

专题4-单细胞数据分析

大纲

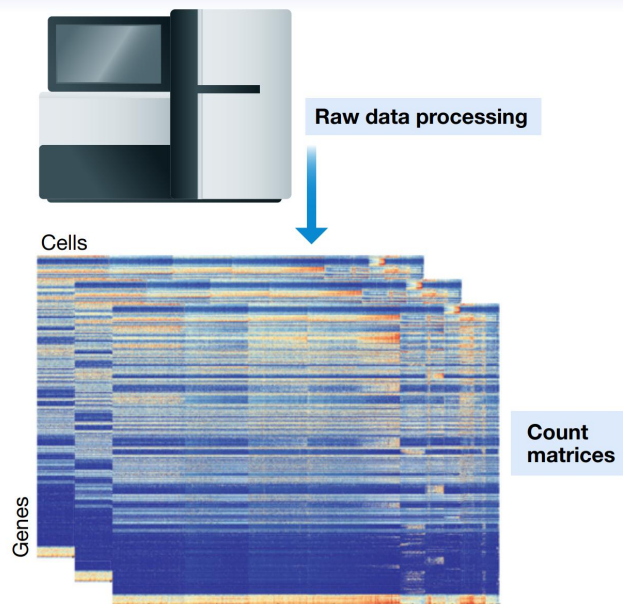
- 单细胞技术
- 单细胞数据分析模式
- 单细胞数据分析专题

单细胞的数据预处理

- 该流程主要就是数据预处理：质控、归一化、数据矫正、特征选择和降维。

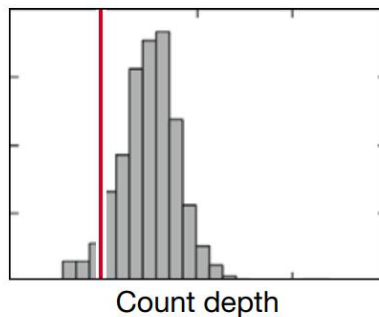
单细胞的数据预处理

PRE-PROCESSING

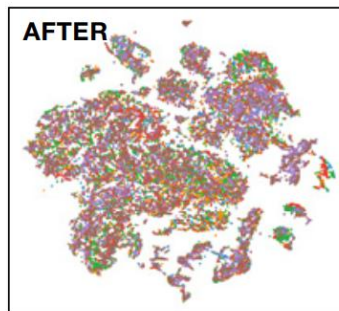
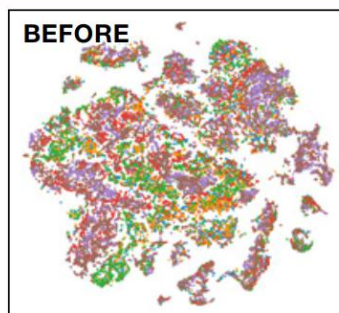


- 质控、归一化、数据校正、特征选择、降维、可视化。

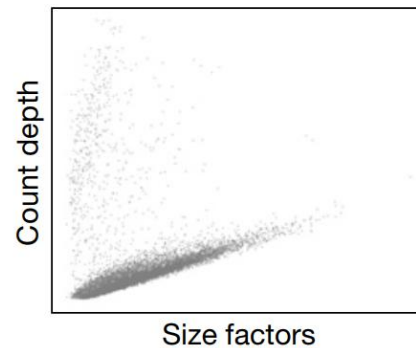
Quality control



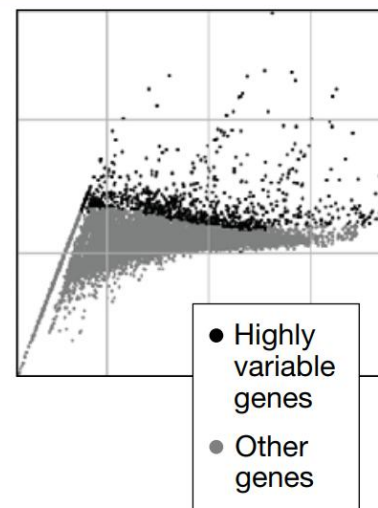
Data correction (e.g. batch)



