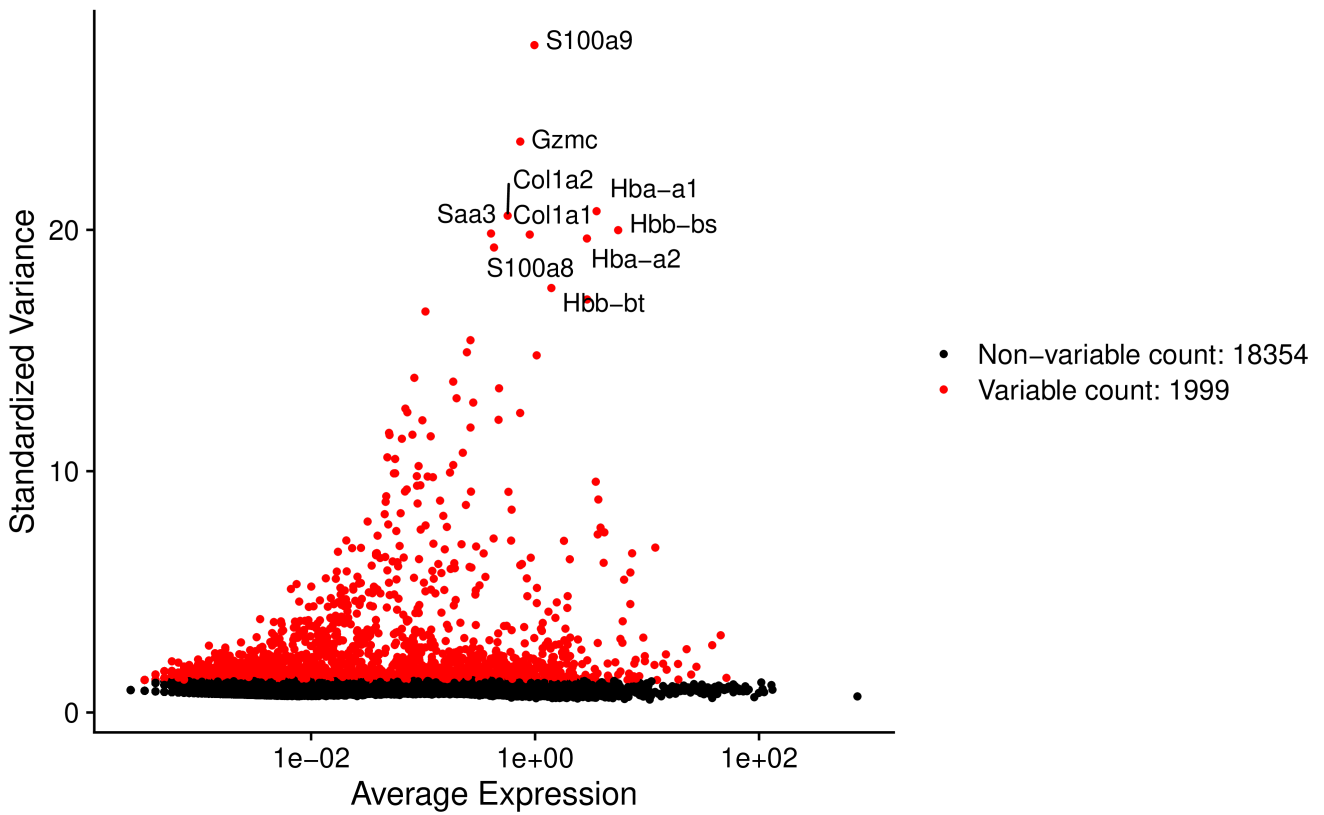


图2.3.1-1 高变基因筛选图

X轴：平均表达量，Y轴：标准化方差，标注了表达量变异程度top 10基因。

2.3.2 主成分分析

维度热图展示了主成分（Principal Components）变异最大的基因及其表达量在细胞数据中的异质性，有助于选择合适的维度进行下游分析。下图展示了top15 PC的降维结果，可以看到随着PC成分的增加，特征越来越不明显，有助于下游选择合适的维度（默认采用前15个维度进行下游分析）。

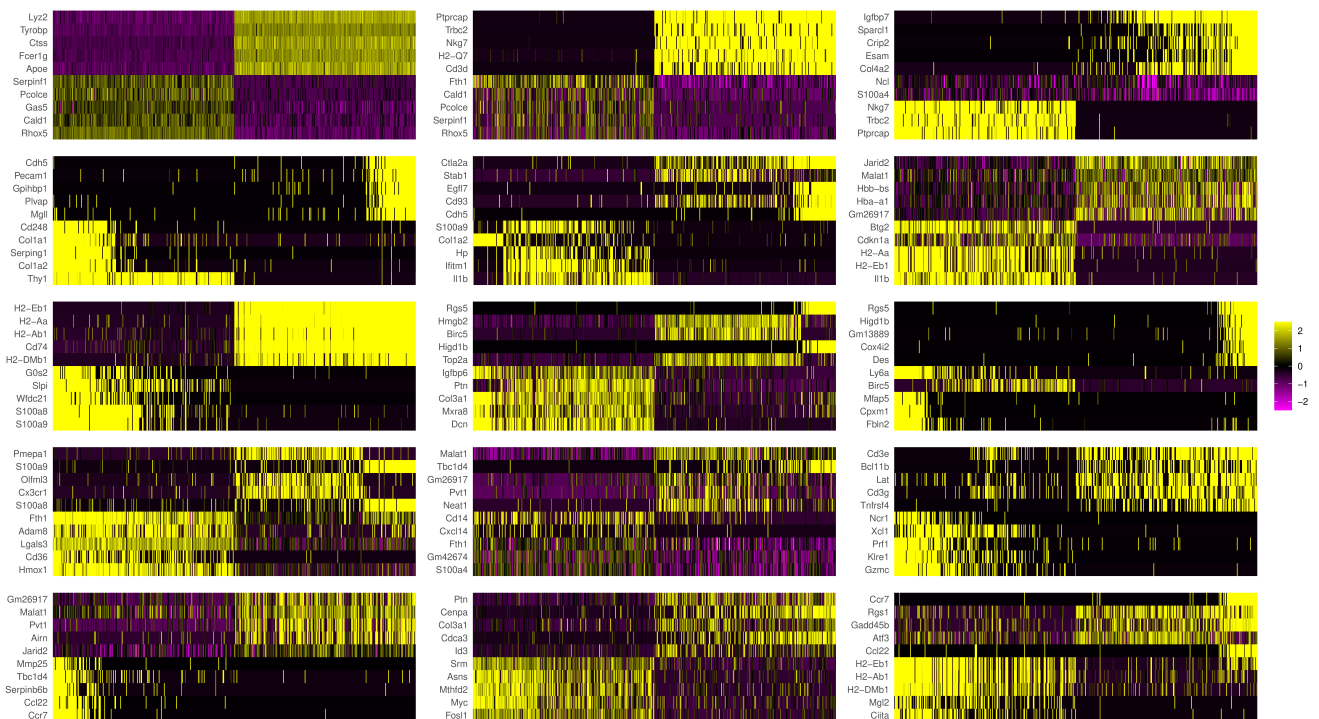


图2.3.2-1 Top15 PC的降维热图

每列表示一个细胞，每行表示一个基因；黄色表示高表达，紫色表示低表达。

2.4 细胞聚类分群

2.4.1 细胞聚类

单个样品聚类分群中每个点代表一个细胞，空间距离近的点表示这些细胞的基因表达模式比较接近；颜色根据 graphbased 聚类算法得到的分类决定，算法认为同一个 cluster 内的细胞，基因表达模式最接近。如果有多个样本，使用 Harmony 算法进行整合聚类，从而去除批次效应，进而进行下游分析。

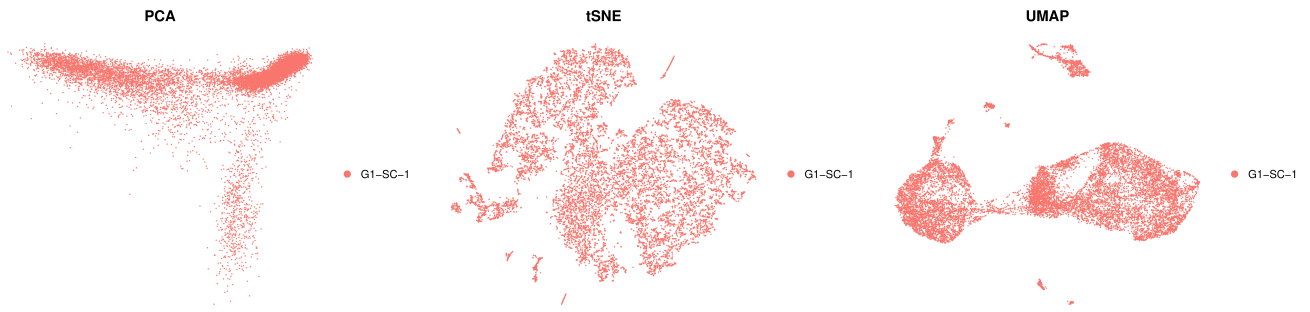


图2.4.1-1 细胞降维聚类

从左到右依次是PCA、tSNE、UMAP图。

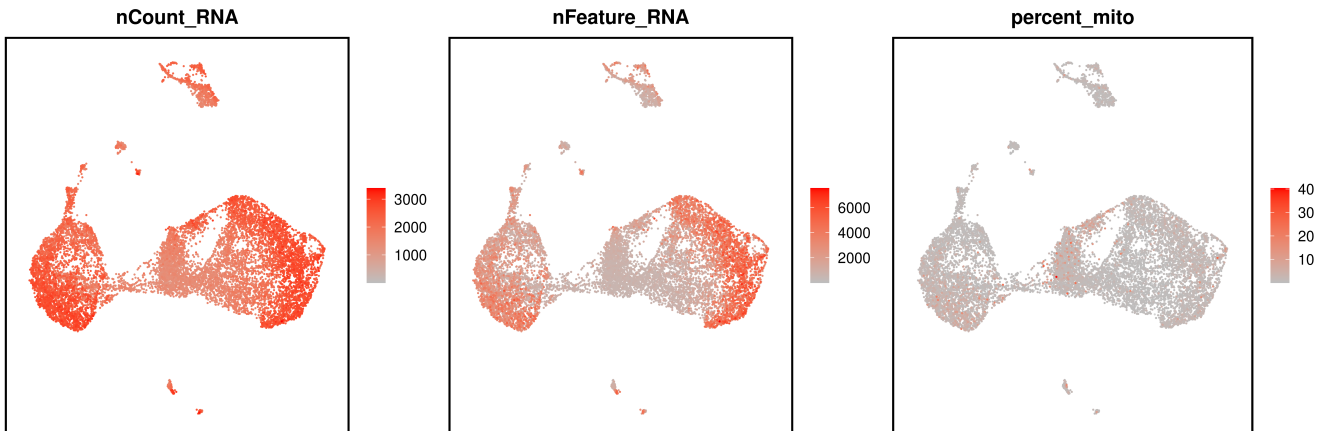


图2.4.1-2 质控结果UMAP图

左图：UMAP图展示了单细胞中总RNA计数的分布。颜色从浅灰到深红表示RNA计数从低到高的变化，其中深红色代表计数最高的细胞群体。中图：UMAP图展示了单细胞中检测到的基因数量的分布。颜色与左图相同，显示了基因数量的分布情况。右图：UMAP图展示了单细胞中线粒体基因表达比例的分布。颜色从浅灰到深红表示线粒体基因表达比例从低到高的变化。

