

下一步分析建议

基于单细胞转录组测序 的下一步分析建议



- www.oebiotech.com -



声明

欧易生物在此声明：

一、本方案仅供客户作为项目研究的参考。

二、本方案不排除以下可能：

- 1) 方案存在未列入其中的项目风险；
- 2) 方案的合理性和可行性存在部分疏漏；
- 3) 项目的实际结果产出与预期存在一定程度的偏差；
- 4) 项目存在其他未知的风险因素。

三、项目实际实施方案，请客户根据具体情况予以确定。



一、分析方向和注意点：

1. 没有对照组的项目
2. 有不同分组的项目
3. 注意点

单细胞转录组数据挖掘视频链接：https://www.bilibili.com/video/BV1Xa41127B6/?spm_id_from=333.337.search-card.all.click

方向1：没有对照的项目（只有一个分组）

下一步分析建议：

1. 细胞类型鉴定：**报告中的reference celltype结果，不建议直接用与后续分析，仅供参考。因此需要老师参考一发表文章的标志基因，进行进一步鉴定后（可以loupe软件鉴定），再安排下游分析（细胞类型鉴定可以联系欧易公司技术支持辅助鉴定）；**

2. 图谱分析：

2.1 有研究基础，用单细胞验证的项目：

1) 如果有重点关注的细胞类型，可以直接对关注细胞做降维聚类分析，探究细胞亚型特征（如果没有，每种细胞类型的细胞数量不低于100个，都可以根据情况做亚型分析）；

2) 关注亚型的基因表达及GO/KEGG富集情况（这部分可以做top marker筛选或富集分析），探究前期关注的基因在亚群中的表达情况，作为验证；

3) 针对具有亚型之间存在功能转化（如naïve T向Exhausted T转化）的细胞，可以做拟时序分析，探究分化/转化轨迹，并挖掘调控基因，关联其他组学看不到的信息，可以考虑通过基因或者通路关联（比如不同组学都验证了同一条通路或基因的关键作用）；

方向1：没有对照的项目（只有一个分组）

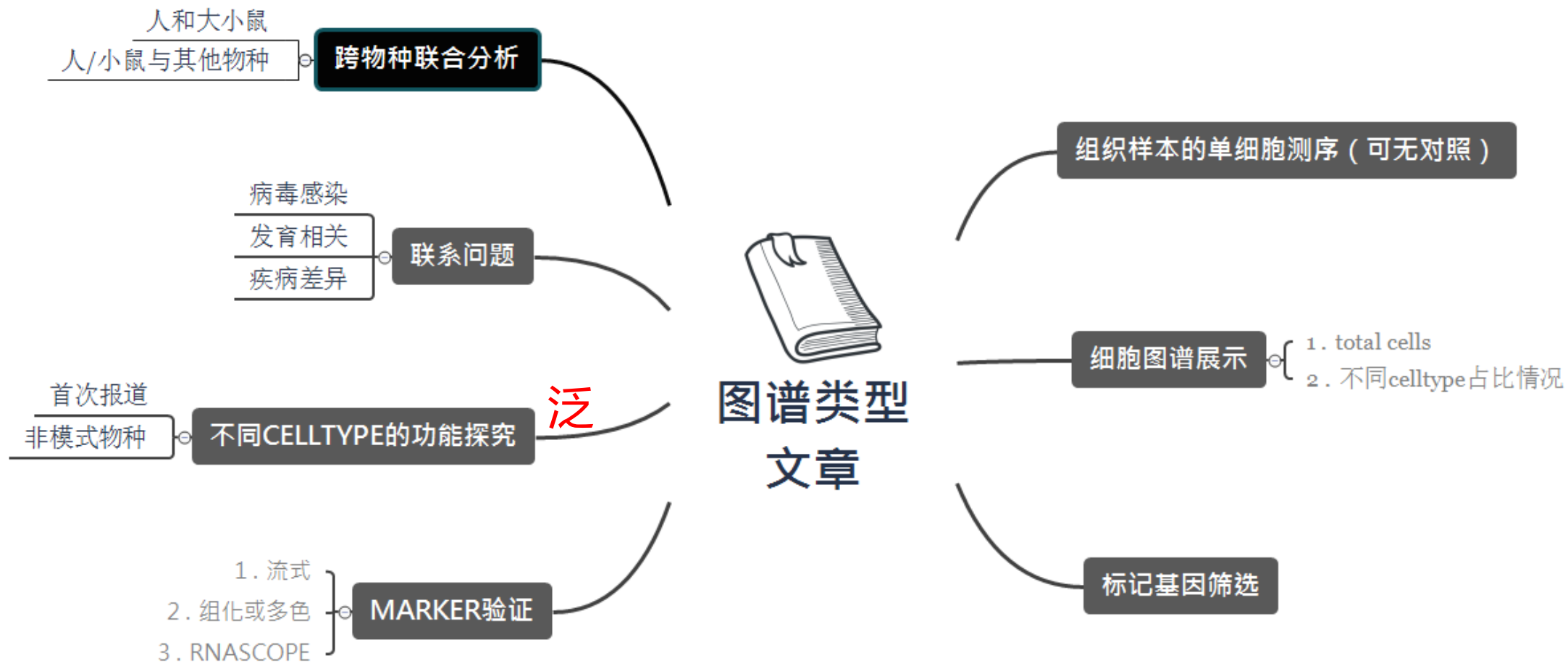
- 4) 关键的信息，受到哪些TF的调控，及这些TF主要是集中在那一群关键的细胞，可以来考虑通过scenic分析来实现；
- 5) 探究不同细胞类型/亚型之间的通讯信号交流，明确这些细胞的变化，受到哪些影响，可以通过cellphoneDB或cellchat实现。