Normalization



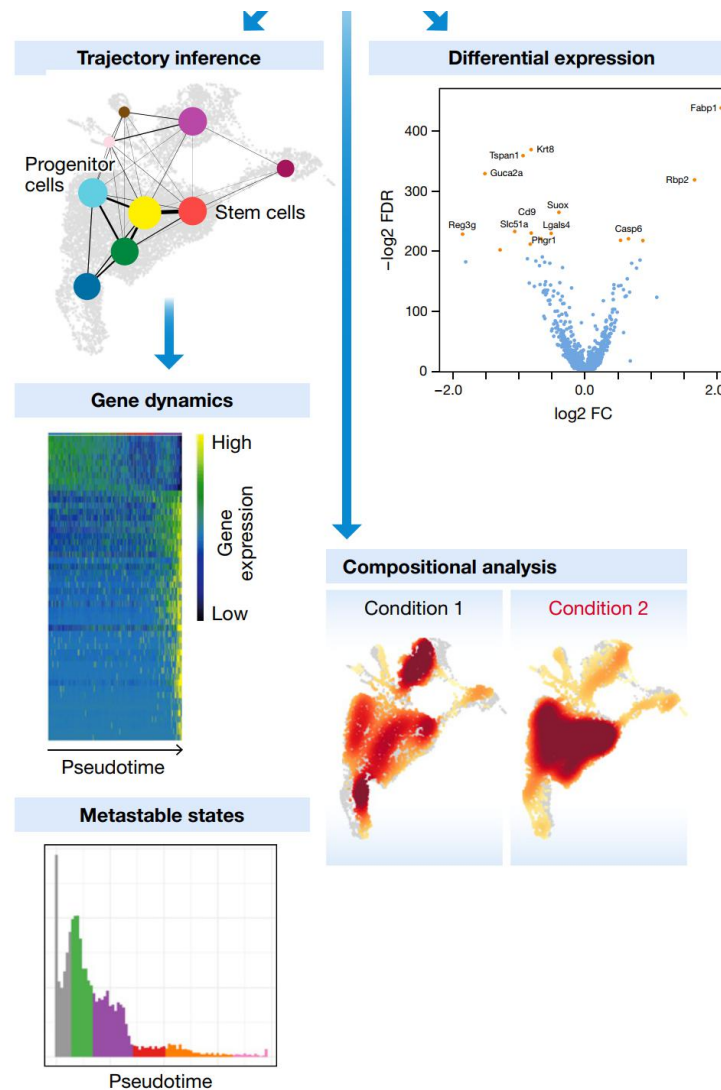
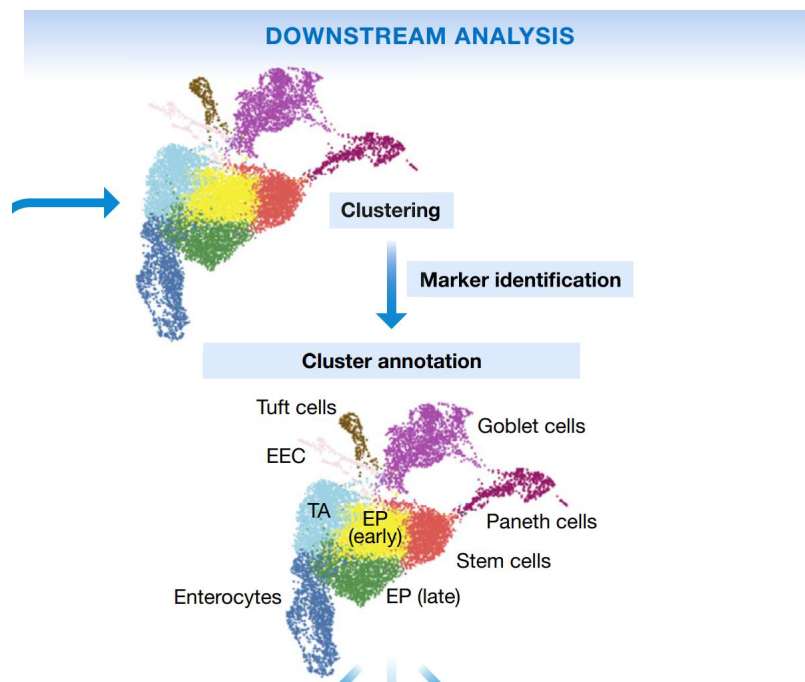
Feature selection