2.4.2 细胞分群

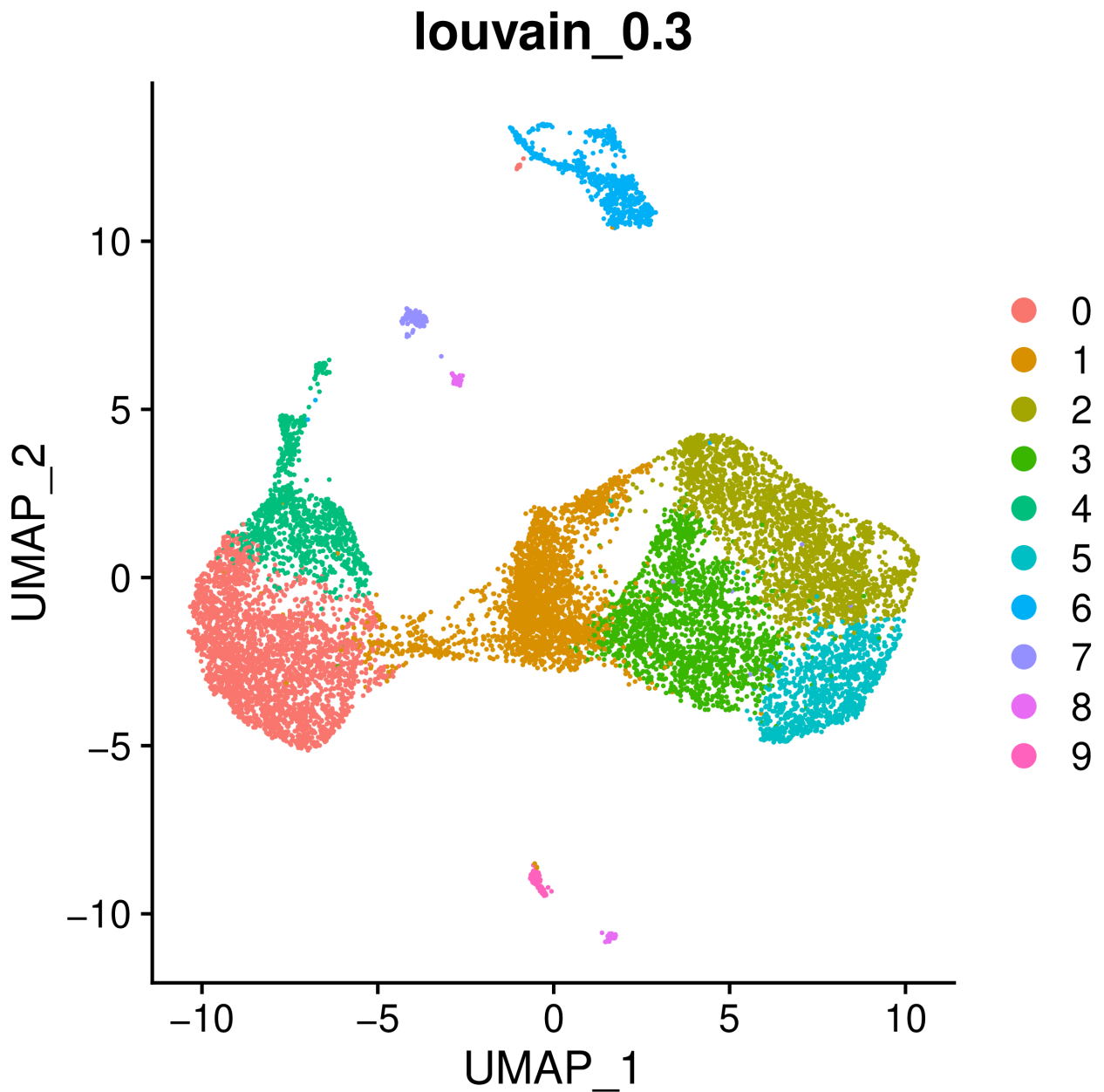


图2.4.2-1 细胞分群

选择分辨率0.3对细胞进行分群，X轴表示的是UMAP第一成分，Y轴表示的是UMAP第二成分。数字代表cluster编号。

通过R包-Seurat分别计算每一类细胞与其他类群的差异表达基因，筛选校正后p-value<0.05且|log2FC| (log2 fold change: 用于评估平均表达量差异倍数) >0.25的差异基因用于后续结果可视化。

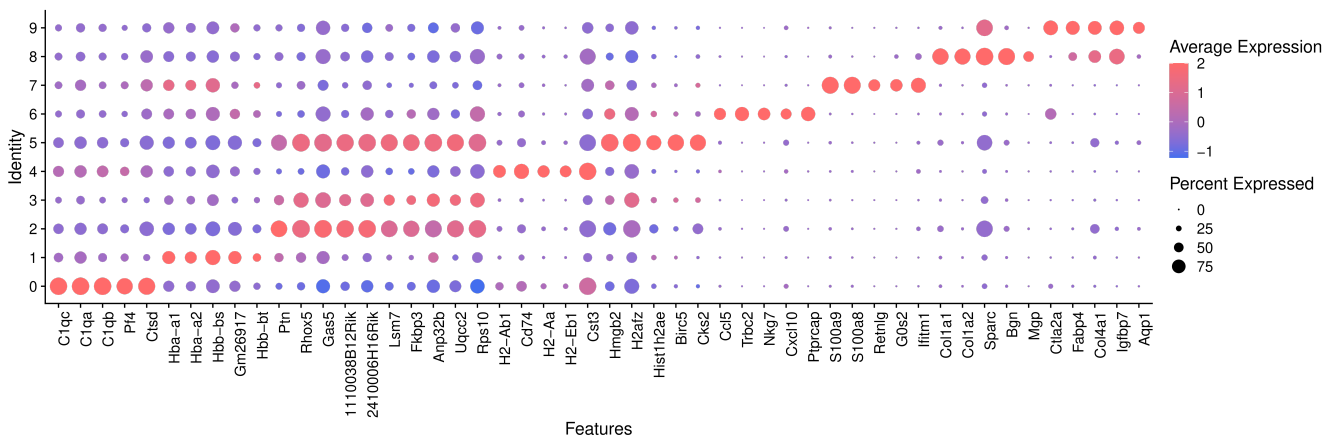


图2.4.2-2 cluster特异marker基因表达量气泡图

X轴为不同的基因，Y轴为不同cluster。

2.4.3 细胞周期分析

细胞周期(cell cycle)是指细胞从一次分裂完成开始到下一次分裂完成所经历的全过程，分为间期与分裂期两个阶段。在细胞周期过程中，细胞的遗传物质复制并均等地分配给两个子细胞。R包-Seurat中的CellCycleScoring函数可以根据每个细胞的S期(或者G2/M期)基因集表达程度进行打分，若显著高表达则判断该细胞处在S期(或者G2/M期)，并进行可视化。对于G1期，通常是在S期和G2/M期之外的细胞群体，这些细胞没有显著的DNA复制活动或细胞分裂活动。细胞周期分析结果如下图：

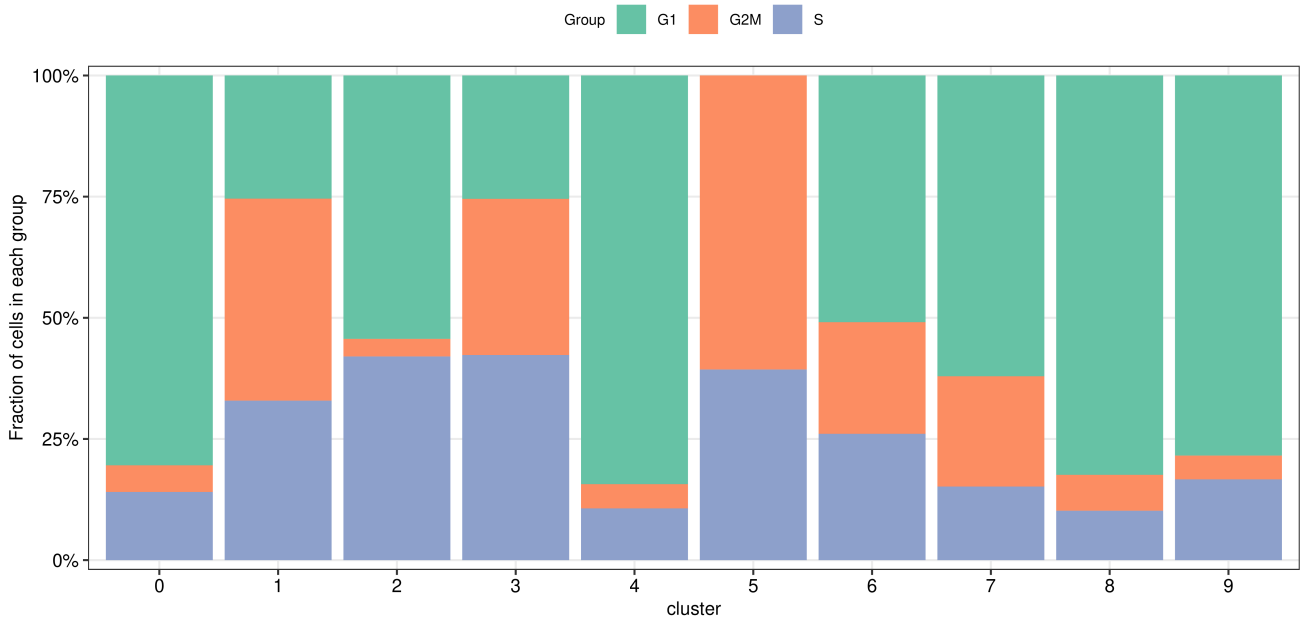


图2.4.3-1 细胞周期分析

X轴为不同的cluster，Y轴为每个细胞亚群中G1期，G2/M期以及S期的细胞占比情况。

2.5 细胞类型注释以及marker基因的鉴定

2.5.1 细胞类型注释

R包-SingleR⁵可以将细胞聚类结果中的每个cluster与参考数据库进行比较，根据该cluster与参考数据库中已知细胞类群基因表达谱的相似性对cluster进行打分，最终得到一个打分矩阵。目前SingleR可以注释的物种为人和小鼠。将注释好的细胞类型对应到cluster编号上，绘制细胞类型注释图。