2.2 没有研究基础-针对难收/珍贵临床样本等：

- 1) 可以采用泛分析，即尽可能详细的分析，将每种细胞类型分别做subcluster分析，探究其基因表达，功能，富集通路等信息
- 2) 其他分析参考上述2) -5) 的分析内容
- 3) 这类项目如果要发高分文章，尽量结合临床，如已知疾病相关基因和各细胞类型/亚型的相关性、基于单细胞筛选的marker，结合TCGA/公开发表的相关数据进行联合分析，预测生存曲线等。

方向1：没有对照的项目（只有一个分组）

技术路线参考



方向2：有不同分组的项目

主要针对：健康/对照，KO/WT，不同时间点，疾病亚型，不同治疗阶段等

下一步分析建议：

1. 细胞类型鉴定：**报告中的reference celltype结果，不建议直接用与后续分析，仅供参考。因此需要老师参考一发表文章的标志基因，进行进一步鉴定后（可以loupe软件鉴定），再安排下游分析（细胞类型鉴定可以联系欧易公司技术支持辅助鉴定）；**
2. 普筛（泛分析）-针对没有明确关注点的老师
 - 1) 基于1的结果，首先统计分析各细胞类型在组间的差异（主要看比例差异，或各细胞类型的差异基因），这些差异的细胞，考虑和要探究的生物学问题具有相关性；
 - 2) 基于差异的细胞，可以考虑进行亚型分析，从subcluster角度出发，分析某种细胞类型的异质性，找到和XXX疾病具有显著相关的细胞亚型，探究其基因表达和功能特征，尤其相关通路和功能的富集情况；
 - 3) 鉴于某些细胞亚型可能存在比较大的组间差异，而这些亚型细胞可能由于其他细胞亚型转化而来，因此，可以考虑做拟时序，进一步挖掘不同功能亚型之间的转化/细胞分化，与疾病进展/发育过程的关系，并筛选轨迹调控的关键基因。这一步也可以考虑基于先验知识，对存在分化关系的细胞之间进行伪时间分析；

方向2：有分组的项目

4) 疾病的发生发展，甚至是发育过程，在机体内是一个连续的动态变化网络，因此在后续分析的过程中，可能会发现很多细胞类型甚至亚型都存在变化，那么这些变化的细胞之间是否存在信号交流，**可以考虑通过细胞通讯分析构建相关细胞的相互作用网络，探究细胞间关键的配体-受体信息，比较XX组和XXX组组间细胞通讯网络构成的差异；**

5) 此外，对于泛分析的项目，可以考虑从宏观角度分析某种细胞类型/亚型中转录因子的调控网络，为下游验证/临床转化的biomarker库储备信息。

3. 有明确分析目标-针对特定细胞类型/分选某种细胞类型分析

- 1) 即使分选的细胞，最好也做一下鉴定，避免分选纯度不高导致杂细胞混进来；
- 2) 针对重点关注的细胞做subcluster分析，探究亚群的异质性及在不同分组中的差异。对于通过单细胞数据验证的老师，可能会基于前期验证的基因，对subcluster的结果进行分类，如分成geneA+XX cells或gene A-XX cells
- 3) 其余分析思路和上述2) -5) 一致。

