

植物单细胞转录组测序研究进展

李益 孙超

(中国医学科学院北京协和医学院药用植物研究所, 北京 100193)

摘要: 单细胞转录组测序是一种在单细胞水平上研究基因表达的技术。多孔板法和液滴法是目前应用于植物研究的两类主要的单细胞转录组技术。首先概述了植物单细胞转录组测序的技术原理和数据分析流程, 然后介绍了植物单细胞转录组的研究进展, 重点阐述了单细胞转录组测序技术在鉴定植物细胞类型、揭示细胞演化轨迹和构建细胞间调控网络中的应用。单细胞转录组测序技术可为植物领域的研究提供新视角, 有助于识别和理解复杂组织中关键的细胞生物学过程。

关键词: 单细胞; 转录组测序; 液滴法; 多孔板法; 细胞类型特异性

DOI: 10.13560/j.cnki.biotech.bull.1985.2020-1011

Research Progress in Single-Cell RNA-Seq of Plant

LI Yi SUN Chao

(Institute of Medicinal Plant Development, China Academy of Medical Sciences & Peking Union Medical College, Beijing 100193)

Abstract: Single-cell RNA-Seq (scRNA-seq) is a technology to analyze gene expression at single cell level. Currently PCR plate-based and droplet-based protocols are 2 major popular ones used in the studies of plant. Here, we firstly outline the technical principles and data analysis procedure of plant scRNA-seq, and then introduce its research status, focusing on how scRNA-seq can be used to detect distinct plant cell types, to reveal cell evolution trajectories and to construct inter-cell regulatory networks. ScRNA-seq may provide a new perspective on plant research, and help to understand and identify critical cellular biological processes in complex tissues.

Key words: single cell; RNA-sequencing; droplet-based protocol; PCR plate-based protocol; cell-type-specific

生物体由不同形态、具有特定功能的细胞构成, 而不同细胞的基因表达模式也是不同的^[1]。常规的植物转录组学研究通常是将植物整个器官或组织均质化后测序, 忽略了细胞的异质性^[2], 虽然有助于在器官或组织水平上解决许多生物学问题, 但无法了解稀有细胞类型或单个细胞的转录过程。细胞捕获、测序技术和生物信息学的飞速发展使得单细胞转录组测序 (Single-cell RNA-seq, scRNA-seq) 成为可能, 并不断发展和完善, 目前已广泛应用于生物学和医学领域, 并在植物学研究中显示出巨大的潜力。在植物中开展单细胞转录组研究有助于深入理解不同细胞类型在发育过程中的作用以及细胞间的调控网络^[3]。本文对植物单细胞转录组测序和数据

分析以及单细胞转录组测序在植物研究中的应用进行了概述。

1 植物单细胞转录组测序技术

scRNA-seq 技术发展迅猛, 从最初只能检测几个细胞到单次实验可以同时检测数 10 万个细胞, 实验效率得到了显著提高^[4]。同时, 建库及测序过程中不同环节的改进使得成本不断降低、有效信息量不断增加。目前已经开发出多种 scRNA-seq 技术, 不同技术的适用范围不同。多孔板法 (PCR plate-based) 和液滴法 (Droplet-based) 是两类成功应用于植物研究的 scRNA-seq 技术 (表 1)。两类方法各有特点, 可根据实验目的和植物样本性质选择合适的技术方法^[5]。

收稿日期: 2020-08-12

基金项目: 中国医学科学院医学与健康科技创新工程 (2016-I2M-2-003)

作者简介: 李益, 女, 硕士研究生, 研究方向: 药用植物功能基因组学; E-mail: norland1126@163.com

通讯作者: 孙超, 男, 博士, 教授, 研究方向: 药用植物功能基因组学和天然产物合成生物学; E-mail: csun@implad.ac.cn

表 1 单细胞转录分析技术

建库方法	分析技术	捕获细胞数目	扩增技术	转录本覆盖率	参考文献
多孔板法	Smartseq2	94-384/plate	PCR	全长	[10]
	CEL-seq2	94-384/plate	IVT	3' 端	[11]
液滴法	Chromium	80 000+	PCR	3' 端	[13]
	Drop-seq	50 000+	PCR	3' 端	[14]
	inDrop	40 000+	PCR	3' 端	[15]

