

网络出版时间:2022-09-06 13:30 DOI:10.13207/j.cnki.jnwafu.2023.03.005  
网络出版地址:https://kns.cnki.net/kcms/detail/61.1390.S.20220905.1754.002.html

# 单细胞转录组测序技术发展及其在甘薯中的应用

赵楠<sup>1,2</sup>,梁雪莲<sup>1</sup>,谢淑燕<sup>2</sup>,张新新<sup>2</sup>,马婷婷<sup>2,3</sup>,邹宏达<sup>2</sup>,  
陈景益<sup>2</sup>,房伯平<sup>2</sup>,黄立飞<sup>2</sup>

(1 仲恺农业工程学院 农业与生物学院,广东 广州 510000;2 广东省农业科学院作物研究所,  
广东省农作物遗传改良重点实验室,广东 广州 510000;3 华南农业大学 植物保护学院,广东 广州 510000)

**[摘 要]** 单细胞转录组测序(single-cell RNA sequencing, scRNA-seq)技术是在单个细胞水平通过对 mRNA 进行高通量测序研究其整体水平基因表达情况的一项新技术。本研究概述了 scRNA-seq 技术的发展历程和创新;详细介绍了 scRNA-seq 技术的单细胞捕获/分选、反转录/PCR 扩增、建库测序和测序数据基础分析等步骤,以及在植物单细胞图谱、非生物胁迫响应机制、细胞分化发育、转录因子调控网络等研究中的应用;剖析了该技术在甘薯中单细胞图谱、分化、抗逆和基因方面的应用潜力;最后对该技术的应用进行了展望。

**[关键词]** 单细胞转录组测序;甘薯育种;细胞异质性

**[中图分类号]** Q943.2

**[文献标志码]** A

**[文章编号]** 1671-9387(2023)03-0039-12

## Development of single-cell RNA-seq technology and its application in sweet potato

ZHAO Nan<sup>1,2</sup>, LIANG Xuelian<sup>1</sup>, XIE Shuyan<sup>2</sup>, ZHANG Xinxin<sup>2</sup>, MA Tingting<sup>2,3</sup>,  
ZOU Hongda<sup>2</sup>, CHEN Jingyi<sup>2</sup>, FANG Boping<sup>2</sup>, HUANG Lifei<sup>2</sup>

(1 College of Agriculture and Biology, Zhongkai University of Agriculture and Engineering, Guangzhou, Guangdong 510000, China;  
2 Guangdong Academy of Agricultural Sciences Institute of Crops, Guangdong Key Laboratory of Crops Genetics and Improvement,  
Guangzhou, Guangdong 510000, China; 3 College of Plant Protection, South China Agricultural University,  
Guangzhou, Guangdong 510000, China)

**Abstract:** Single-cell RNA-seq(scRNA-seq) technology is new to study overall gene expression at the single cell level by high-throughput sequencing of mRNA. This paper summarized the development and innovation of single cell transcriptome sequencing technology. The steps of single cell capture or sorting, reverse transcription/PCR amplification, database construction and sequencing data basic analysis of scRNA-seq technology were introduced in detail, and its application in the study of scRNA-seq technology in plant single cell map, abiotic stress response mechanism, cell differentiation and development, and transcription factor regulatory network were introduced. The application potential of single cell map, differentiation, stress resistance and gene in sweet potato was also analyzed. Finally, the application of this technology was prospected.

**Key words:** single-cell RNA sequencing (scRNA-seq); sweet potato breeding; cell heterogeneity

**[收稿日期]** 2022-02-06

**[基金项目]** 广东省重点领域研发计划项目(2020B020219001);国家甘薯产业技术体系(CARS-10)项目

**[作者简介]** 赵楠(1998—),女,山西长治人,硕士,主要从事甘薯抗病育种研究。E-mail:zhaonan202109@163.com

**[通信作者]** 黄立飞(1979—),男,山西左权人,副研究员,博士,主要从事甘薯抗病育种研究。E-mail:hlfl57@163.com

细胞是生物体的基本结构和功能单位,在其生长发育过程中,不同形态、具有特定功能的细胞转录组信息的变化呈多样性。传统转录组测序(bulk RNA sequencing, bulk RNA-seq)技术是在群体水平上对聚集的细胞进行分析,得出的平均数据往往反映的是数量占优势的细胞,缺失了低丰度信息,无法揭示细胞之间基因表达的异质性。单细胞转录组测序(single-cell RNA sequencing, scRNA-seq)技术是在单个细胞水平进行高通量测序,以高通量和单分子分辨率研究单个细胞表达谱,揭示细胞群体的异质性,避免单个细胞的基因表达信号被群体的平均值所掩盖。目前 scRNA-seq 技术已经在生物学和医学领域<sup>[1-2]</sup>进行了应用,在植物研究中也显示出巨大的应用潜力。甘薯(*Ipomoea batatas* L. Lam.)是我国重要的粮食和经济作物,作为无性繁殖的六倍体植物( $2n=6x=90$ ),它具有自交和同群杂交不亲和性,相比其他作物基因组相关研究难度较大,基因组数据匮乏,因此甘薯分子生物学发展受到限制。目前 scRNA-seq 技术已在植物领域进行了应用,取得了较大突破,但尚未见在甘薯中应用的报道,在甘薯中开展 scRNA-seq 有助于了解稀有细胞类型或单个细胞的转录过程以及细胞间的调控网络,系统跟踪单个细胞的动态变化,加深对细胞状态、转录本质以及基因表达调控的认识,揭示细胞间的调控网络,从而对抗病反应乃至免疫机制有进一步理解。本研究对 scRNA-seq 技术原理以及在植物研究中的应用进行概述,探讨了该技术在甘薯研究上的应用前景。