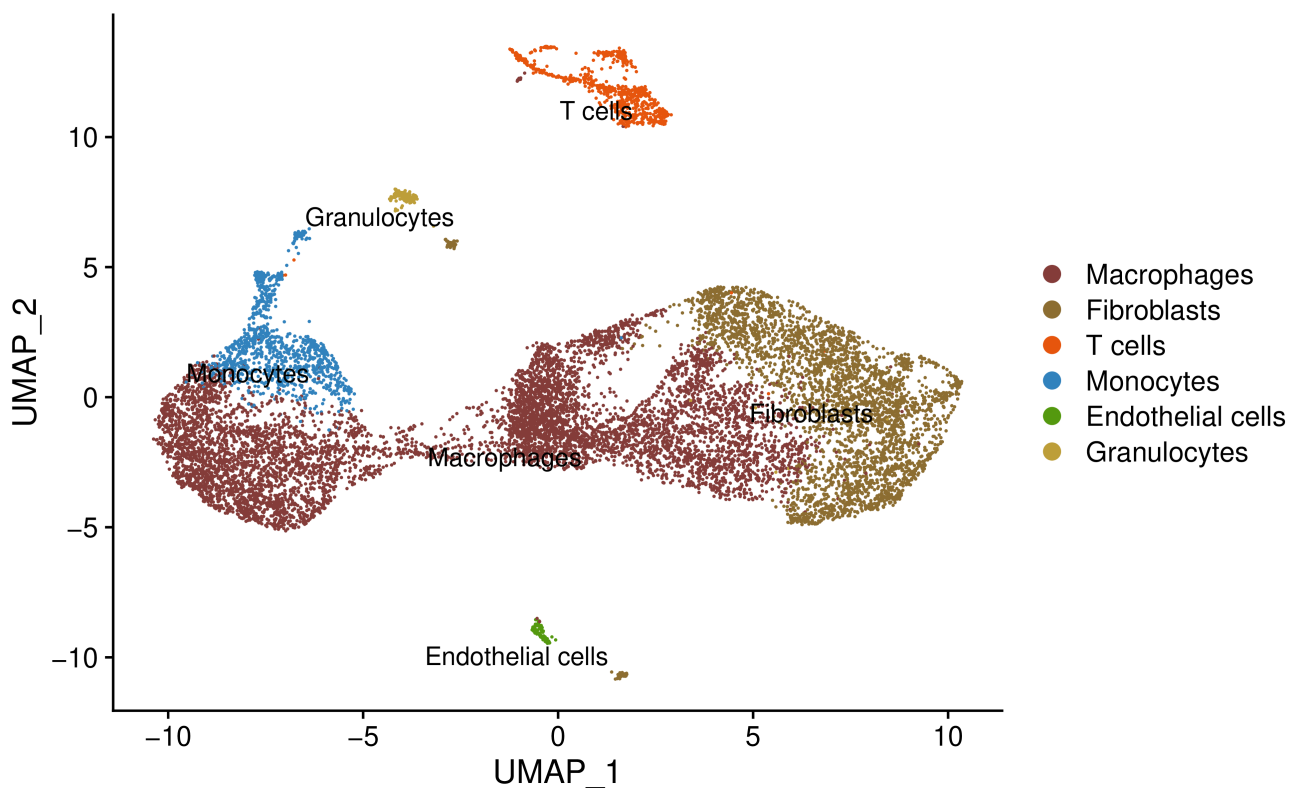


图2.5.1-1 细胞类型注释图

X轴表示的是UMAP第一成分，Y轴表示的是UMAP第二成分。图中不同颜色代表注释到的不同细胞类型。

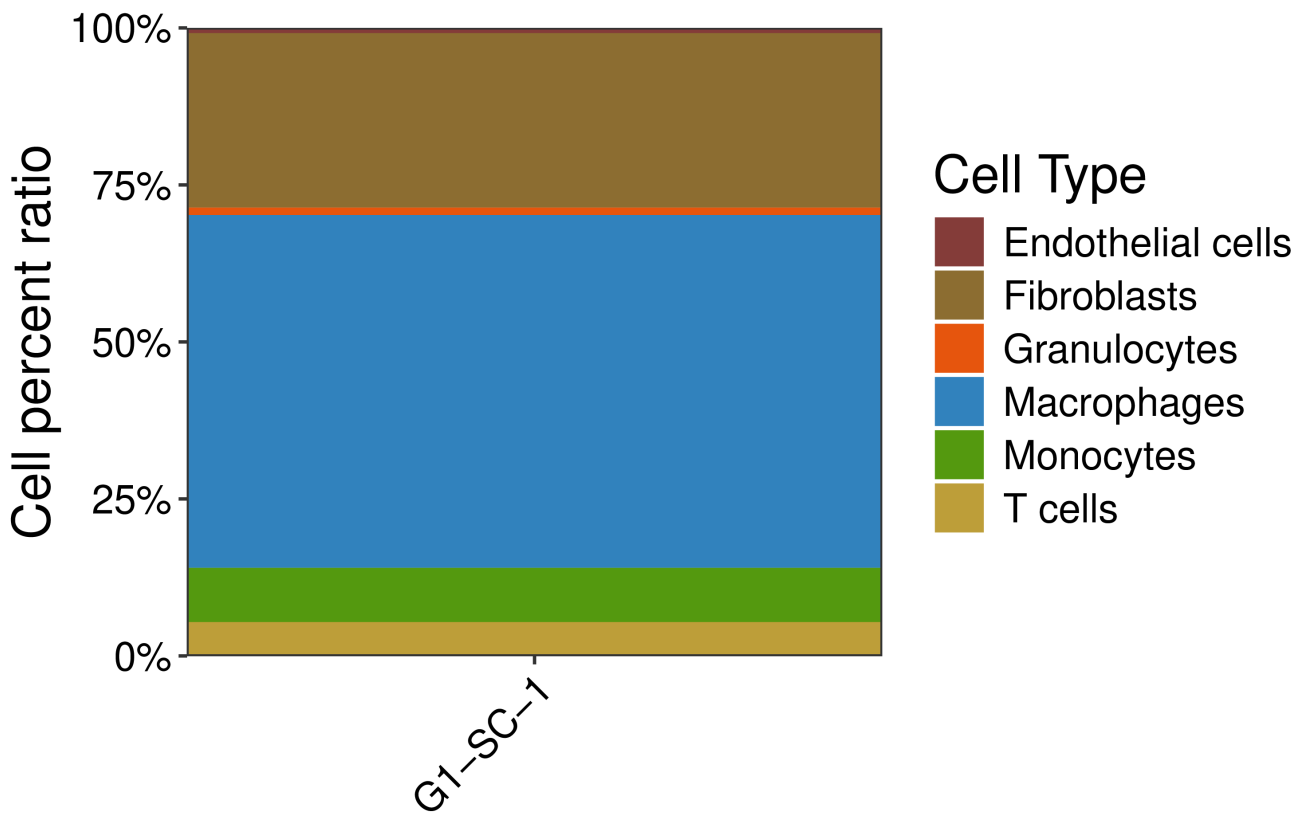


图2.5.1-2 细胞类型比例图

X轴表示的样本名称，Y轴表示的是不同细胞类型的占比。

2.5.2 注释后marker基因的鉴定

通过R包-Seurat分别计算每一类细胞与其他类群的差异表达基因，筛选校正后p-value<0.05且|log₂FC| (log₂ fold change: 用于评估平均表达量差异倍数) >0.25的差异基因用于后续结果可视化。

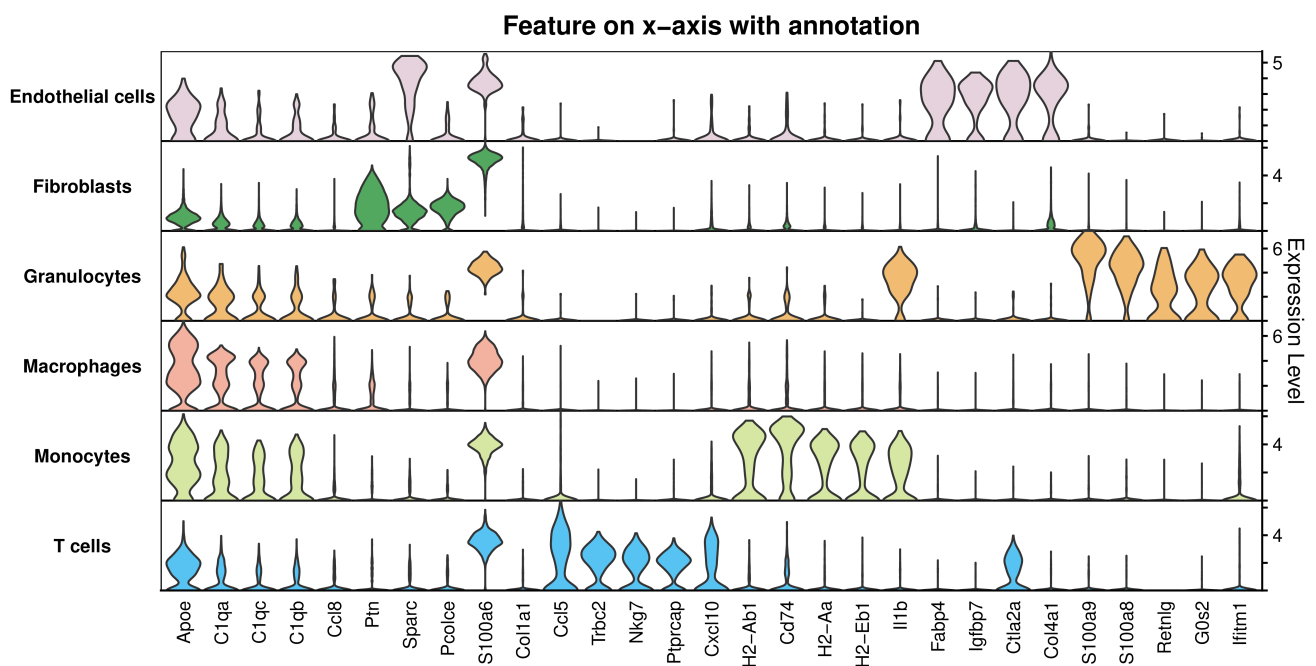


图2.5.2-1 cluster特异marker基因表达量小提琴图

X轴为不同的基因，Y轴为不同细胞类型。图中展示了每个细胞类型TOP 5基因。

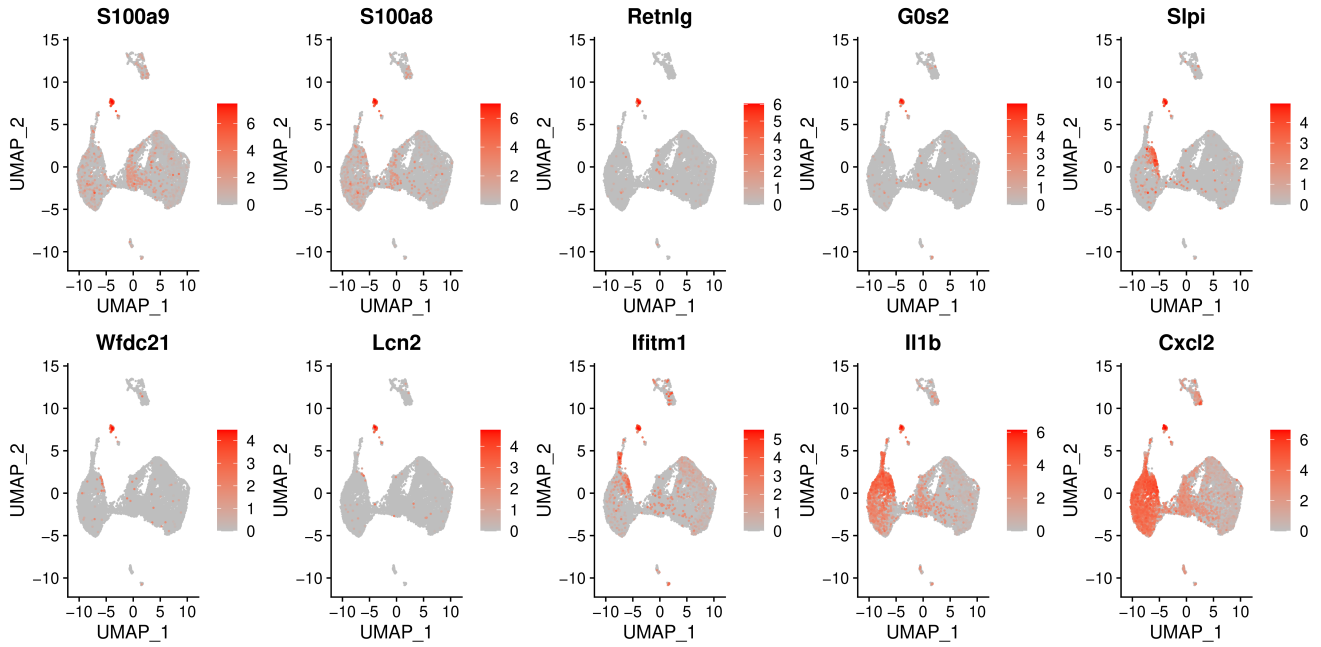


图2.5.2-2 cluster特异marker基因UMAP图

X轴表示的是UMAP第一成分，Y轴表示的是UMAP第二成分。颜色越红，表示平均表达量越高。

其余cluster特异marker基因详见src/summary/BGI_result/4.cell_annotation/4.marker_featureplot

随后通过热图展示了每个cluster中变化量大的基因，筛选矫正后p-value<0.05且|log2FC|>0.25的基因。

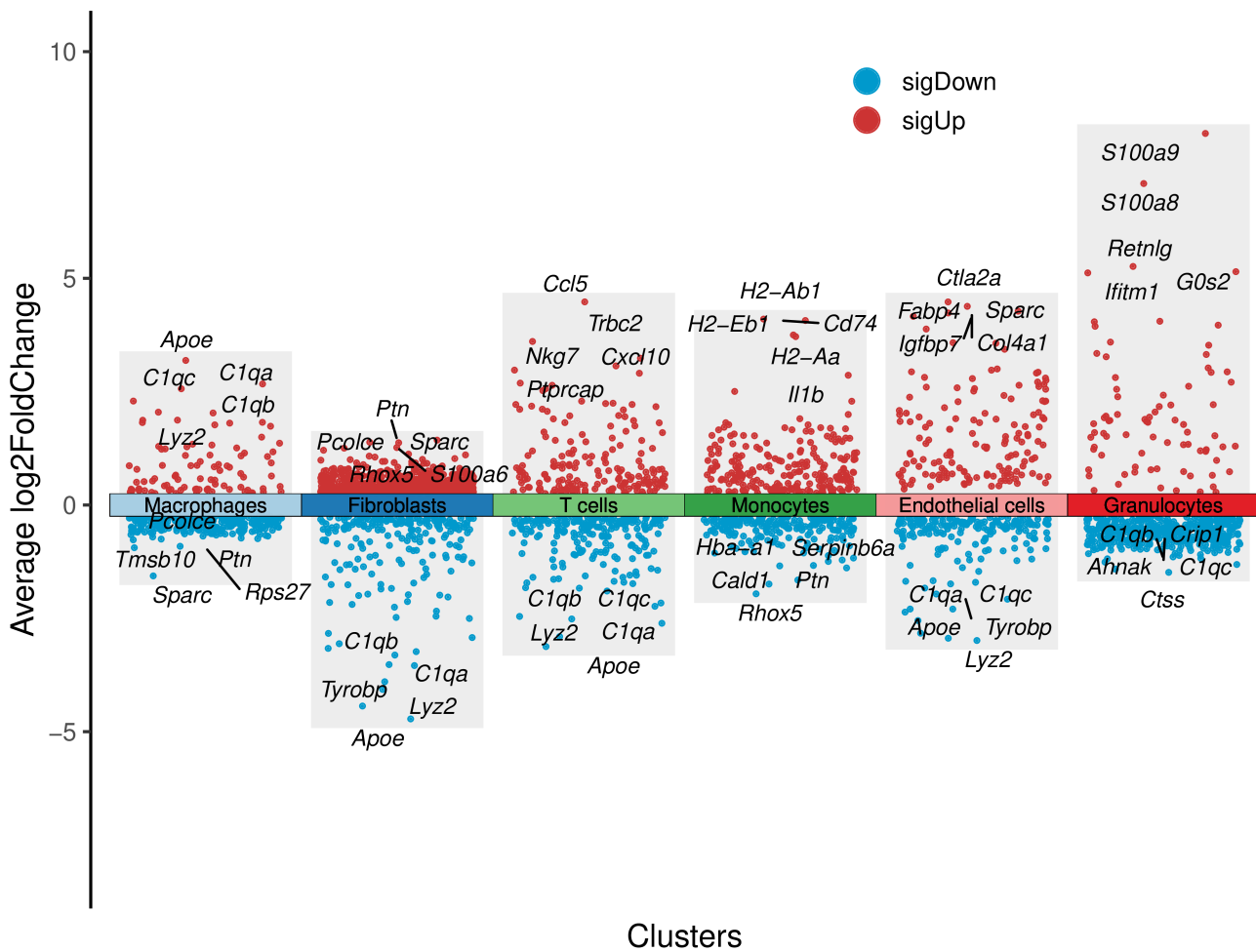


图2.5.2-3 cluster差异基因火山图

不同颜色表示不同的细胞类型，展示了经 $p\text{-value} < 0.05$ 且 $|\log_2FC| > 0.25$ 的基因，红色表示上调基因，蓝色表示下调基因。

2.5.3 marker基因富集分析

使用clusterProfiler对筛选出的marker基因进行GO富集分析与KEGG富集分析，挑选 $p\text{-value} < 0.05$ 的通路进行展示。