方向2：有分组的项目

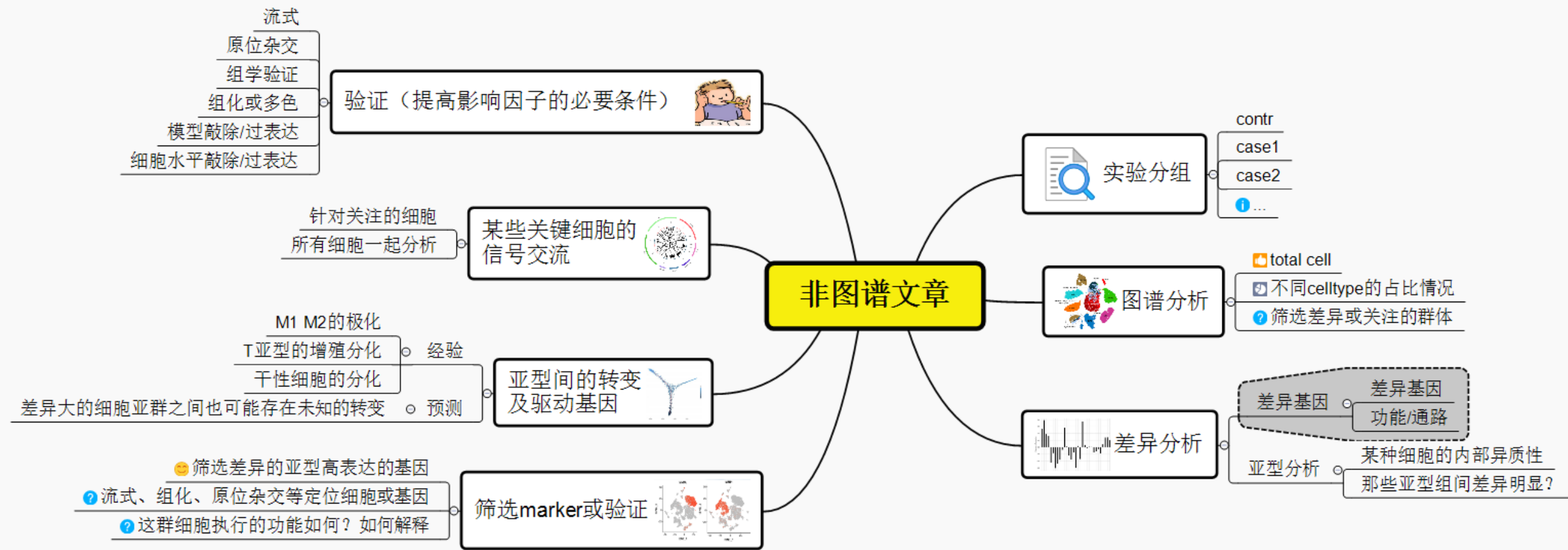
4. 后续验证（提高文章影响因子的重要方式）

1) 简单的验证，可增加免疫荧光实验，流式等；

2) 深入的验证，可在细胞层面进行敲降，过表达，共培养等；或针对关键的基因进行敲除，过表达等造模实验。

方向2：有分组的项目

技术路线参考



注意点

1. 细胞亚型鉴定和命名原则是什么？

亚型命名，因其个性化程度比较高，**需要老师自己完成**

- 1) 有明确的细胞亚型名字的，可以直接参考文献中的基因进行鉴定，如T亚型可分为naïve, memory, Cytotoxic等；
- 2) 对于没有公认亚型名称，且存在异质性的细胞，可以考虑直接按照无监督聚类得到的subcluster进行分析
- 3) 基于2)，也可以从top基因中挑选相关基因，对其subcluster命名，如geneA+XX cells；
- 4) 对于有明确关注基因的老师，可以考虑通过对应的基因，将subcluster按照geneA+/hi/lo XX cells进行标记。

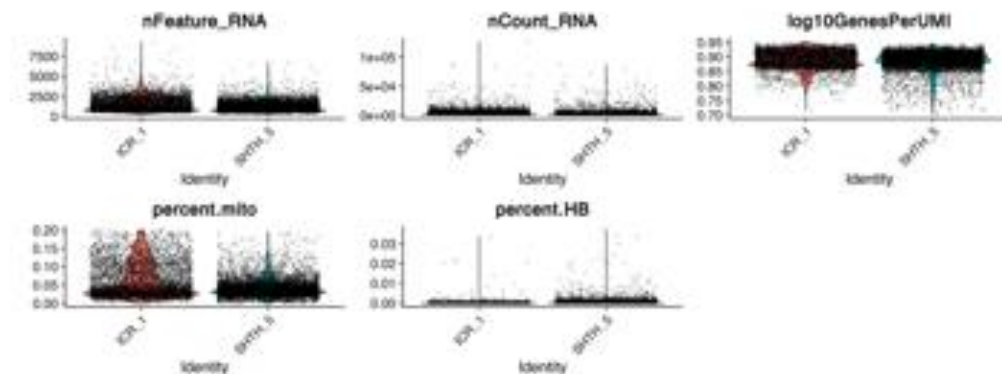
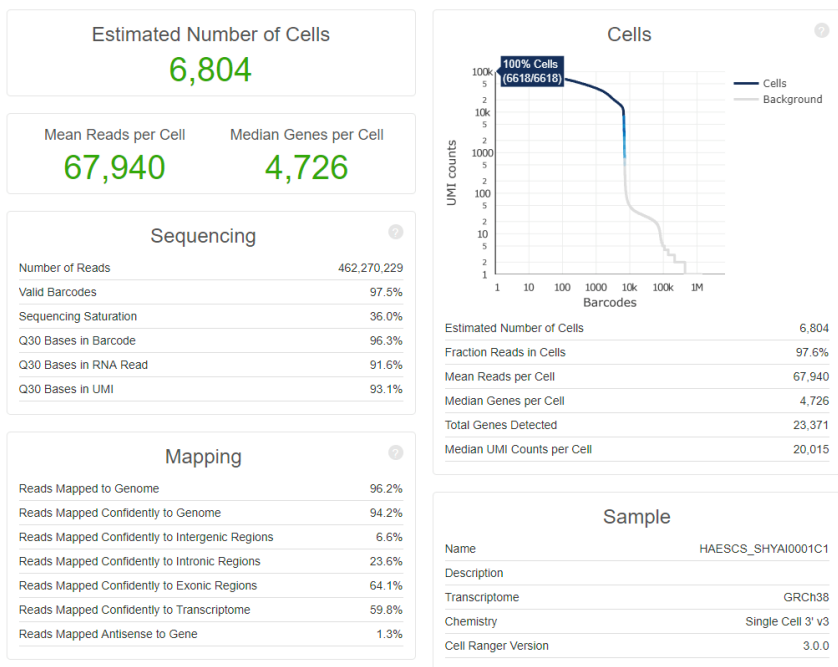
2. 什么情况下可以考虑用拟时序分析？