Visualization

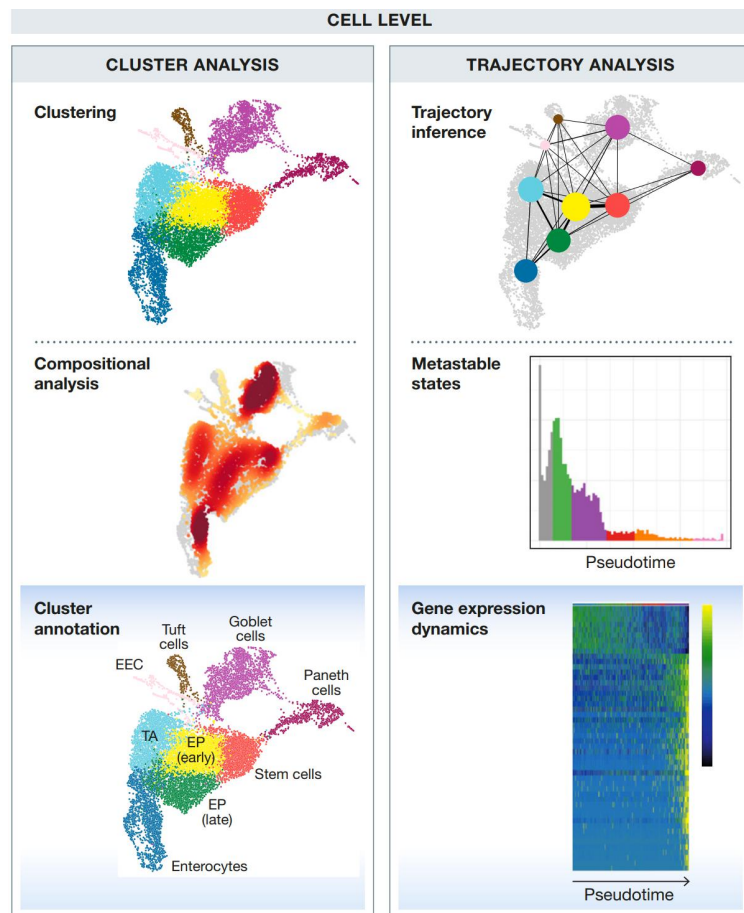


单细胞下游分析



下游分析：
聚类分析、
伪轨迹分析、
差异表达分析、
基因动力学、
亚稳态分析、
成分分析等
可分为 cell-level,
gene-level。

cell-level



cell level 集中描述两种结构：
聚类clusters，通过细胞在不同的类
中，解释数据的异质性
(heterogeneity)
轨迹trajectories: 动态过程中的一个
剪影。

单细胞技术

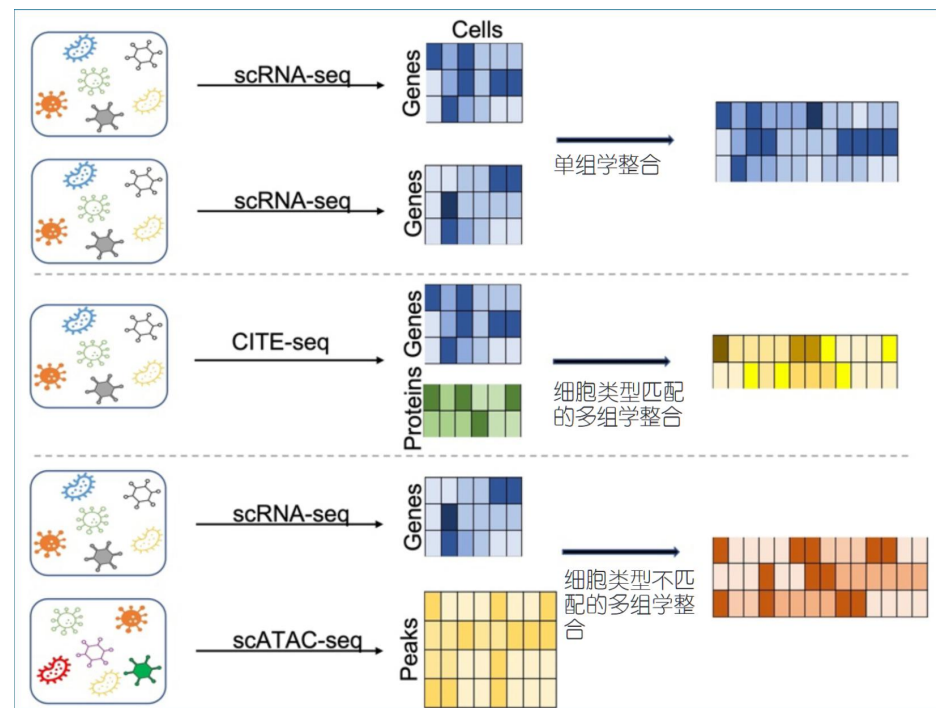
- 这里特别强调一下下游分析，该教程将下游分析主要分为细胞和基因水平。
 - 细胞水平的下游分析，研究对象主要是单细胞。围绕单细胞开展聚类分析、伪轨迹分析、成分分析、亚稳态分析、聚类注释和基因表达动力学。
 - 基因水平的下游分析，研究对象主要是基因。围绕基因开展差异表达分析、基因集分析和基因调控网络构建。对于细胞水平的下游分析，该教程忽略了细胞之间的通讯分析。另外，对于基因水平的下游分析，该教程忽略了基因模块分析。

细胞类型分析

- 用于预测细胞类型，识别复杂和罕见的细胞类型群，评估不同细胞系的发育轨迹，揭示在各种疾病和治疗环境中的细胞变异。
- 细胞类型的可解释性。改善单细胞水平生物学方面的研究和可解释性。

单细胞多组学分析

- 单组学数据整合：跨多个批次和技术生成的，由技术误差导致测量RNA表达出现偏差。也称，批次校正。目标去除批次之间的技术误差，同时保留批次内部和批次之间的生物差异。
 - 常见软件： Seurat, LIGER等。
- 多组学数据整合
 - 细胞类型匹配的
 - 细胞类型不匹配



细胞相互作用网络

- 挑战： 缺少互相作用网络的标准， 以及方法的评价框架。
 - 侧重配体-受体相互作用的下游通路分析
 - 没有用到空间转录组信息。
 - 没有对细胞内传达通路以及转录因子-靶基因为核心元件的细胞特异调控网络。

参考文献

- <https://www.nature.com/articles/s41596-020-00409-w>
- <https://pubmed.ncbi.nlm.nih.gov/33835452/>
- <https://www.embopress.org/doi/epdf/10.15252/msb.20188746>