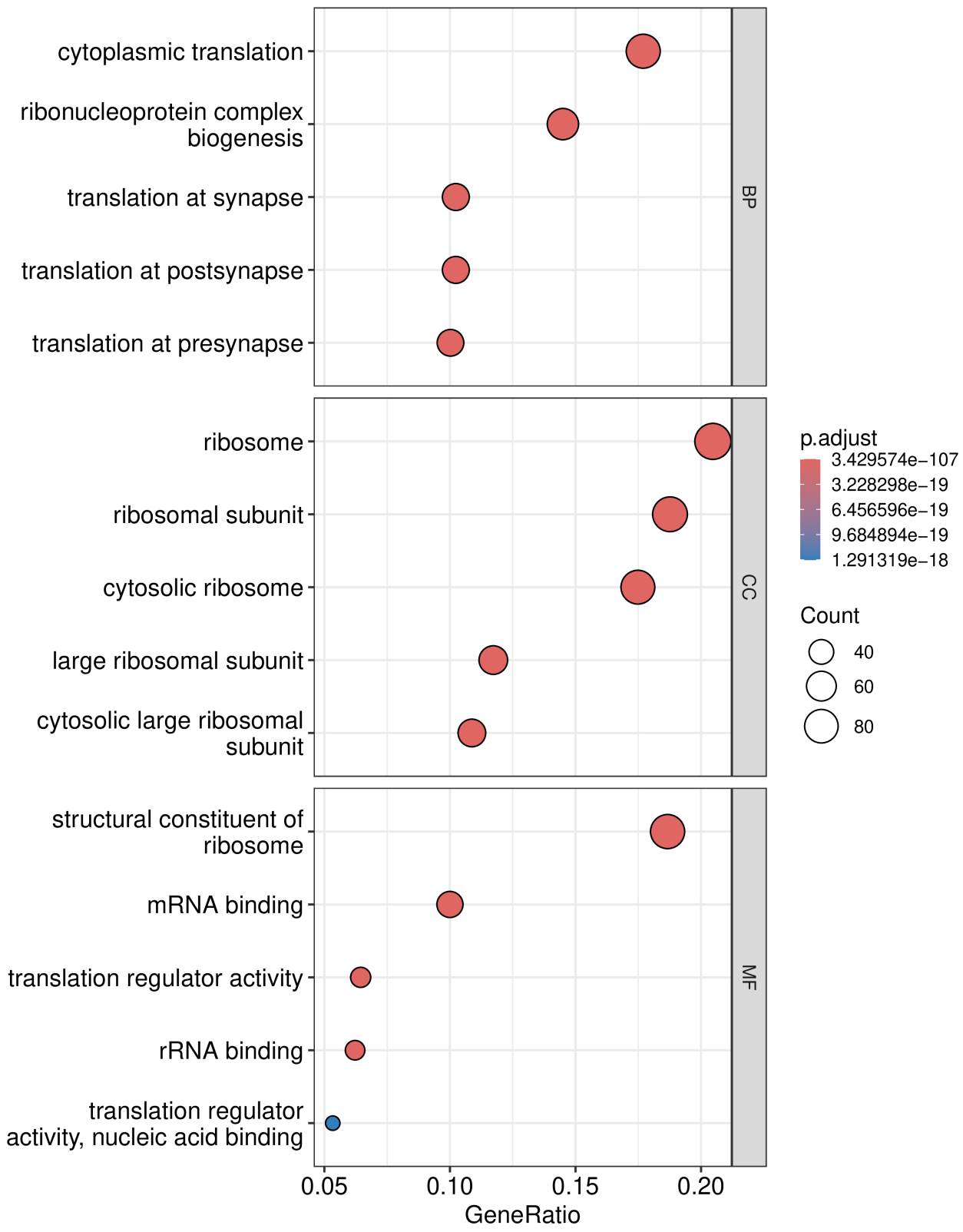


图2.5.3-1 marker基因GO功能分析气泡图

X轴代表基因占比，Y轴代表GO功能分类。BP：生物过程；CC：细胞组分；MF：分子功能。

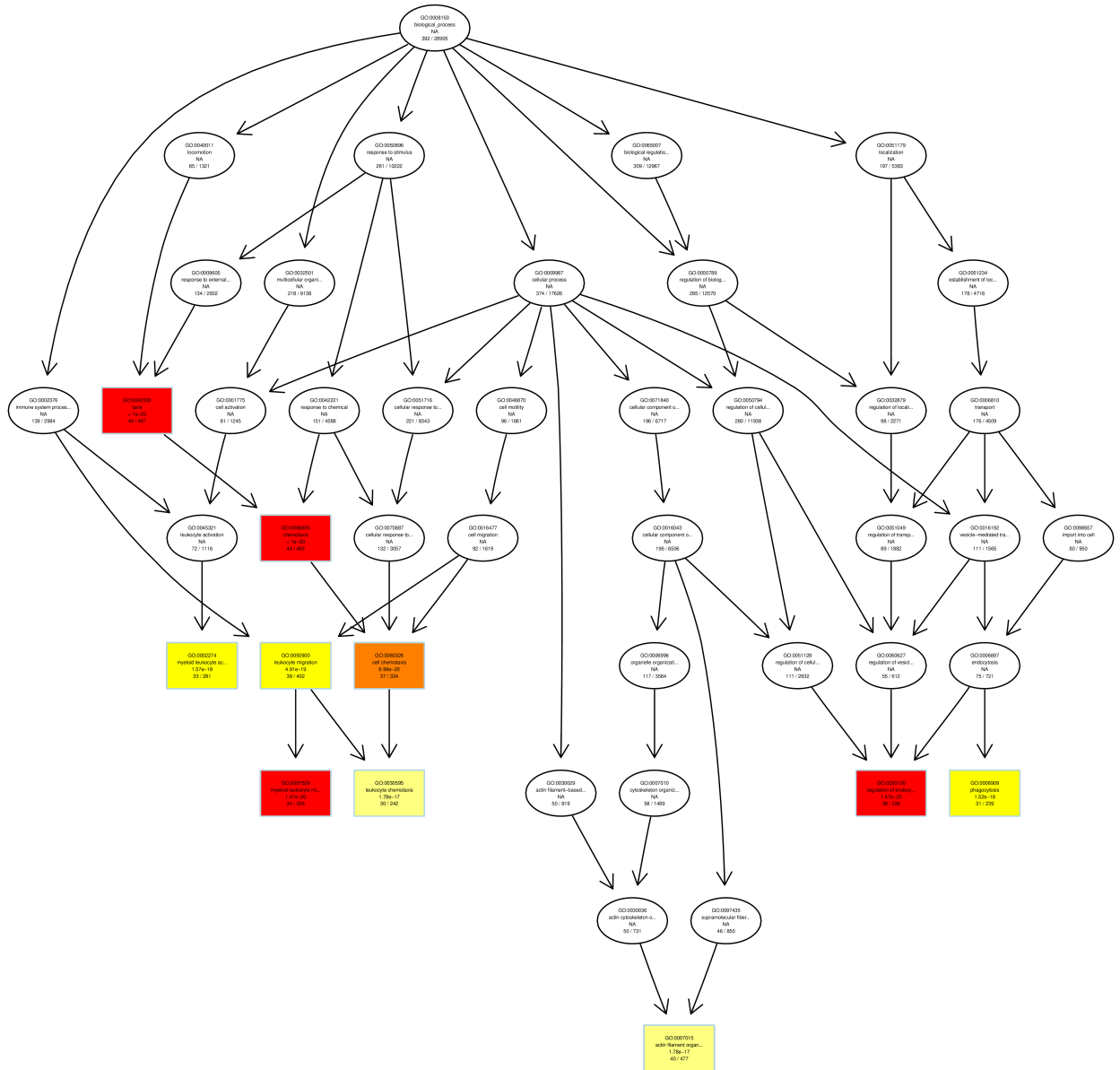


图2.5.3-2 基因GO功能分析有向无环图

有向无环图展示了通路之间的从属关系。

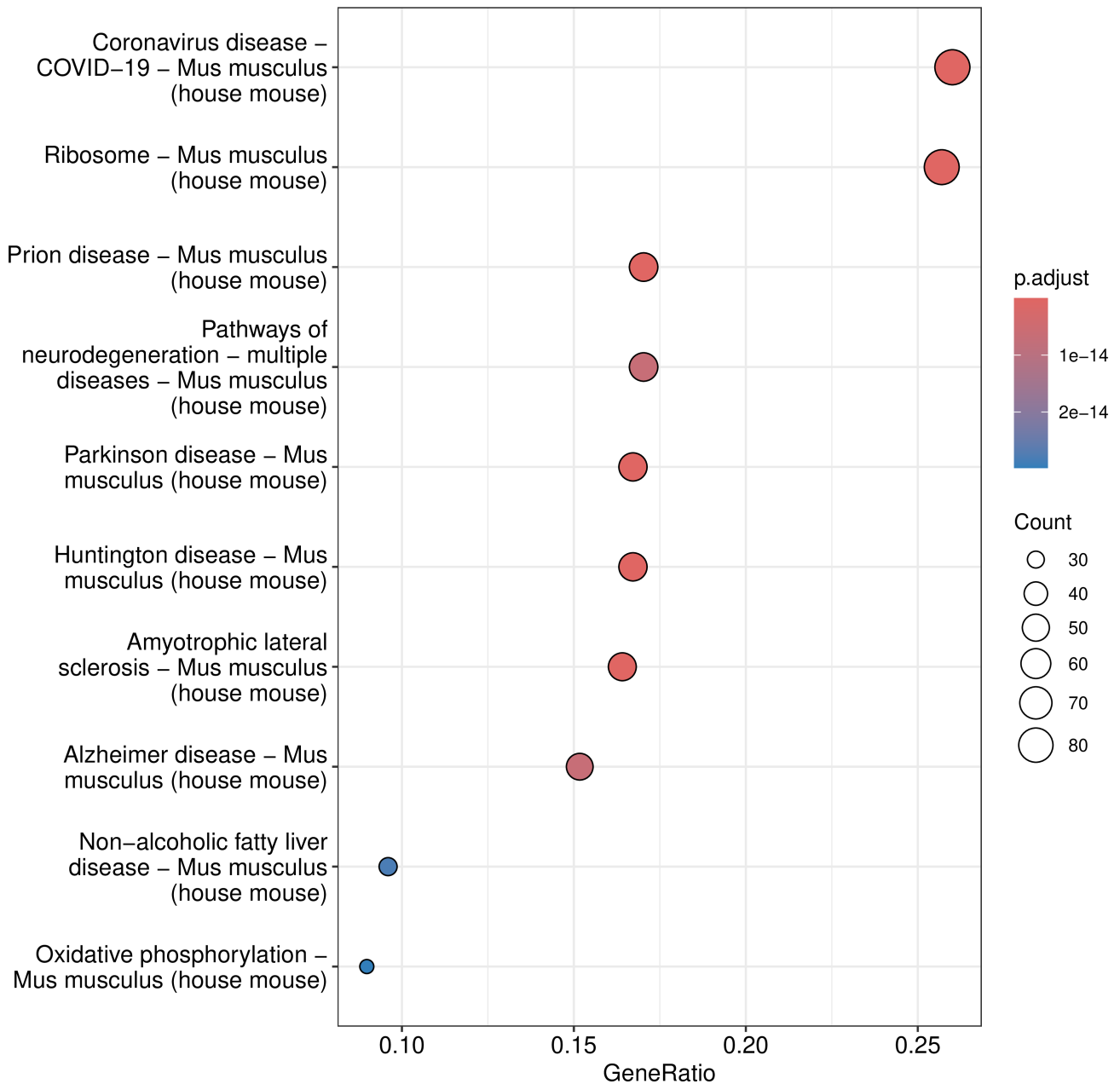


图2.5.3-3 marker基因KEGG功能分析气泡图

X轴代表基因占比，Y轴代表KEGG通路。

其余cluster富集分析详见src/summary/BGI_result/4.cell_annotation/5.marker_function

2.6 轨迹分析

拟时序分析 (pseudotime analysis) 也称为细胞轨迹分析 (trajectory analysis)，是指根据不同细胞类群之间基因差异表达的情况，获得细胞谱系的发育结果，构建细胞随着一个虚拟时间顺序的变化，以此重现细胞随时间变化而变化的过程。

2.6.1 拟时序分析

单细胞水平的研究使人们可以描述复杂生理过程和高度异质性细胞群体的转录调控。这些研究有助于发现识别特定细胞亚型的基因、标记生物过程中间状态的基因，以及在两种不同的细胞命运之间过渡态的基因。在许多单细胞研究中，单个细胞以不同步的方式执行基因表达过程。实际上，每个细胞都是正在研究的转录过程的一个瞬间。Monocle [6](#)引入了在伪时间（拟时间）内对单个细胞排序的策略，利用单个细胞的非同步进程，将它们置于与细胞分化等生物学过程相对应的轨迹上。Monocle利用先进的机器学习技术(反向图嵌入)从单细胞数据中学习显式的主图（展现细胞转录特征相似性关系的图）来对细胞进行排序，Monocle的机器学习算法可以依据上述降维图形，学习描述细胞如何从一种状态过渡到另一种状态的轨迹。