1.1 多孔板法

对于稀有类型细胞或者细胞量较少的样本，可以考虑多孔板法。因此，基于多孔板法的 scRNA-seq 技术适用于研究特定类型或稀有的细胞，如生殖细胞。通常结合毛细管口吸法、激光显微切割或流式细胞荧光分选技术来分选单个细胞^[6]。这种方法首先需要将少量的细胞分选到含有 PCR 引物的 64/96 孔板中，然后对单个细胞进行独立构建测序文库并测序^[7]。基于多孔板的方法最大的优点是捕获效率高，但由于其细胞通量较低，测序成本高，限制了其大规模应用。目前在植物中成功开展应用的有 Smartseq2^[8] 和 CEL-seq2^[9]。Smart-seq2 支持全长转录本测序，灵敏度高，是检测低表达的转录本的最佳选择^[10]。CEL-seq2 采用体外转录线性扩增 (IVT) 的建库方法，主要优势是减少了 PCR 指数扩增所造成的偏差，扩增后 DNA 的双端深度测序能够准确检测两条链的序列^[11]。

1.2 液滴法

基于液滴法的 scRNA-seq 通过微流控芯片，利用液滴直接分选单个细胞，可以无差别获得组织上千个细胞，使得植物单细胞转录组研究实现了从少量细胞到高通量的飞跃^[12-13]。2015 年，哈佛大学的两个团队将微流控技术引入 scRNA-seq 中，分别开发出 Drop-seq^[14] 和 inDro^[15] 两种技术。随后，10× Genomics 公司于 2017 年推出一个基于液滴法的商业化单细胞分析系统 Chromium，使得 scRNA-seq 的应用得到了迅速发展。2019 年，先后有多篇文章报道利用 Chromium 平台对拟南芥根开展了研究，证明了高通量单细胞转录组测序同样可用于研究植物^[16-20]。基于液滴法的 scRNA-seq 都运用了相似的技术原理，在微流控设备中，水流中包含悬浮状态

下的细胞，裂解缓冲液中包含了用条码 (Barcodes) 标记的微珠，这两股流体汇集在一起后穿过油体通道，最终形成一个油滴包裹的凝胶珠。一旦液滴包裹成功，细胞立即被裂解，释放出与微珠表面引物结合的 RNA，在微珠表面反转录成 cDNA，生成包含成千上万个单细胞的 cDNA 文库^[21-22]。

2 植物单细胞转录组分析流程

相比于传统转录组测序，scRNA-seq 产生的数据量更为庞大复杂，分析和解释数据也是 scRNA-seq 分析中的重点^[23]。scRNA-seq 分析的具体步骤可能会由于生物学问题不同而有所不同，但大多数分析中使用的核心流程是一致的，包括数据预处理、数据的降维和聚类以及数据下游分析 3 部分^[24-26] (图 1)。

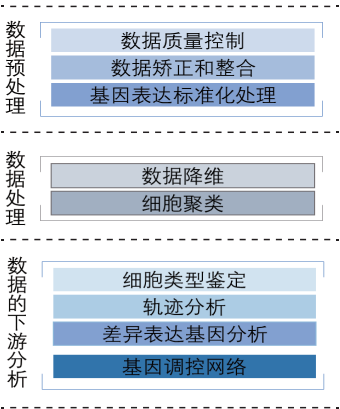


图 1 单细胞转录组数据分析流程

2.1 数据预处理

数据的预处理包括数据的质控、数据矫正和整合以及对数据的标准化处理 3 部分^[27-30]。细胞异常破裂、死亡、捕获细胞的位置没有细胞或者含有多个细胞会导致产生低质量数据，因此在正式分析前需要对原始数据进行质量控制以剔除这部分数据^[31]。数据整合可以消除实验过程中的生物因素、技术因素以及不同批次引起的数据偏差，尽可能地展示单个细胞的真实表达情况^[32-33]。通过多步过滤的数据即可用于构建高精度的基因-细胞表达矩阵，用于后续分析。