- 1) 已知存在分化关系的细胞
- 2) 针对细胞亚型，组间存在差异的亚型之间

3. 如何基于细胞间通讯分析结果筛选想要的信息？

- 1) 锁定细胞类型
- 2) $P < 0.05$ ，表达量/通讯概率越高越好

二、分析内容参考——单细胞转录组之数据质控和标准化分析



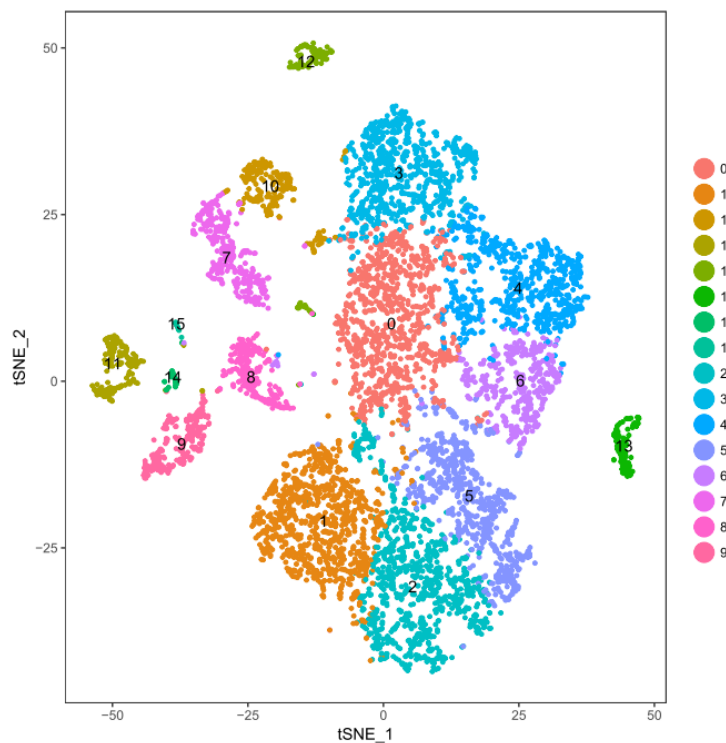
二次质控

- 剔除异常数据；
- 剔除多细胞、双细胞或者未结合上细胞的数据；
- 剔除死细胞等。

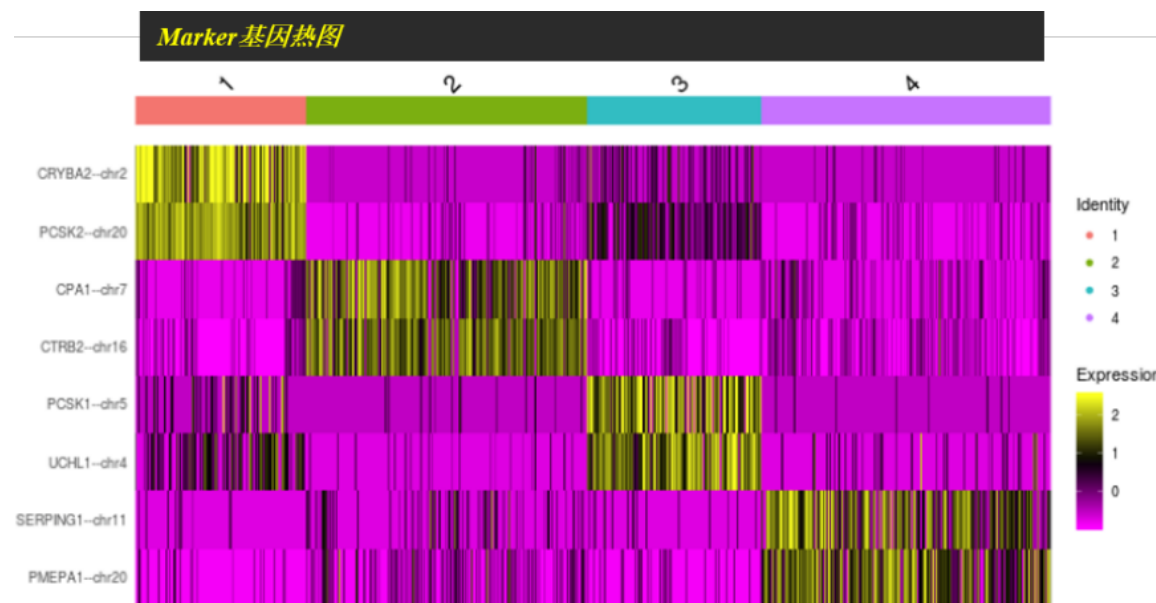
10x Genomics 官方软件 Cell Ranger

参考：分样本进行单细胞转录组测序，数据质控和标准化分析（墨卓或其他平台分析思路基因一致）。

二、分析内容参考——单细胞转录组之降维聚类(marker鉴定)