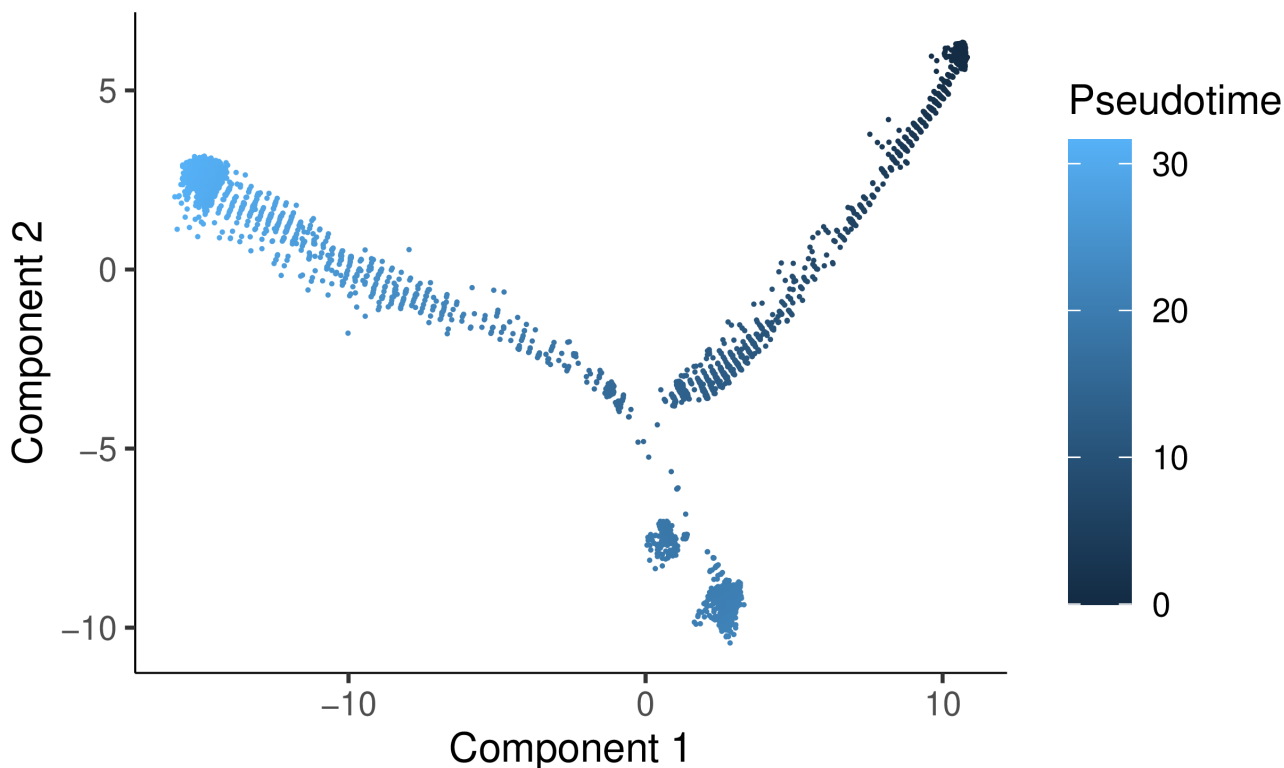


图2.6.1-1 Pseudotime发育轨迹图

X轴代表细胞的拟时序，指示细胞在某一生物过程（如发育、分化等）中的相对进程。Y轴反映的是细胞的基因表达差异，表示细胞在一些主要特征上的变异性。Pseudotime数值从小到大就是推断的发育顺序。图上的点颜色越深，时间越早；颜色越浅，时间越晚。

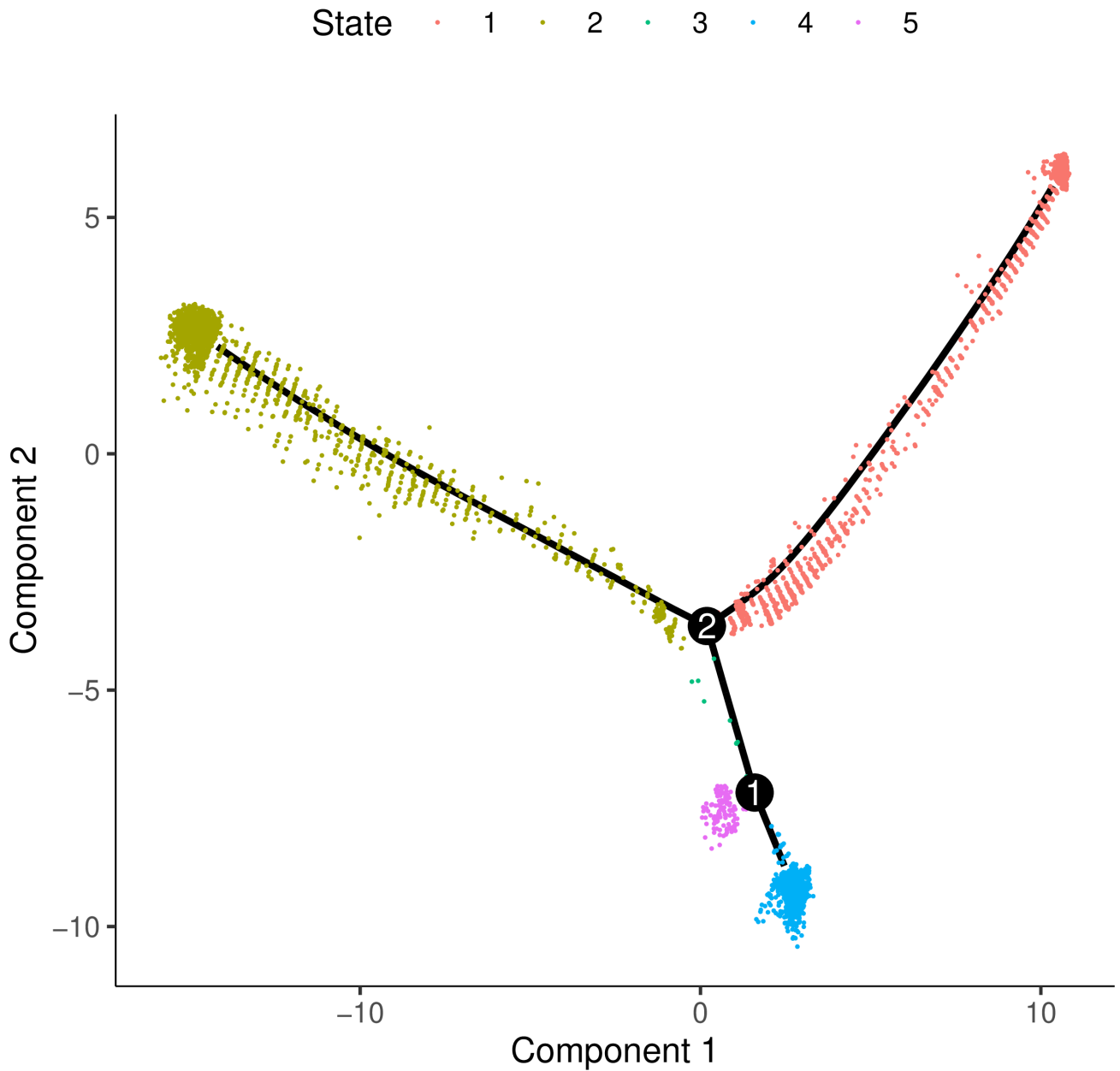


图2.6.1-2 State发育轨迹图

X轴代表细胞的拟时序，指示细胞在某一生物过程（如发育、分化等）中的相对进程。Y轴反映的是细胞的基因表达差异，表示细胞在一些主要特征上的变异性。state是发育的不同阶段，数值越小越靠前。

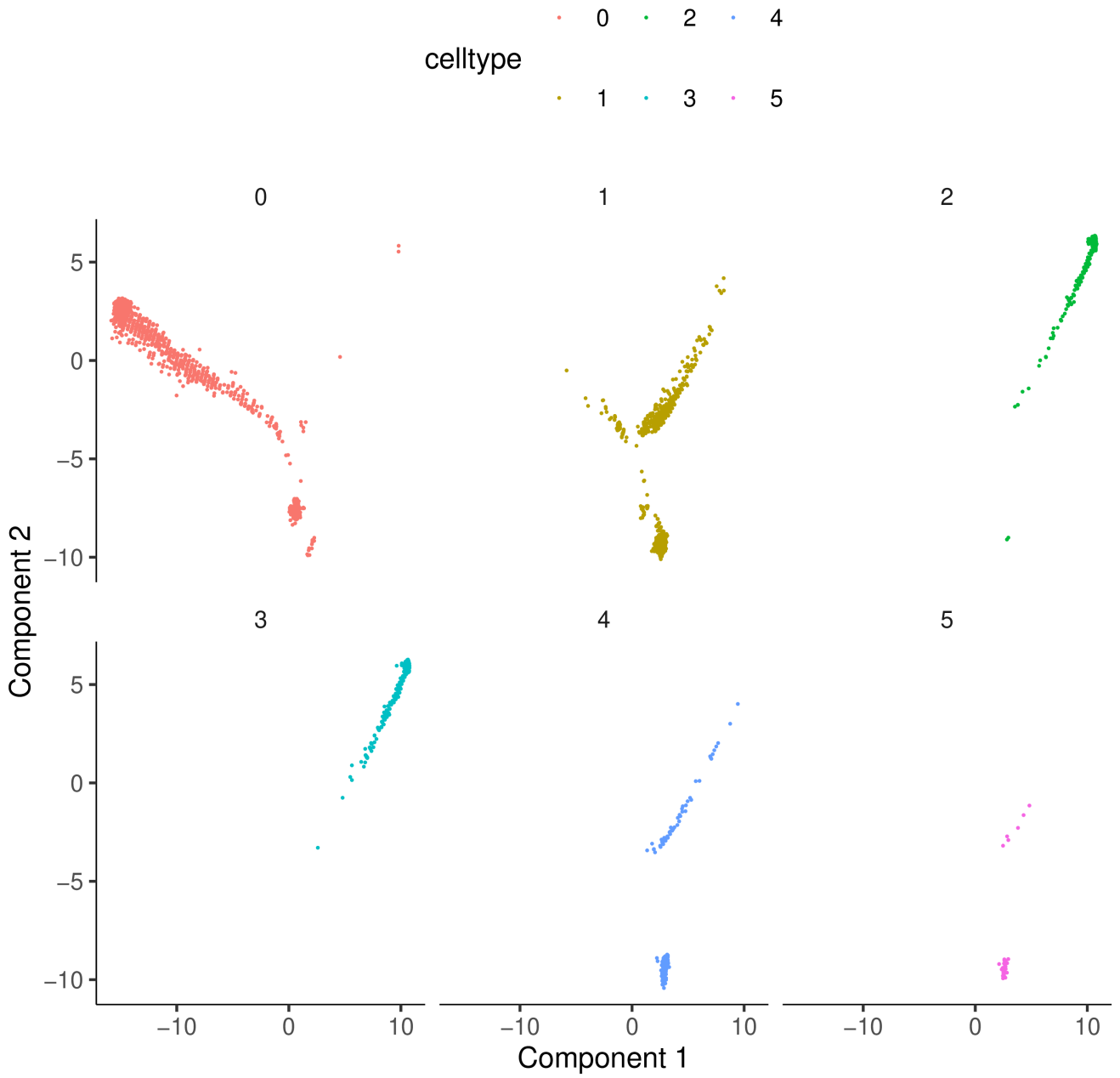


图2.6.1-3 celltype发育轨迹图

X轴代表细胞的拟时序，指示细胞在某一生物过程（如发育、分化等）中的相对进程。Y轴反映的是细胞的基因表达差异，表示细胞在一些主要特征上的变异性。celltype分面图可以看到具体的细胞类型在时间轨迹图上的分布。

下图是以样本形式的着色轨迹图

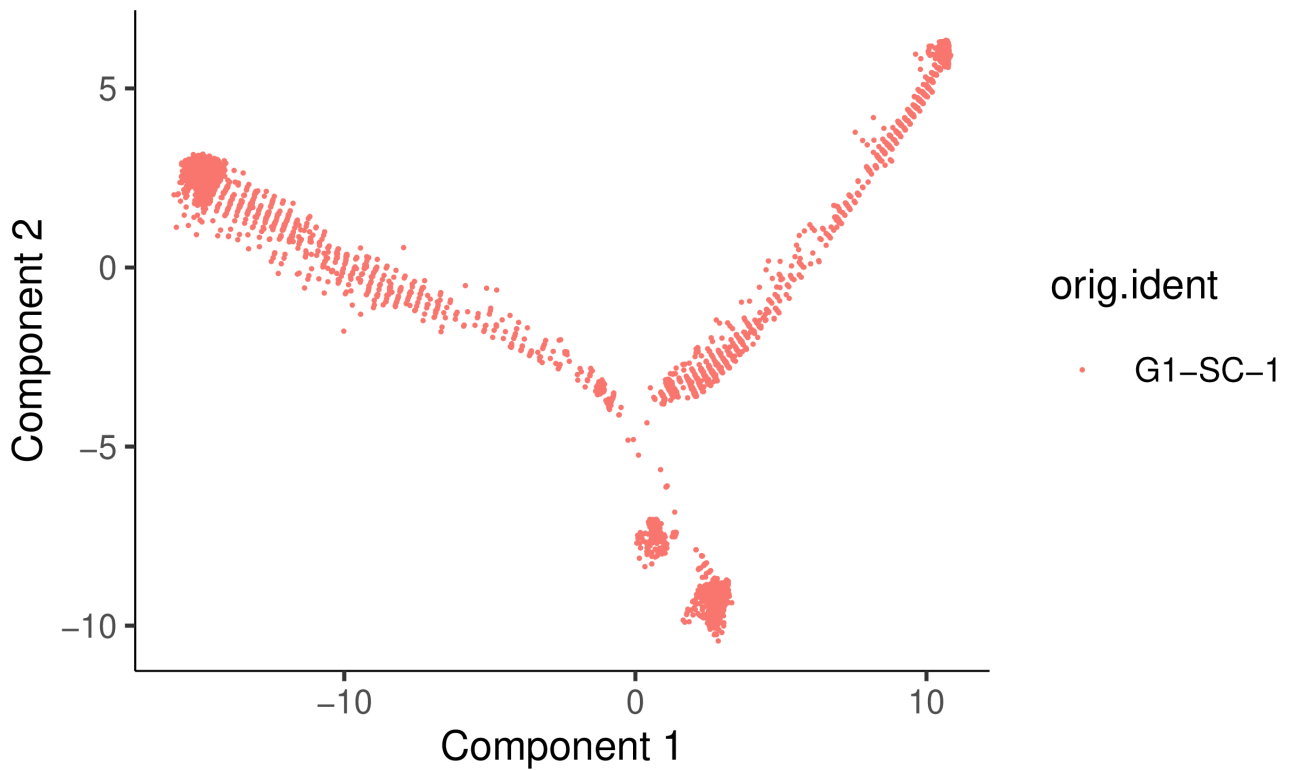


图2.6.1-4 样本形式发育轨迹图

X轴代表细胞的拟时序，指示细胞在某一生物过程（如发育、分化等）中的相对进程。Y轴反映的是细胞的基因表达差异，表示细胞在一些主要特征上的变异性。不同颜色代表不同的样本。