2.2 降维和聚类分析

scRNA-seq 数据具有高维性，涉及数千个基因

和大量细胞, 当在一个高维基因表达空间中比较细胞时, 细胞间的距离变得更加均匀, 使得区分群体间或者群体内的差异非常困难。首先在数千个细胞的基因表达量数据中, 选取其中高度可变的基因 (Highly variable genes, HVGs), 比使用所有的基因, 选择 HVGs 更为有效^[26, 34]。然后采取主成分分析 (Principal component analysis, PCA) 降低数据集的高维度和复杂度, PCA 可将数据投射到较少的独立的线性维度中, 从而捕捉到可能的最大方差。采用 t 分布随机领域嵌入 (t-Distributed stochastic neighbor embedding, tSNE) 或均匀流形近似和投影 (Uniform manifold approximation and projection, UMAP) 对细胞进行进一步降维, 这两种都是非线性降维方法, 可以有效地将高维度数据转换成二维图像^[35-36]。PCA 降维后的数据传递到 t-SNE 与 UMAP 进行二维可视化展示, 细胞之间的基因表达模式越相似, 在 t-SNE/UMAP 图中的距离也越接近。接下来可采用 k-means 算法或图聚类算法 (Graph-based) 进行聚类分析, 将表达相似的细胞聚在一起, 形成不同的细胞亚群^[24, 37] (Cell cluster)。Seurat 是用于分析单细胞数据的常用软件, 它使用基于图聚类的算法, 通过计算细胞间的差异性, 优化细胞间聚类关系距离的权重值 (通过设定软件中的阈值), 实现对细胞的聚类。

2.3 数据下游分析

数据下游分析包含细胞水平和基因水平的分析^[25]。细胞水平的分析又分为细胞类型鉴定和轨迹分析, 其中也涉及到基因水平的分析。基因水平分析包括差异表达分析、基因集分析和基因调控网络分析。目前, 主要有两种方法用于鉴定细胞类型, 一种是人工鉴定方法, 综合利用样本信息、组织类型、细胞状态、表面 marker 和差异表达基因, 并结合已知数据库的信息, 进行细胞类型注释^[38]。CellMarker^[39] 和 panglaoDB^[40] 是两个常用数据库, 提供了人和小鼠细胞注释集。对于其他物种, 则需要根据报道的文献来确定标记基因 (marker 基因)。另一种方法是利用自动化鉴定工具对细胞进行注释, 目前, 已经开发出近 30 种自动化鉴定工具, 包括 Scamp^[41]、SingleR^[42]、cellassign^[43] 等。

自动化工具利用已知类型的细胞样本的基因表达谱以及 marker 基因作为参考数据集, 基于单细胞与参考数据集表达谱的相似性, 对细胞类型进行自动化注释。细胞聚类、注释、重新聚类或子聚类以及重新注释过程的反复迭代非常耗时, 自动化鉴定方法提高了细胞注释的效率, 但也降低了其准确性。对于较小的数据集, 可以优先考虑人工注释的方法, 随着单细胞转录组测序样本数和细胞数的增加, 可以结合多种方法, 如首先使用自动化工具进行粗略注释, 然后利用人工注释对结果进行补充完善。在分选单细胞的过程中可以捕获到处于中间状态的细胞 (从一种状态到另一种状态的细胞), scRNA-seq 提供了一个很好的机会来组装发育过程中的演化轨迹。在细胞的演化进程中, 细胞的转变可能表现出不同的速率, 意味着不应随着时间来评估基因表达的变化, 而是应该依赖于发育过程中的进展。拟时 (Pseudotime) 分析, 又称细胞轨迹 (Cell trajectory) 分析, 根据测序细胞之间表达模式的相似性对单细胞沿着轨迹进行排序, 以此推断出发育过程细胞的分化轨迹或细胞亚型的演化过程^[44]。Saelens 等^[45] 对 45 种轨迹推断方法的准确性、可扩展性、稳定性和可用性 4 个方面进行了比较, 评估结果发现当前轨迹推断方法之间存在很大的互补性, 不同的工具有不同的使用范围。Monocle 是一款常用的拟时分析软件, 其计算细胞的相关性得到最小生成树, 找到最小路径, 然后把其他的所有数据点投射到最小路径, 最终得到细胞分化轨迹图的算法^[46-47]。