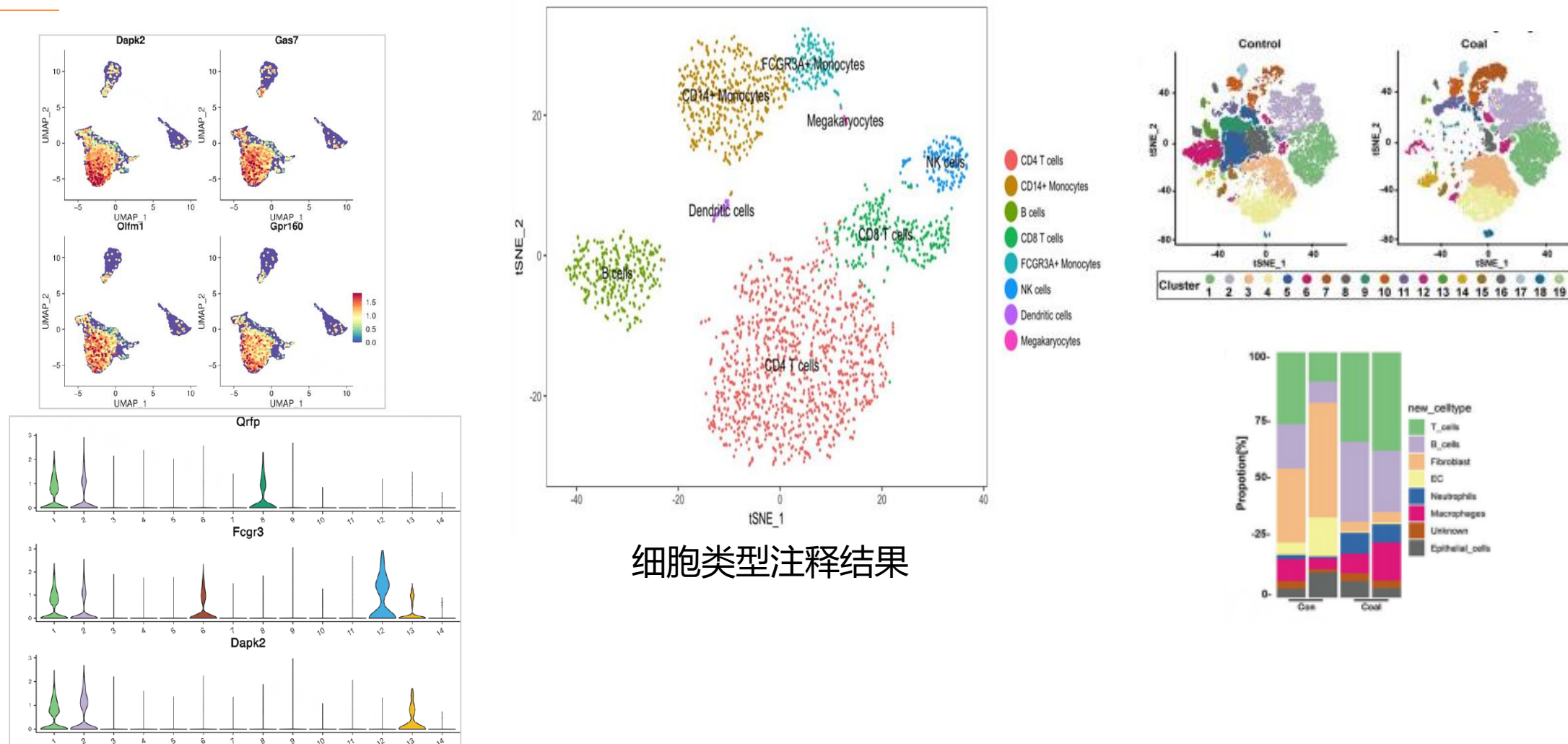
tSNE降维聚类图



Marker基因热图

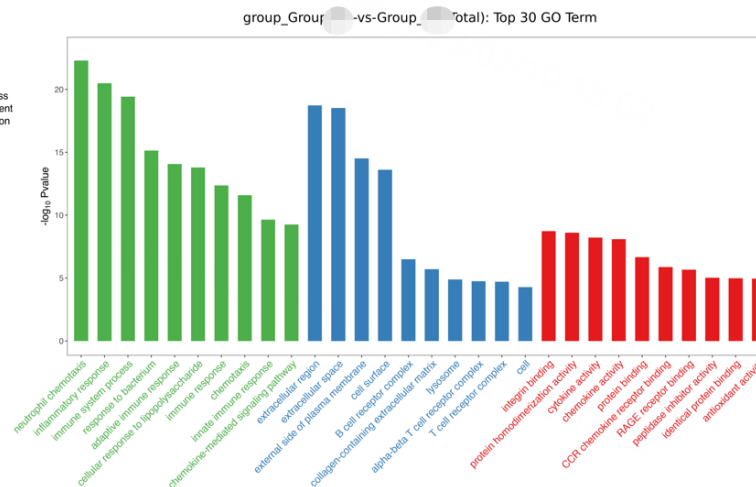
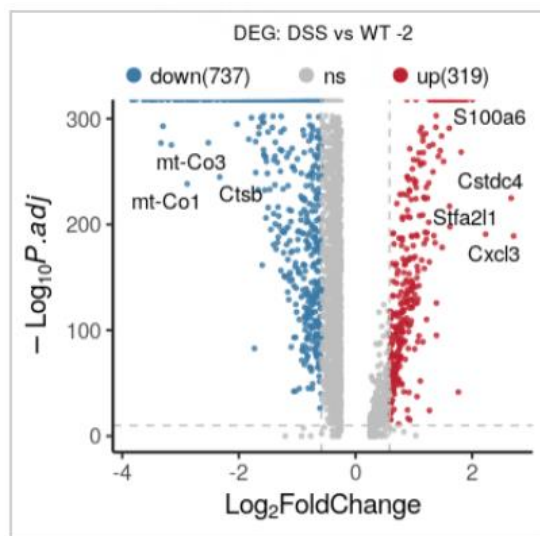
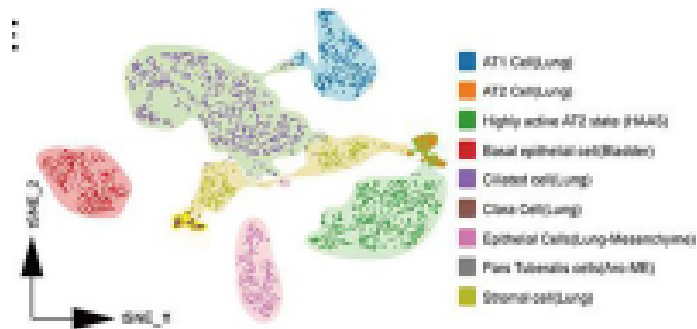
参考：将合格的细胞按照基因表达谱相似性进行聚类分析，不同群细胞以不同颜色区分，并鉴定marker基因。