2.6.2 基因拟时序

采用对数尺度（Log-scale），更好地展示基因表达的动态范围，尤其是低表达和高表达的差异。不同的基因有不同的表达水平范围，从1到100或更高，显示了基因在不同阶段的活跃程度。黑色曲线表示基因表达随拟时序变化的平滑趋势线。它帮助展示基因在细胞发育过程中的整体表达趋势。

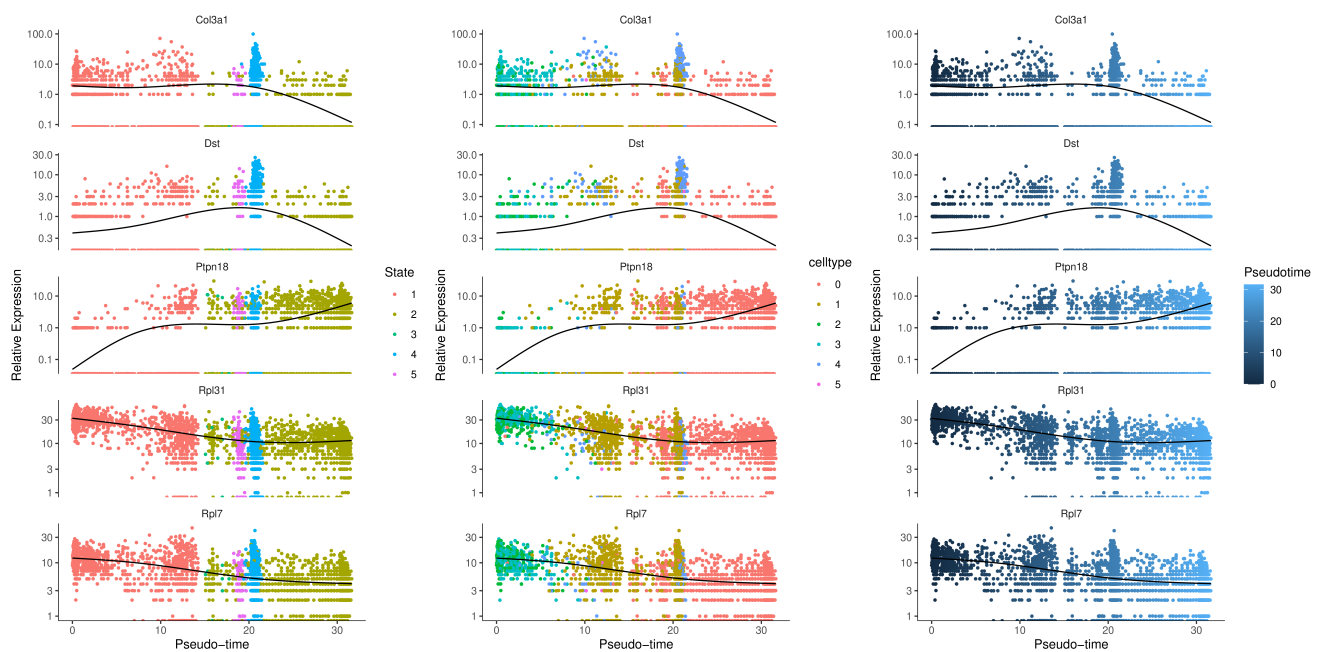


图2.6.2-1 基因拟时序点图

横轴表示拟时序 (Pseudotime)，反映了细胞在发育或状态转换中的相对顺序。数值越大，表示细胞处于发育或分化的更晚阶段。纵轴表示基因的相对表达水平。Pseudotime数值从小到大就是顺序就是推断的发育顺序。

单细胞轨迹中通常包括决策点分支，它们的出现是因为细胞的表达模式不同。当细胞做出选择时，或者遗传、化学或环境扰动时，就会表现出不同的基因表达模式。Monocle可以模拟出每个细胞所处的分化时间，并寻找随着分化时间逐渐升高或降低的基因表达热图和分布图。BEAM (Branched expression analysis modeling) 是一种统计方法，用于寻找以依赖于分支的方式调控的基因。

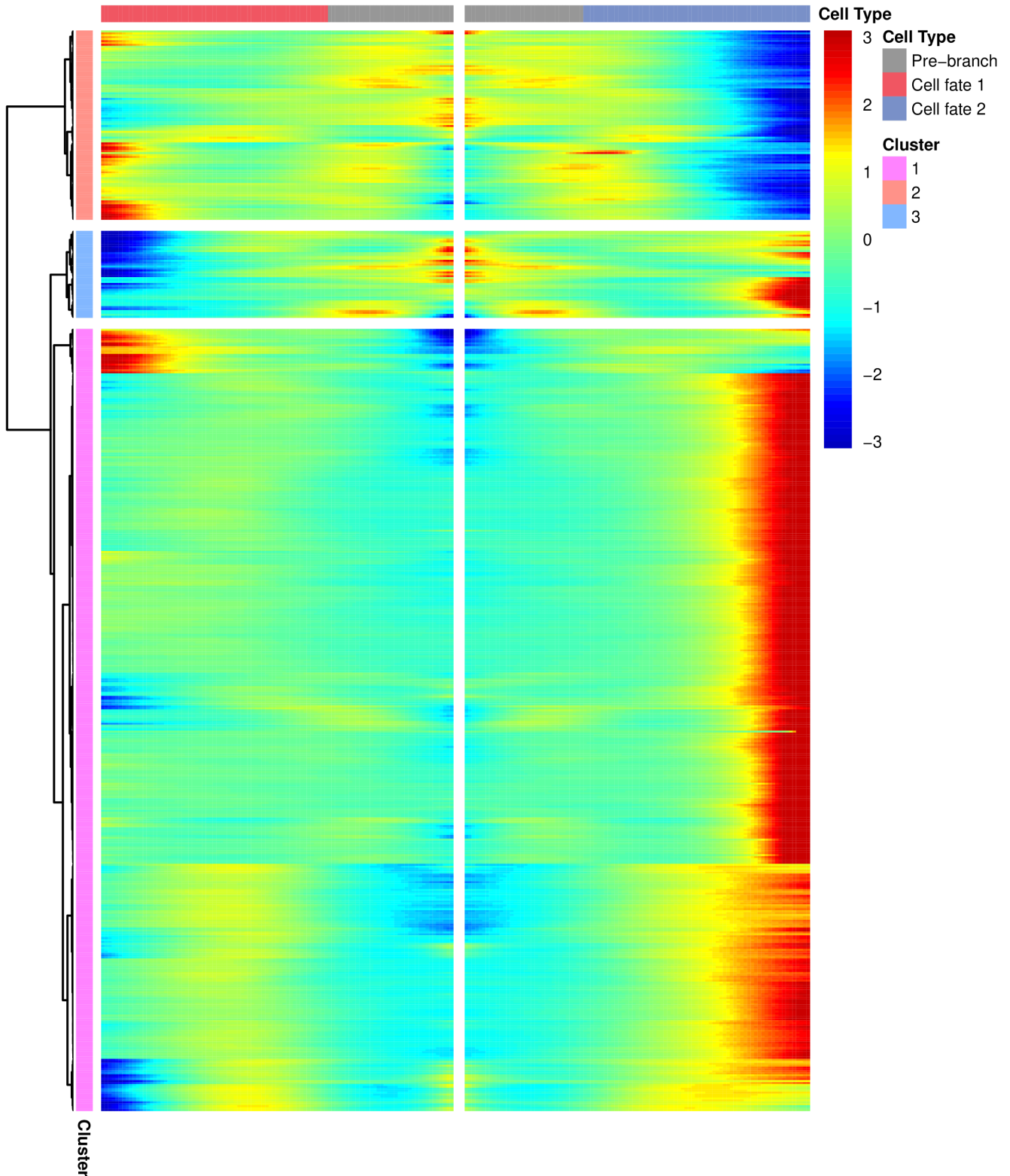


图2.6.2-2 BEAM统计分析图

该热图显示的是同一时间点两个谱系的变化。热图的横坐标展示了在特定决策点（例如决策点1）上细胞的拟时序分布情况。每个决策点代表细胞在拟时序上的不同阶段或路径选择。celltype代表细胞的不同类型或状态。其中，Pre-Branch 指的是在决策点之前的细胞状态或阶段。Cell Fate表示在拟时序决策点之后的细胞状态，反映了细胞在经历拟时序过程后的最终状态。

2.7 PAGA分析

PAGA⁷ (Partition-based graph abstraction) 是一种用于单细胞转录组数据的拟时序 (pseudotime) 分析方法，它旨在通过分析细胞状态的连通性和过渡，揭示细胞群体之间的动态变化过程。主要用于处理细胞在复杂发育过程中的轨迹推断。与传统的拟时序方法相比，PAGA 通过网络图的形式展示了细胞群体的关系，并通过图算法推断细胞群体间的过渡和发展路径。与传统的拟时序分析方法（如 Monocle 或 Slingshot）不同，PAGA 不强制要求每个单独细胞都具有明确的拟时序位置。PAGA 更侧重于细胞群体间的关系，并通过细胞群体间的动态联系来推断细胞状态的变化。