基因水平上的分析主要是通过比较细胞亚型之间差异基因的表达和功能富集, 从而进一步解释细胞的异质性^[48]。差异基因分析实际上是贯穿单细胞研究的重点分析内容, 亚群特征基因分析、处理组之间的基因动态变化、分化路径上的基因动态变化, 本质上都是差异基因分析。目前的基因差异表达分析软件有各自的优缺点, Wang 等^[49] 对比了 11 种基因差异表达分析软件发现, 传统的基因差异分析工具 (DESeq2, edgeR) 与单细胞差异分析工具性能表现相当, 尤其在检测灵敏度上表现良好, 但此类软件的运行时间较长, 对于大数据量的单细胞转录组的基因差异分析来说, 算法的运行时间通常是一种重要的考虑因素。在单细胞差异表达分析工具中,

DEsingle^[50]和SigEMD^[51]可以同时保证检测灵敏性和准确性,但是运行效率仍然比较低。另外,MAST (Model-based Analysis of Single-cell Transcriptomics) 软件利用 hurdle 模型消除 dropout (基因在某些细胞完全没有表达,同时在另外一些细胞有高表达的现象)的影响,在性能和效率上可以达到较好的平衡^[25, 52]。

3 单细胞转录组测序在植物研究中的应用

基于多孔板法的 scRNA-seq 可用于研究稀有细

胞,Efroni 等^[8]和 Nelms 等^[9]分别利用 Smart-seq2 和 Cell-seq2 捕捉到了愈伤组织和生殖细胞在进入分化阶段前的瞬时变化。基于液滴法的 scRNA-seq,尤其是 Chromium 平台的高细胞通量为植物单细胞转录组研究带来了新的突破口,使得研究样本从少量细胞向组织器官转变。总的来说,scRNA-seq 可以通过捕获单个细胞的基因的表达情况(表 2),来揭示细胞的异质性,细胞的分化轨迹以及细胞对环境变化的响应机制。

表 2 植物单细胞转录组研究概况

建库方法	分析技术	物种	样本类型	细胞数	总基因数(转录本数)	基因数(转录本数)/细胞	参考文献
多孔板法	Smart-seq2	拟南芥	根	238	—	—	[8]
	Cell-seq2	玉米	生殖细胞	144	101 245*	—	[9]
	Cell-seq2	拟南芥	静止中心细胞	24	—	14 000	[53]
			中柱细胞	7	—	4 312	[53]
液滴法	Chromium	拟南芥	根(3个生物学重复)	7 522	>22 000	5 000	[16]
	Chromium	拟南芥	根(2个生物学重复)	4 727	16 975	4 276	[17]
	Chromium	拟南芥	根	3 121	22 419	2 445	[18]
			根(热胁迫)	1 009	21 237	1 009	[18]
			根	1 079	22 971	4 079	[18]
			根	7 695	23 161	—	[19]
	Chromium	拟南芥	子叶	12 844	—	—	[20]
	Drop-seq	拟南芥	根	6 102	—	>1 000*	[54]
			根(1%蔗糖处理)	6 096	—	—	[54]
	Drop-seq	拟南芥	根	374	—	—	[55]

注：* 为转录本数

3.1 鉴定细胞类型

2019 年 2 月,首篇利用高通量单细胞测序的植物根尖单细胞图谱文章发表在 *Plant Physiology* 上。Ryu 等^[16]选择拟南芥幼苗根尖组织为样本,利用 Chromium 平台,共获得了 7 552 个细胞的转录组数据。通过 Seurat 对这些细胞进行降维聚类分析,得到 9 个主要的细胞亚群,随后利用 86 个已知特异性表达的标记基因集对不同的细胞亚群进行注释,同样地,利用木质部和韧皮部标记基因集区分了中柱细胞内的不同细胞亚型,证实了高通量 scRNA-seq 在植物研究中的可行性和有效性。Denyer 等^[17]选用了相同的测序平台对拟南芥根组织进行了测序,利用相似的有监督分类方法注释细胞类型,并构建了含报告基因的转基因拟南芥株系,结果显示内皮层组织内的细胞确实有报告基因绿色荧光蛋白的表