二、分析内容参考——单细胞转录组之细胞类型鉴定和占比



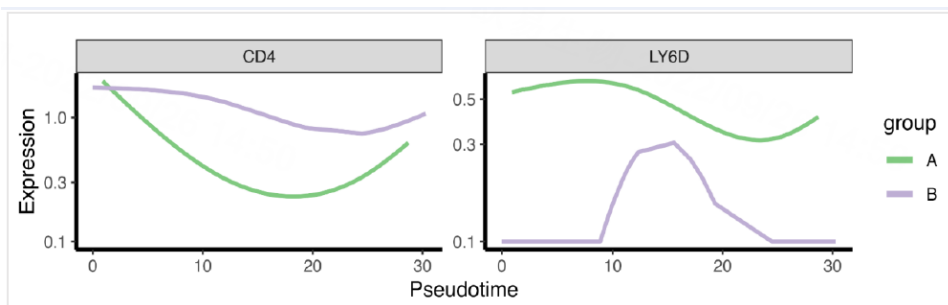
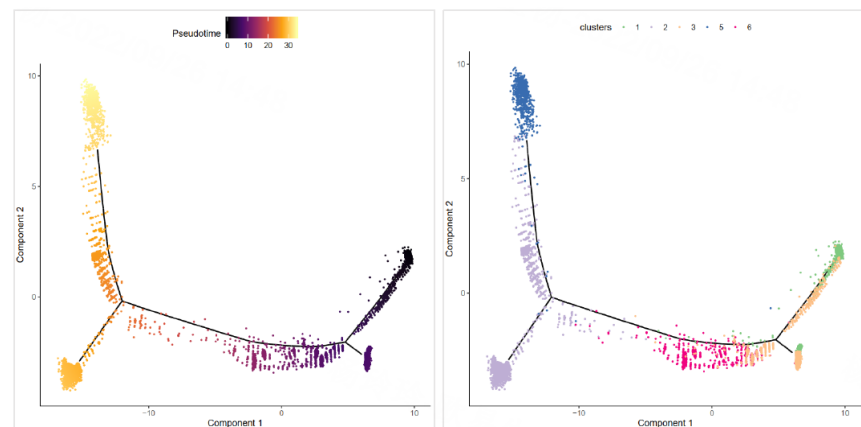
参考：以人工注释为主，数据集鉴定为辅，鉴定各组样本细胞类型，探究不同分组间的细胞组成和占比等。后续可针对这些差异的细胞类型进行亚型的划分。

二、分析内容参考——单细胞转录组之差异基因和富集分析

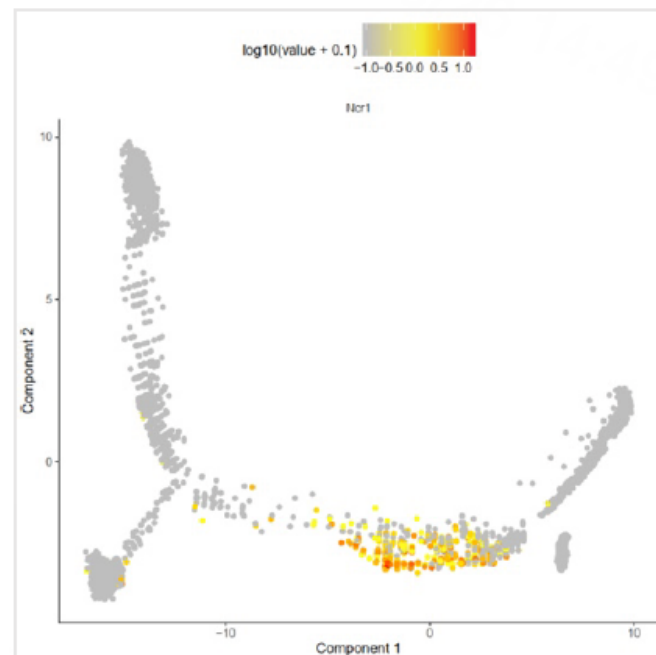
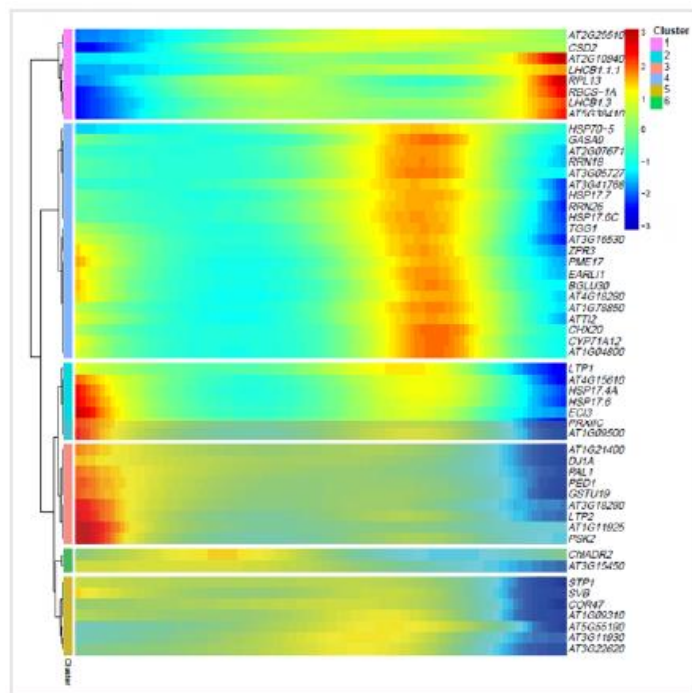