PAGA LTNN-graph

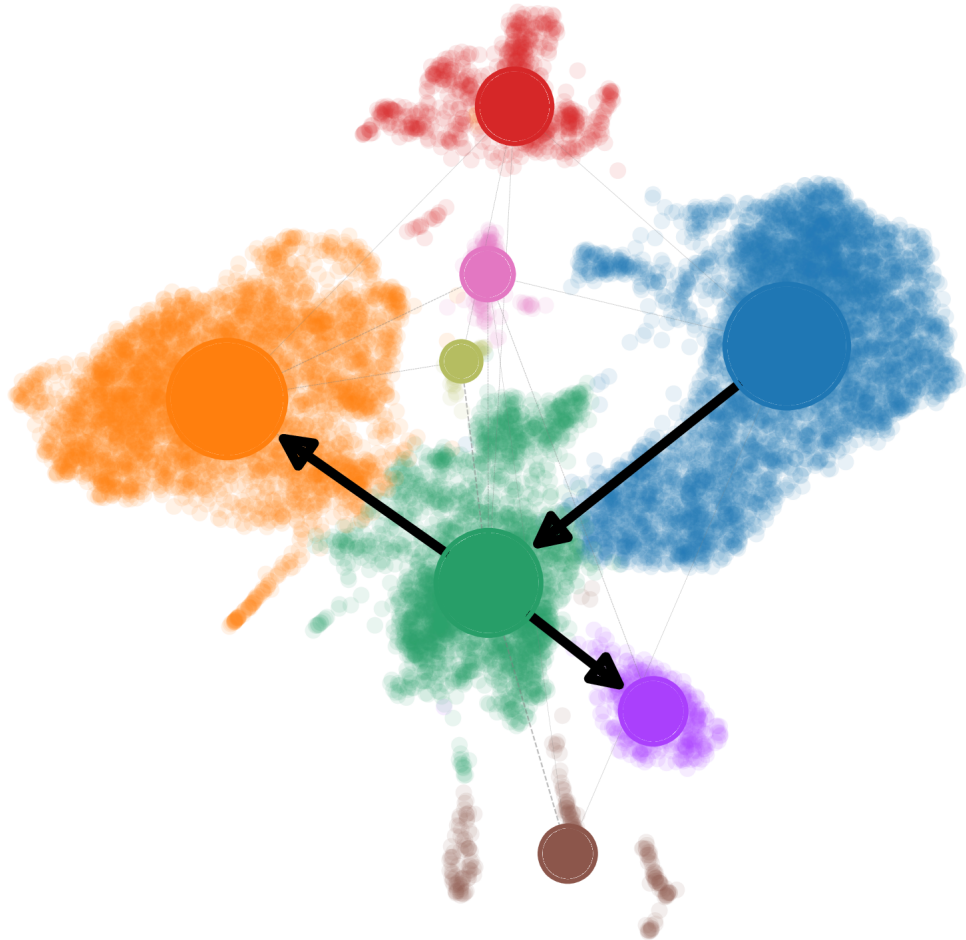


图2.7 PAGA拟时序分析图

箭头方向代表分化方向。

2.8 RNA velocity分析

RNA velocity⁸ (RNA速率)，即RNA丰度的时间导数，可以通过区分标准单细胞RNA测序中未剪接和剪接的mRNA来预测估计 (mRNA剪接动力学)。

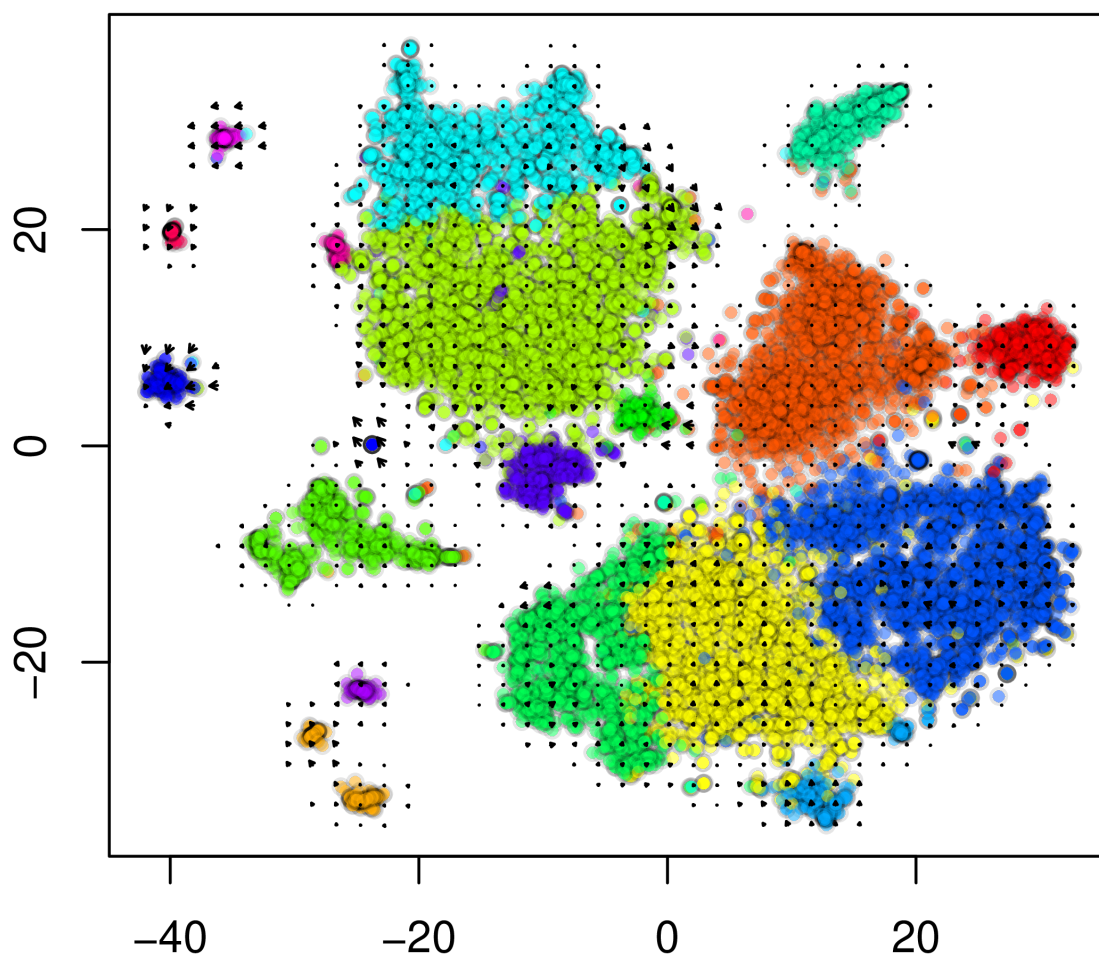


图2.8 RNA velocity分析图

X轴表示的是tSNE第一成分，Y轴表示的是tSNE第二成分。箭头方向表示发育方向。不同颜色代表不同的细胞类型

2.9 细胞通讯分析

细胞通讯分析是一种研究不同细胞类型之间如何通过信号分子（如配体和受体）进行相互交流和调控的分析方法。它在揭示细胞间相互作用的机制，理解组织和器官如何协调运作方面具有重要意义。

2.9.1 细胞间受体配体分析

使用R包-CellChat⁹来分析和可视化细胞间的通讯关系。CellChat含有高质量的信号分子互作数据库，对细胞进行差异表达分析和统计测试，推断不同细胞之间特定的通讯信号并可视化。这个部分的可视化主要是统计不同细胞类群之间的受体-配体对数（interactions）和通讯强度（strength）。

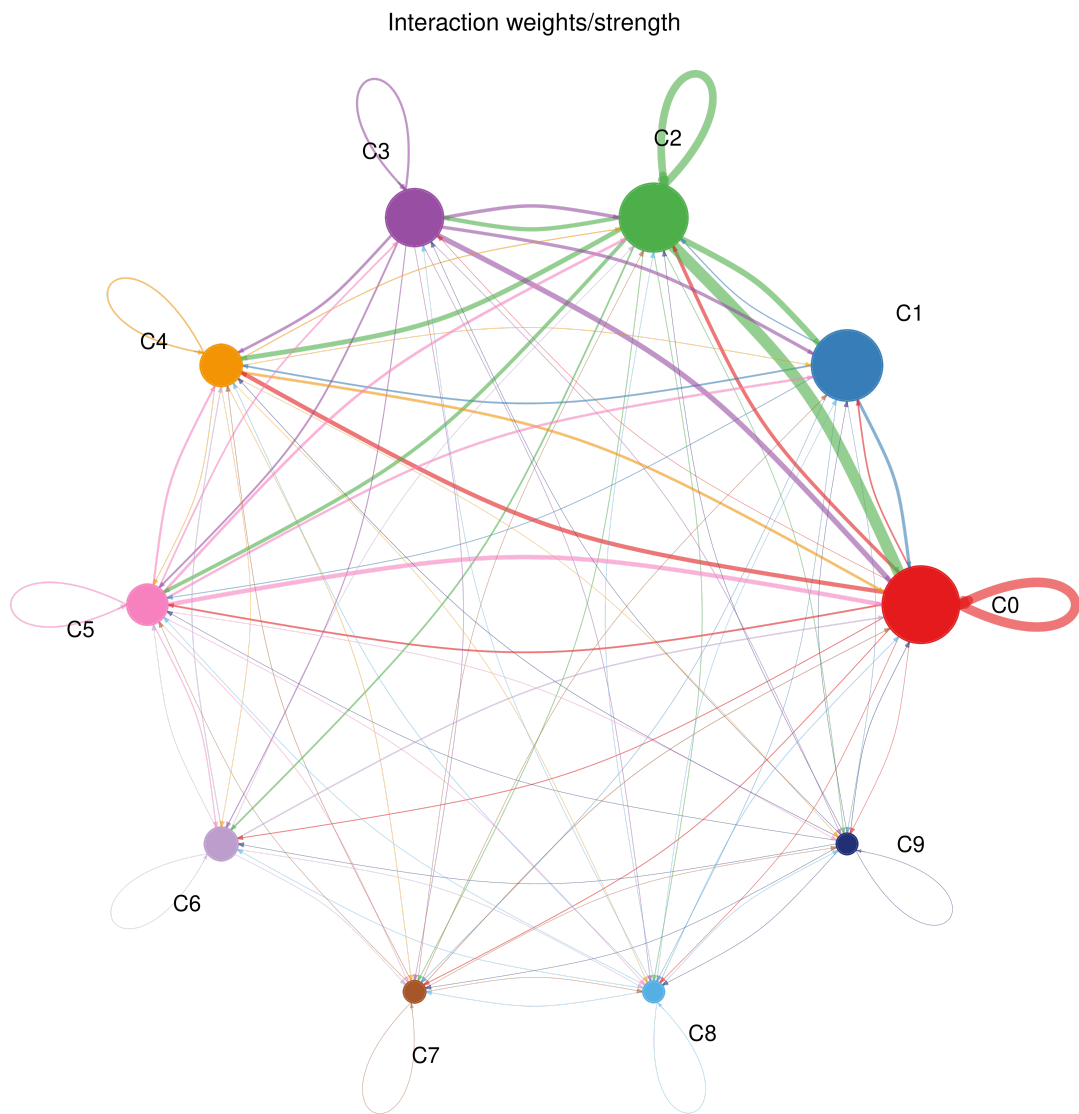


图2.9.1-1 细胞类型之间互动强度图

线越粗越互动越强，线的颜色和点的颜色相同表示点为source细胞，另一端为target细胞。

2.9.2 多个配体-受体介导的细胞互作关系分析

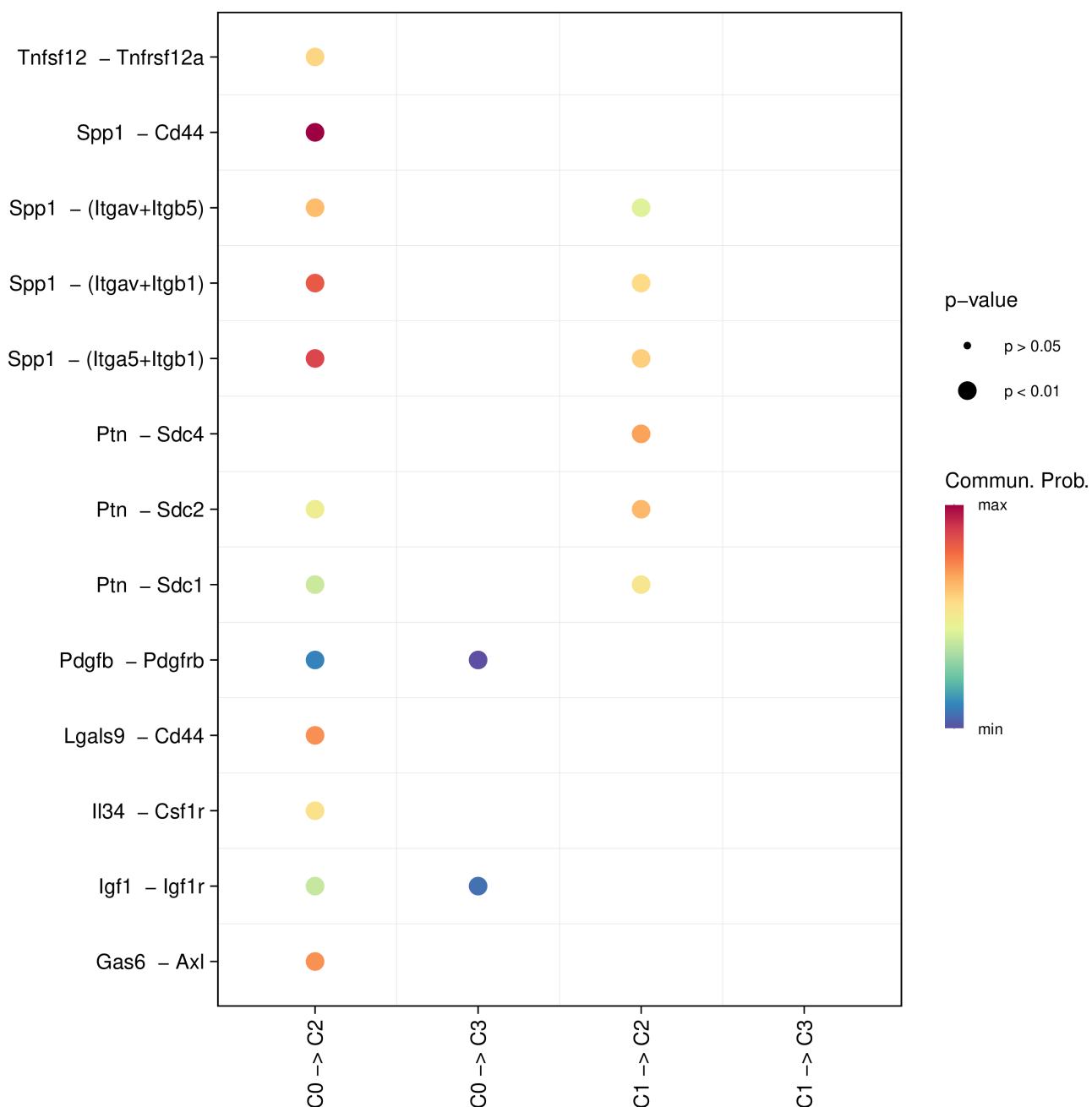


图2.9.2-1 细胞间配体-受体互作图

X轴代表不同细胞之间的互作方向。Y轴代表特定的配-受体对。圆点的不同颜色代表通讯概率。点的大小代表通讯显著性的P值。

2.10 SCENIC单细胞转录因子分析

细胞命运的控制依赖于一系列通常被称为转录因子 (Transcription Factors, TFs) 的蛋白质, TF通过与DNA特定区域结合, 促进或阻碍DNA的转录过程。了解TF对于了解细胞功能及疾病进程等生命活动有重要意义。2017年11月, SCENIC¹⁰发表于Nature Methods, 作为主流的单细胞转录因子分析工具, SCENIC在单细胞转录组数据构建基因调控网络、转录因子分析以及细胞状态鉴定方面表现优异。目前主要有R和python两个版本, 支持分析人、小鼠和果蝇的数据。