达。有监督分类方法仅适用于极少数有参考数据集的植物,作者还采用了无监督分类方法,通过定义聚簇中特异性的标记基因标准,在聚簇之间的差异基因集中进行筛选,获取了数百个自定义的标记基因,并从中挑选了 10 个特异性高且此前未报导过与根发育相关的标记基因,通过报告基因株系进行验证发现,有 8 个基因的表达模式与预测一致。有关拟南芥根组织的研究结果表明,利用 scRNA-seq 数据可以鉴别不同细胞类型,如中柱鞘细胞、韧皮部筛管和不同表皮细胞亚型,也能检测到静止中心(Quiescent centre, QC)这种数目稀少的细胞群。

3.2 揭示细胞分化轨迹

利用拟时序分析可以推导出具有分化/演化关系的细胞亚群间可能的分化路径。Shulse 等^[54]通过 Monocle 推断了内皮层细胞的发育过程,结果显示发

育早期的细胞亚群沿着两支轨迹曲线分化。与内皮层发育相关的 798 个基因在拟时间序列上呈现早期、中期、晚期 3 种表达模式,且调节内皮细胞分化初始阶段的相关基因在拟时间序列的早期阶段表达,而晚期表达的基因主要与木质素代谢和细胞连接组分合成相关。对多项拟南芥单细胞的研究,同样利用 Monocle 解析了根组织中的表皮^[16]、内皮层^[54]、根毛^[56]、分生组织^[17]、根冠^[19]细胞分化轨迹和各类细胞分化过程中基因的动态变化。Nelms^[9]收集了玉米的 144 个生殖细胞,通过单细胞分析重塑了玉米雄性细胞进入减数分裂的发育过程。在减数分裂前期,转录组图谱发生了两次急剧变化,通过比较转录组图谱与染色体细胞形态学的关系发现,第一个转录水平转变发生在第一次减数分裂前期的细线期,第二个转变发生在偶线期,表明减数分裂期间转录表达的改变不仅与核事件相关,而且与细胞形态相关。通过拟时序分析还能找出驱动细胞亚群分化的关键基因。分生组织到根毛细胞的发育过程中的基因表达谱显示,与根毛发育相关的细胞扩张和细胞重组的基因在拟时间序列的中期表达,以此推测这些基因可能是驱动根毛分化的特定基因。Turco 等^[55]研究木质部细胞分化到终端分化的分子机制,基于全根表达谱数据和单细胞数据,表明了 VND7 是启动根细胞向木质部细胞急剧转换的关键因子,确定了 4 个候选 VND7 下游靶基因。

3.3 分析细胞间调控网络

高通量 scRNA-seq 技术的可提供单细胞分辨率的转录组信息,有助于发现新的发育调控因子。为了进一步研究在根毛细胞发育过程中基因的相互作用,Denyer 等^[17]调取了单细胞数据中 239 个转录因子的动态表达数据,构建了精细的基因调控网络 (Gene regulatory network, GRN), GRN 分析结果显示了参与根毛发育过程中的关键因子以及相互作用关系,进一步将 GRN 的转录因子过滤至 25 个核心组分,发现了一系列负反馈调控的转录因子。Liu 等^[20]采用 scRNA-seq 技术解析了拟南芥气孔谱系细胞发育进程中的转录组动态模式,对气孔谱系细胞早期发育阶段中转录因子进行了筛选,结果显示调控植物幼苗生长、发育和应激响应的重要转录

因子显著高表达,转录因子的调控网络显示 BPC、WRKY33 作为核心转录因子不仅参与调控功能基因,还与其他转录因子相互作用。此外, Jean-Baptiste 等^[18]对幼苗进行热胁迫处理,分析了单细胞水平热激响应转录调控网络发现,响应高温的热激基因在不同细胞类型中都有表达,但仍有一些基因在不同细胞类型中存在显著的差异表达,植物对内外源信号的响应的细胞异质性应该广泛存在。