参考：可重点关注前述差异的细胞细胞类型/亚型。分析在这些细胞/细胞亚型中差异基因或通路并加以挖掘和验证。

二、分析内容参考——单细胞转录组之拟时序分析



拟时序分析

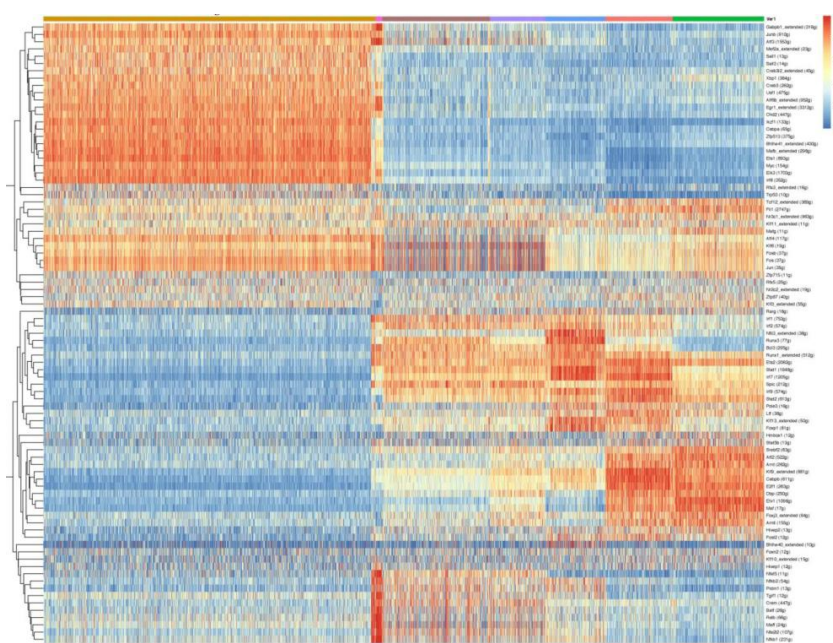


图中的每个点代表一个细胞，具有相似细胞状态的细胞被聚到一起，不同颜色标注的是不同的细胞状态，每个分支点代表着一个可能的细胞生物学过程决策点。

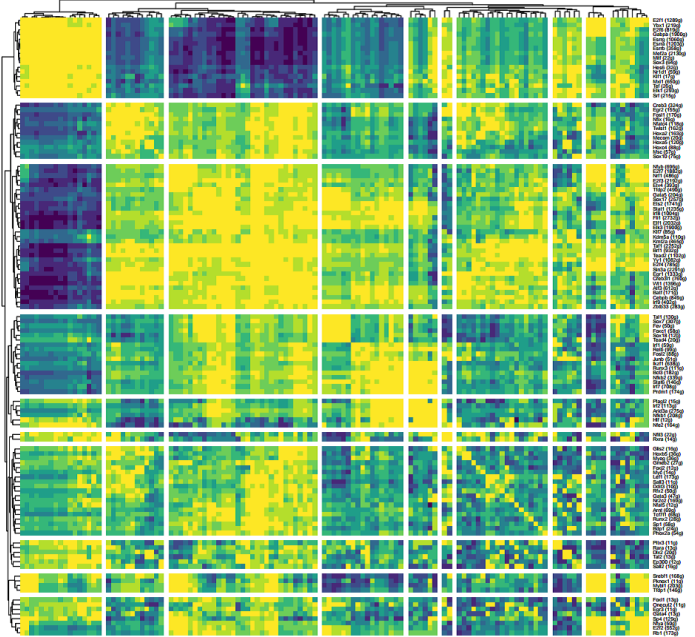
参考：针对前述差异的细胞类型/重点关注的细胞，进行拟时序分析，可以帮助判断不同处理之间指定细胞类型的变化轨迹，相对应的在这个过程中基因的变化，而这些变化的基因是否与疾病发生相关呢？推测疾病发生过程中细胞亚型的变化轨迹。