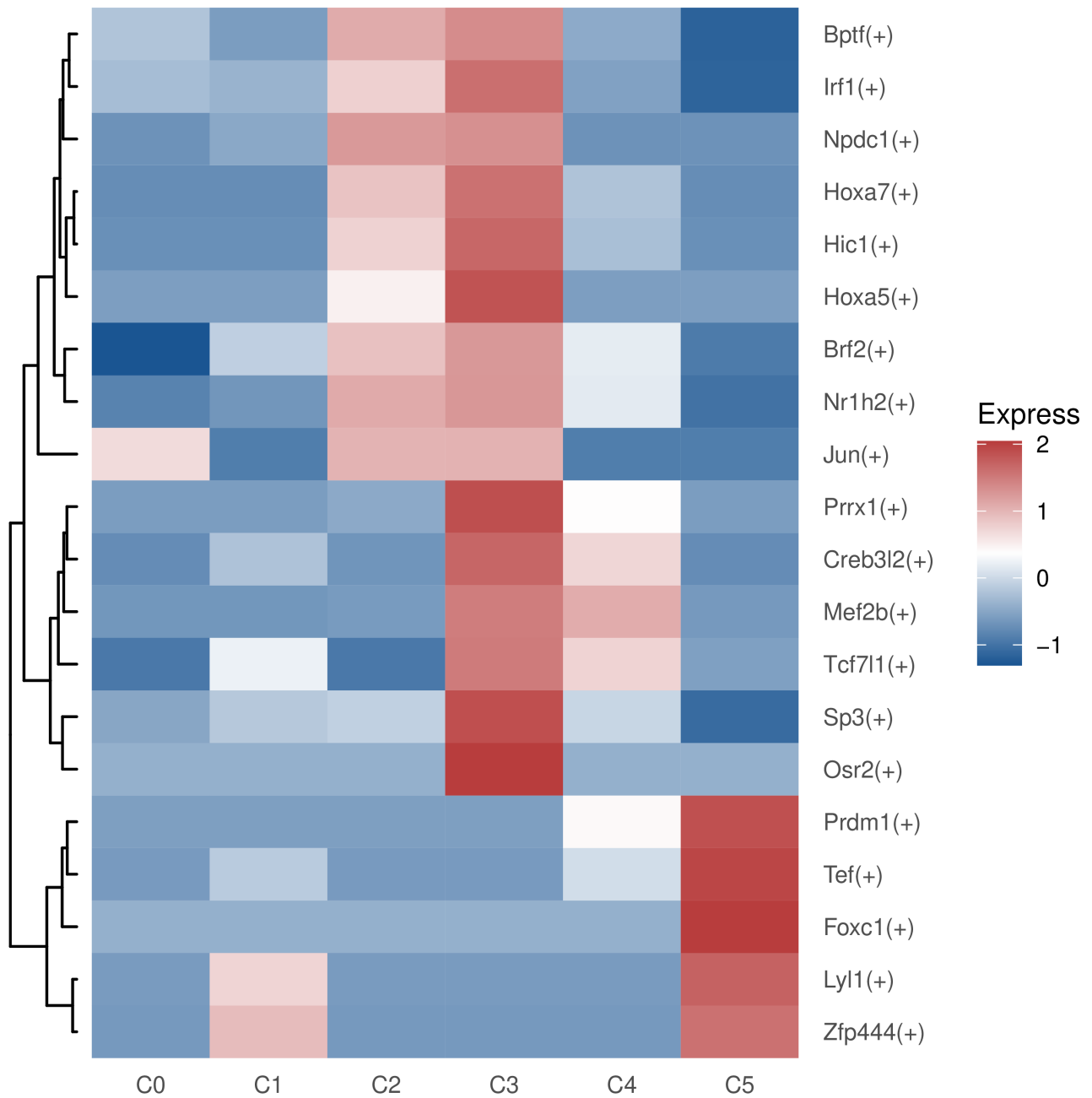


图2.10-1 转录因子热图

横轴代表每一个细胞类型，纵轴是pySCENIC选出来的、所有的regulon，颜色则代表了regulon的活跃程度。越接近深红色，活跃程度越高，越接近深蓝色，活跃程度越低。此外，这张图还分别展示了对每一个regulon和每一个细胞类型的聚类结果。

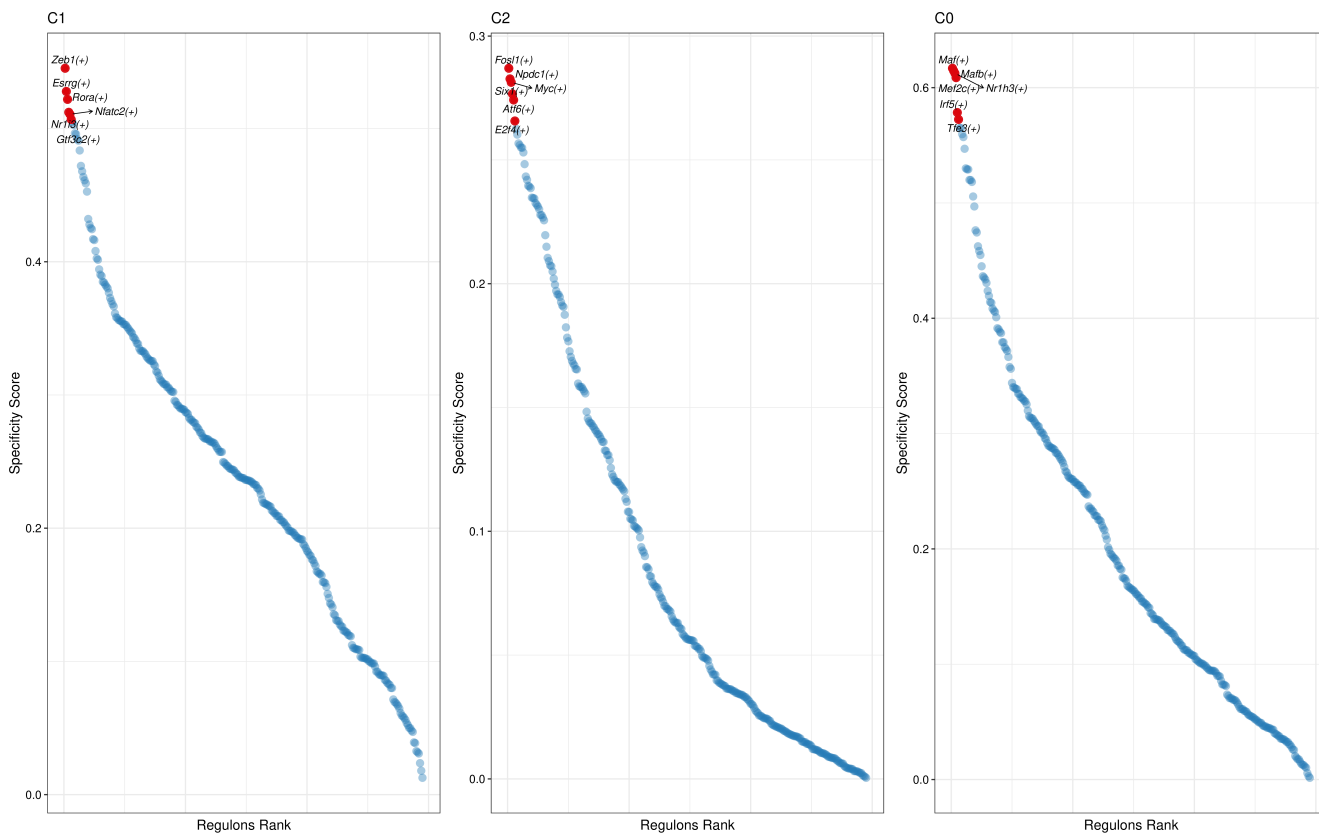


图2.10-2 单个cluster中regulon的RSS分布图

X轴代表细胞类型。Y轴代表RSS值，即转录因子活性分数。每个图展示了Top 5 TF，用红色点标注。

3 分析软件说明

本项目用到的分析软件见下表：

表4 分析软件列表

分析工具名称	版本	分析内容	说明
dnbc4tools	v2.1.2	reads质控,序列比对	dnbc4tools是华大开发的一款用于单细胞数据的分析工具
DoubletFinder	v2.0.3	鉴定双细胞	依据泊松分布的统计原理可以计算每个样本doublets的数量,结合之前的细胞pANN值排序,过滤doublets。
SoupX	v1.6.2	去除基因组污染	环境RNA污染简单来说就是由于实验原因导致某些基因的转录本扩散到大多数细胞,使得部分基因在大多数细胞中均出现高表达的现象,从而影响我们对细胞类型的定义以及其它分析。
Seurat	v4.3.0.1	单细胞分群、差异基因分析,细胞通讯分析	Seurat是一个专用于单细胞测序分析与时空组测序分析的R包,其功能涵盖了数据过滤、数据标准化、去批次效应、细胞聚类分析、差异基因分析、细胞周期分析等,功能全面强大、在单细胞测序与时空组测序中获得了广泛应用。
SingleR	v2.4.1	细胞类型注释	SingleR是一个用于对单细胞RNA-seq测序数据进行细胞类型自动注释的R包。它通过给定的具有已知类型标签的细胞样本作为参考数据集,对测试数据集中与参考集相似的细胞进行标记注释。
monocle	v2.32.0	拟时序分析	Monocle是一个用来进行拟时序分析的R包,根据不同细胞类群之间其基因表达差异的情况,获得细胞谱系发育结果,构建细胞随着一个虚拟时间顺序的变化轨迹,以此重现细胞随时间变化的过程。
CellChat	v1.6.1	细胞通讯分析	CellChat是一种利用scRNA-seq数据进行推断,可视化和分析细胞间的通讯关系的R包。该软件含有高质量的信号分子互作数据库; CellChat使用质量作用模型,对细胞进行差异表达分析和统计测试,推断不同细胞之间特定的通讯信号。
pySCENIC	v0.12.1	转录因子分析	SCENIC全称Single-Cell rEgulatory Network Inference and Clustering,即转录因子分析,用于构建GRN, gene regulatory network基因调控网络。
omicverse	v1.6.8	PAGA分析	基于分区的图抽象(Partition-based graph abstraction)生成单个细胞的拓扑结构并保留映射。高维基因表达数据降维后计算邻域关系的相关距离度量来表示kNN图(pca和欧氏距离)。
velocyto.R	v0.6	RNA速率分析	细胞中的RNA在转录后会存在剪切成熟的过程,根据对不同成熟度的RNA的分析,确定细胞在发育过程中的先后。

4 参考文献

1. Liu C, Wu T, Fan F, et al. A portable and cost-effective microfluidic system for massively parallel single-cell transcriptome profiling[J]. BioRxiv, 2019: 818450. [↵](#)
2. McGinnis C S, Murrow L M, Gartner Z J. DoubletFinder: doublet detection in single-cell RNA sequencing data using artificial nearest neighbors[J]. Cell systems, 2019, 8(4): 329-337. e4. [↵](#)
3. Young M D, Behjati S. SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data[J]. Gigascience, 2020, 9(12): gaaa151. [↵](#)

4. Hao Y, Hao S, Andersen-Nissen E, et al. Integrated analysis of multimodal single-cell data[J]. Cell, 2021, 184(13): 3573-3587. e29. [↵](#)
5. Aran D, Looney A P, Liu L, et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage[J]. Nature immunology, 2019, 20(2): 163-172. [↵](#)
6. Qiu X, Mao Q, Tang Y, et al. Reversed graph embedding resolves complex single-cell trajectories[J]. Nature methods, 2017, 14(10): 979-982. [↵](#)
7. Van de Sande B, Flerin C, Davie K, et al. Zeng Z, Ma Y, Hu L, et al. OmicVerse: A single pipeline for exploring the entire transcriptome universe[J]. bioRxiv, 2023: 2023.06. 06.543913. [↵](#)
8. La Manno G, Soldatov R, Zeisel A, et al. RNA velocity of single cells[J]. Nature, 2018, 560(7719): 494-498. [↵](#)
9. Jin S, Guerrero-Juarez C F, Zhang L, et al. Inference and analysis of cell-cell communication using CellChat[J]. Nature communications, 2021, 12(1): 1088. [↵](#)
10. Van de Sande B, Flerin C, Davie K, et al. A scalable SCENIC workflow for single-cell gene regulatory network analysis[J]. Nature protocols, 2020, 15(7): 2247-2276. [↵](#)

5 联系我们

西安浩瑞基因成立于2019年，公司引入了三代测序平台——2台PacBio Revio和7台Sequel II设备，致力于深耕动植物基因组学、转录组和微生物组学研究的科研技术服务。2024年，与华大智造携手共建西北首家DCSLab组学前沿实验室，引入DNBSEQ-T7测序平台，开展基于二代测序的单细胞转录组、时空转录组等前沿技术服务。凭借专业的一站式多组学技术，为广大科研客户提供专业、高效、可靠的组学科研技术服务

联系方式

热线电话：+86 029-89303503

官方网站：www.xahorizon.cn

邮箱：project@xahorizon.cn

地址：陕西省西安市沣东新城中兴深蓝科技产业园A区2号楼3层