4 挑战与前景

植物单细胞转录组研究的最大技术挑战是将细胞从适当的组织中分离出来,并获得大量的细胞用于高通量分析。植物细胞有细胞壁的保护,必须先制备成原生质体才能制备单细胞悬液。目前尚未开发出可以适用于任何植物的通用的制备原生质体的方法,这也是植物单细胞研究样本单一的原因。因此在制备植物单细胞悬液过程中,可根据植物组织的特性,优化酶解条件以分离原生质体。单细胞测序产生的背景噪声数据和不同样本间产生的批次效应是处理单细胞数据时的难点。为此,多种生物信息学工具已被开发并成功应用于 scRNA-seq 分析。多篇植物单细胞转录组测序文章证实了高通量 scRNA-seq 方法的在植物研究中的可行性和有效性,预示着植物研究进入了单细胞时代。未来植物单细胞技术发展的主要趋势是提高植物单细胞分离效率,实现多样本多组织研究。近期,有科学家提出了植物细胞图谱计划 (Plant Cell Atlas)^[57],高通量 scRNA-seq 技术是其不可或缺的重要一环。可以预见单细胞 ChIP-seq、单细胞 ATAC-seq、单细胞 Hi-C 等单细胞测序技术也会加入植物单细胞研究的队列,从而使高精度研究单细胞基因调控模型成为可能。

参考文献

- [1] Libault M, Pingault L, et al. Plant systems biology at the single-cell level [J]. *Trends Plant Sci*, 2017, 22 (11): 949-960.
- [2] 文路, 汤富酬. 单细胞转录组高通量测序分析新进展 [J]. *遗传*, 2014, 36 (11): 1069-1076.
Wen L, Tang FC. Recent progress in single-cell RNA-Seq analysis [J]. *Hereditas*, 2014, 36 (11): 1069-1076.
- [3] Efroni I, Birnbaum KD. The potential of single-cell profiling in