二、分析内容参考——单细胞转录组之SCENIC分析

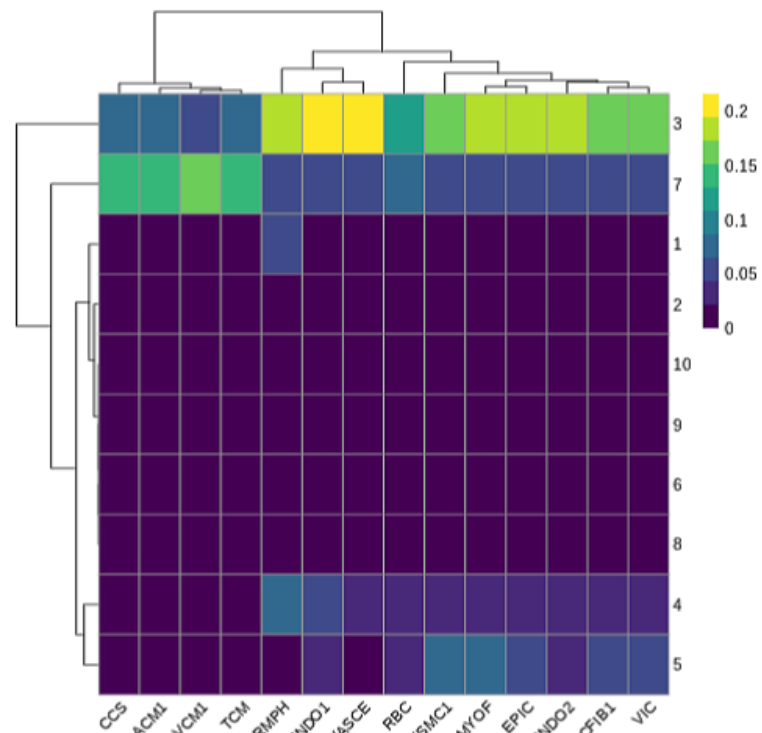
1. 识别单细胞转录组数据中的转录因子及其靶基因，鉴定Regulon;
2. 找到共同调控下游基因，负责细胞功能的高关联性Regulon;
3. 构建不同细胞类型中的基因调控网络;
4. 确定维持细胞特性的重要调控子。



Regulon活性热图



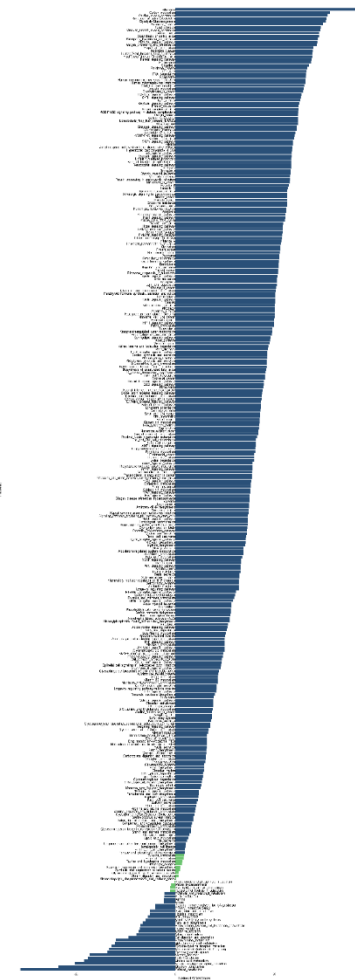
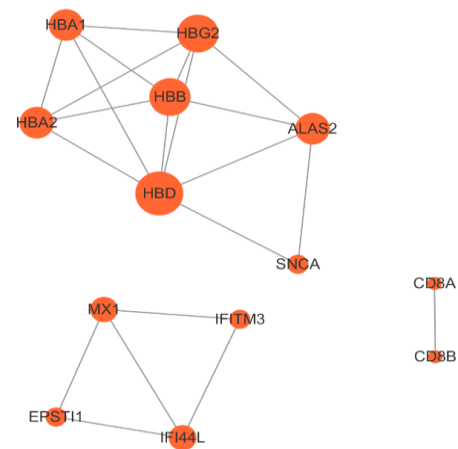
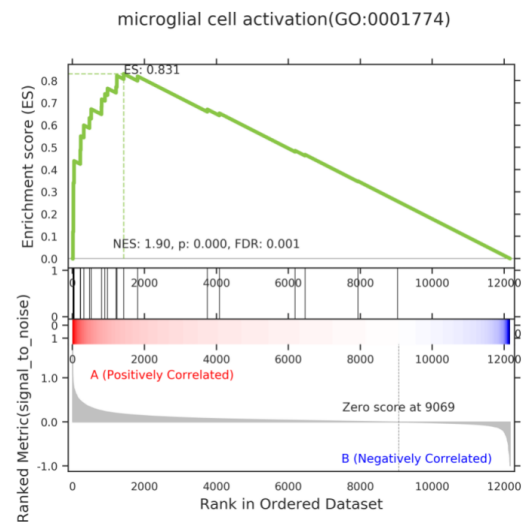
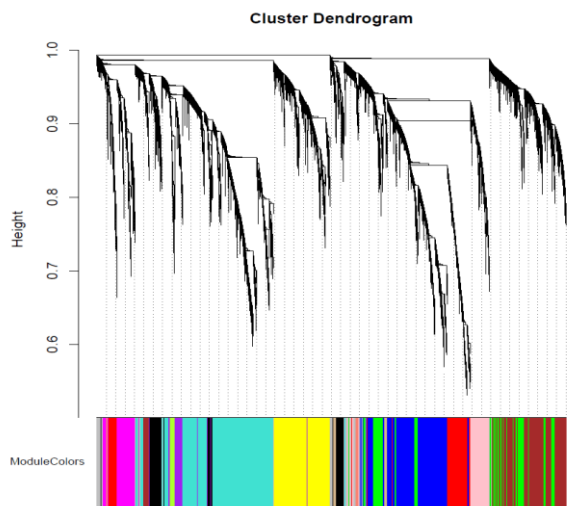
Regulon关联特异性指数(CSI)聚类图



CSI模块的活性热图

参考：探究关键细胞群中相关的调控子，寻找关键转录因子，从该角度进行解析。

二、分析内容参考——其他个性化分析



WGCNA (基因共表达)、PPI (蛋白互作)、GSEA/GSVA (基因集富集/变异分析)，可根据需求选择。

THANKS FOR LISTENING

感谢观看