- plants [J]. *Genome Biology*, 2016, 17 (1): 65.
- [4] Kolodziejczyk AA, Kim JK, Svensson V, et al. The technology and biology of single-cell RNA sequencing [J]. *Molecular Cell*, 2015, 58 (4): 610-620.
- [5] Baran-Gale J, Chandra T, Kirschner K. Experimental design for single-cell RNA sequencing [J]. *Briefings in Functional Genomics*, 2017, 17 (4): 233-239.
- [6] Jaitin DA, Kenigsberg E, Keren-Shaul H, et al. Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types [J]. *Science*, 2014, 343 (6172): 776-779.
- [7] Andrews TS, Hemberg M. Identifying cell populations with scRNASeq [J]. *Mol Aspects Med*, 2018, 59: 114-122.
- [8] Efroni I, Mello A, Nawy T, et al. Root regeneration triggers an embryo-like sequence guided by hormonal interactions [J]. *Cell*, 2016, 165 (7): 1721-1733.
- [9] Nelms B, Walbot V. Defining the developmental program leading to meiosis in maize [J]. *Science*, 2019, 364 (6435): 52-56.
- [10] Picelli S, Faridani OR, et al. Full-length RNA-seq from single cells using Smart-seq2 [J]. *Nature Protocols*, 2014, 9 (1): 171-181.
- [11] Hashimshony T, Senderovich N, Avital G, et al. CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq [J]. *Genome Biology*, 2016, 17 (1): 77.
- [12] Zhang X, Li T, Liu F, et al. Comparative analysis of droplet-based ultra-high-throughput single-cell RNA-seq systems [J]. *Molecular Cell*, 2019, 73 (1): 130-142.
- [13] Zheng GXY, Terry JM, Belgrader P, et al. Massively parallel digital transcriptional profiling of single cells [J]. *Nature Communications*, 2017, 8 (1): 14049.
- [14] Macosko EZ, Basu A, Satija R, et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets [J]. *Cell*, 2015, 161 (5): 1202-1214.
- [15] Klein Allon M, Mazutis L, Akartuna I, et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells [J]. *Cell*, 2015, 161 (5): 1187-1201.
- [16] Ryu KH, Huang L, Kang HM, et al. Single-cell RNA sequencing resolves molecular relationships among individual plant cells [J]. *Plant Physiology*, 2019, 179 (4): 1444-1456.
- [17] Denyer T, Ma X, Klesen S, et al. Spatiotemporal developmental trajectories in the *Arabidopsis* root revealed using high-throughput single-cell RNA sequencing [J]. *Developmental Cell*, 2019, 48 (6): 840-852.
- [18] Jean-Baptiste K, McFaline-Figueroa JL, Alexandre CM, et al. Dynamics of gene expression in single root cells of *Arabidopsis thaliana* [J]. *The Plant Cell*, 2019, 31 (5): 993-1011.
- [19] Zhang TQ, Xu ZG, Shang GD, et al. A single-cell RNA sequencing profiles the developmental landscape of *Arabidopsis* root [J]. *Mol Plant*, 2019, 12 (5): 648-660.
- [20] Liu Z, Zhou Y, Guo J, et al. Global dynamic molecular profiling of stomatal lineage cell development by single-cell RNA sequencing [J]. *Mol Plant*, 2020, 13 (8): 1178-1193.
- [21] Prakadan SM, Shalek AK, Weitz DA. Scaling by shrinking: empowering single-cell 'omics' with microfluidic devices [J]. *Nature Reviews: Genetics*, 2017, 18 (6): 345-361.
- [22] Rich-Griffin C, Stechemesser A, Finch J, et al. Single-cell transcriptomics: A high-resolution avenue for plant functional genomics [J]. *Trends Plant Sci*, 2020, 25 (2): 186-197.
- [23] Dal Molin A, Di Camillo B. How to design a single-cell RNA-sequencing experiment: pitfalls, challenges and perspectives [J]. *Brief Bioinform*, 2018, 20 (4): 1384-1394.
- [24] Wu Y, Zhang K. Tools for the analysis of high-dimensional single-cell RNA sequencing data [J]. *Nat Rev Nephrol*, 2020, 16 (7): 408-421.
- [25] Luecken MD, Theis FJ. Current best practices in single-cell RNA-seq analysis: a tutorial [J]. *Molecular Systems Biology*, 2019, 15 (6): e8746.
- [26] Chen G, Ning B, Shi T. Single-cell RNA-seq technologies and related computational data analysis [J]. *Front Genet*, 2019, 10: 317.
- [27] AlJanahi AA, Danielsen M, Dunbar CE. An Introduction to the analysis of single-cell RNA-sequencing data [J]. *Molecular Therapy Methods & Clinical Development*, 2018, 10: 189-196.
- [28] Weinreb C, Wolock S, Klein AM. SPRING: a kinetic interface for visualizing high dimensional single-cell expression data [J]. *Bioinformatics*, 2018, 34 (7): 1246-1248.
- [29] Ziegenhain C, Vieth B, et al. Quantitative single-cell transcriptomics [J]. *Brief Funct Genomics*, 2018, 17 (4): 220-232.
- [30] Ding B, Zheng L, Zhu Y, et al. Normalization and noise reduction for single cell RNA-seq experiments [J]. *Bioinformatics*, 2015, 31 (13): 2225-2227.
- [31] van den Brink SC, Sage F, Vértessy Á, et al. Single-cell sequencing

- reveals dissociation-induced gene expression in tissue subpopulations [J]. *Nature Methods*, 2017, 14 (10): 935-936.
- [32] Butler A, Hoffman P, Smibert P, et al. Integrating single-cell transcriptomic data across different conditions, technologies, and species [J]. *Nature Biotechnology*, 2018, 36 (5): 411-420.
- [33] Tran HTN, Ang KS, Chevrier M, et al. A benchmark of batch-effect correction methods for single-cell RNA sequencing data [J]. *Genome Biology*, 2020, 21 (1): 12.
- [34] Brennecke P, Anders S, Kim JK, et al. Accounting for technical noise in single-cell RNA-seq experiments [J]. *Nature Methods*, 2013, 10 (11): 1093-1095.
- [35] van der Mauten L, Hinton G. Visualizing data using t-SNE [J]. *J Mach Learn Res*, 2008, 9 (2605): 2579-2605.
- [36] Diaz-Papkovich A, Anderson-Trocmé L, et al. UMAP reveals cryptic population structure and phenotype heterogeneity in large genomic cohorts [J]. *PLoS Genetics*, 2019, 15 (11): e1008432.
- [37] Duò A, Robinson MD, Soneson C. A systematic performance evaluation of clustering methods for single-cell RNA-seq data [J]. *F1000Research*, 2018, 7 : 1141.
- [38] Abdelaal T, Michielsen L, Cats D, et al. A comparison of automatic cell identification methods for single-cell RNA sequencing data [J]. *Genome Biology*, 2019, 20 (1): 194.
- [39] Zhang X, Lan Y, Xu J, et al. CellMarker : a manually curated resource of cell markers in human and mouse [J]. *Nucleic Acids Research*, 2019, 47 (D1): D721-D728.
- [40] Franzén O, Gan LM, Björkegren JLM. PanglaoDB : a web server for exploration of mouse and human single-cell RNA sequencing data [J]. *Database (Oxford)*, 2019 (2019): baz046.
- [41] Kiselev VY, Yiu A, Hemberg M. scmap : projection of single-cell RNA-seq data across data sets [J]. *Nature Methods*, 2018, 15 (5): 359-362.
- [42] Aran D, Looney AP, 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.
- [43] Zhang AW, O' Flanagan C, Chavez EA, et al. Probabilistic cell-type assignment of single-cell RNA-seq for tumor microenvironment profiling [J]. *Nature Methods*, 2019, 16 (10): 1007-1015.
- [44] Haghverdi L, Büttner M, Wolf FA, et al. Diffusion pseudotime robustly reconstructs lineage branching [J]. *Nature Methods*, 2016, 13 (10): 845-848.
- [45] Saelens W, Cannoodt R, Todorov H, et al. A comparison of single-cell trajectory inference methods [J]. *Nature Biotechnology*, 2019, 37 (5): 547-554.
- [46] Trapnell C, Cacchiarelli D, Grimsby J, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells [J]. *Nat Biotech*, 2014, 32 (4): 381-386.
- [47] Qiu X, Hill A, Packer J, et al. Single-cell mRNA quantification and differential analysis with Censur [J]. *Nature Methods*, 2017, 14 (3): 309-315.
- [48] Alessandri L, Arigoni M, Calogero R, et al. Differential expression analysis in single-cell transcriptomics [J]. *Methods in Molecular Biology*, 2019, 1979 : 425-432.
- [49] Wang T, Li B, Nelson CE, et al. Comparative analysis of differential gene expression analysis tools for single-cell RNA sequencing data [J]. *BMC Bioinformatics*, 2019, 20 (1): 40.
- [50] Miao Z, Deng K, Wang X, et al. DESingle for detecting three types of differential expression in single-cell RNA-seq data [J]. *Bioinformatics*, 2018, 34 (18): 3223-3224.
- [51] Wang T, Nabavi S. SigEMD : A powerful method for differential gene expression analysis in single-cell RNA sequencing data [J]. *Methods*, 2018, 145 : 25-32.
- [52] Finak G, McDavid A, Yajima M, et al. MAST : a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data [J]. *Genome Biology*, 2015, 16 (1): 278.
- [53] Efroni I, Ip PL, Nawy T, et al. Quantification of cell identity from single-cell gene expression profiles [J]. *Genome Biology*, 2015, 16 (1): 9.
- [54] Shulze CN, Cole BJ, Ciobanu D, et al. High-throughput single-cell transcriptome profiling of plant cell types [J]. *Cell Rep*, 2019, 27 (7): 2241-2247.
- [55] Turco GM, Rodriguez-Medina J, Siebert S, et al. Molecular mechanisms driving switch behavior in xylem cell differentiation [J]. *Cell Reports*, 2019, 28 (2): 342-351.
- [56] Yi F, Gu W, Chen J, et al. High temporal-resolution transcriptome landscape of early maize seed development [J]. *Plant Cell*, 2019, 31 (5): 974-992.
- [57] Rhee SY, Birnbaum KD, Ehrhardt DW. Towards building a plant cell atlas [J]. *Trends Plant Sci*, 2019, 24 (4): 303-310.