## 1 scRNA-seq 技术

### 1.1 scRNA-seq 技术的发展历程

2009 年, Tang 等<sup>[3]</sup>首次利用 scRNA-seq 技术分析小鼠单个生殖细胞早期发育阶段的特征,实现了在单个细胞中检测 mRNA 全基因组。此后, scRNA-seq 技术不断创新升级和优化,在技术水平和应用层面上都有进一步发展。

借助单细胞标记的逆转录测序以及全转录本扩增方法的建立和优化, scRNA-seq 技术逐渐成熟。Islam 等<sup>[4]</sup>和 Hashimshony 等<sup>[5]</sup>建立了单细胞标记的逆转录测序方法,即单细胞标记反转录测序(single-cell tagged reverse transcription sequencing, STRT-seq)技术和细胞表达经线性扩增和测序(cell expression by linear amplification and sequencing, CEL-seq)技术,该方法是在单独的管中挑取单个细

胞,其 mRNA 转化为 cDNA,利用模板转换机制,反转录同时引入条形码和上游引物结合序列,所有 cDNA 汇合后测序,包括片段化、接头连接和 PCR 扩增。2012 年, Ramsköld 等<sup>[6]</sup>开发了一项具有里程碑意义的 RNA 转录本 5' 端的开关机制(switching mechanism at 5' end of the RNA transcript, Smart-seq)技术,对全转录本进行扩增,实现了选择性转录本异构体和单核苷酸变异(single nucleotide variants, SNV)的检测。Picelli 等<sup>[7]</sup>在 Smart-seq 技术基础上使用锁核酸(locked nucleic acid, LNA)修饰的模板转换寡核苷酸(template-switching oligo, TSO)引物序列来促进模板转换,进而扩增出完整的 cDNA 序列,即 RNA 转录本 5' 端的开关机制 2 (switching mechanism at 5' end of the RNA transcript 2, Smart-seq2)技术;通过增加  $MgCl_2$  浓度,甜菜碱可在一定程度上提高 cDNA 产量;去除 cDNA 纯化步骤则可显著提高产率,但该技术只能对具有 poly(A)尾巴的 mRNA 进行扩增和检测。

随着微孔技术、微流体技术、微液滴和原位条形码技术的发展和完善, scRNA-seq 技术进入了高通量与自动化的全新时代。研究者开发出利用荧光激活细胞分选(fluorescence-activated cell sorting, FACS)分离单细胞到单独井中的大量平行的单细胞 RNA 测序(massively parallel single-cell RNA sequencing, MARS-seq)技术<sup>[8]</sup>、基因表达流式细胞测序(gene expression cytometry sequencing, Cyto-seq)的蜂窝板技术<sup>[9]</sup>,以及利用微流体装置将带有条形码的微珠和细胞一起装入微小液滴的 Fluidigm C1 技术平台(Fluidigm® C1™, Fluidigm C1)<sup>[10]</sup>、分析液滴中单个细胞的 mRNA 转录本测序(analysis of mRNA transcripts from individual cells in droplets, Drop-seq)技术<sup>[11]</sup>和带液滴的高通量单细胞标记测序(high-throughput single-cell labeling with indexing droplets, inDrop)技术<sup>[12]</sup>。其中 inDrop 技术捕获细胞的比例高于 Drop-seq 技术,故分析极少量的组织样本时采用 inDrop 技术可能更有优势。2017 年,  $10\times$  基因组学技术平台( $10\times$  Genomics chromium,  $10\times$  Genomics)基于携带唯一 DNA 标签编码(barcod)的微滴,通过微流控技术,将这些标签分配给单个细胞,实现了长片段 DNA 测序和单细胞分辨率的转录组测序。此外,还出现了对单细胞悬液进行不断分池,并加入标签,最终实现对每个细胞加入特异性标签的基于裂池连接的转录组测序(split pool ligation-based transcriptome sequen-

cing,SPLIT-seq)技术<sup>[13]</sup>。纵观这些年的单细胞测序方法的发展历程(图 1)可知,单细胞测序技术通

量越来越高,即在单位时间和一定费用内,可以获得更多的单细胞信息<sup>[14]</sup>。

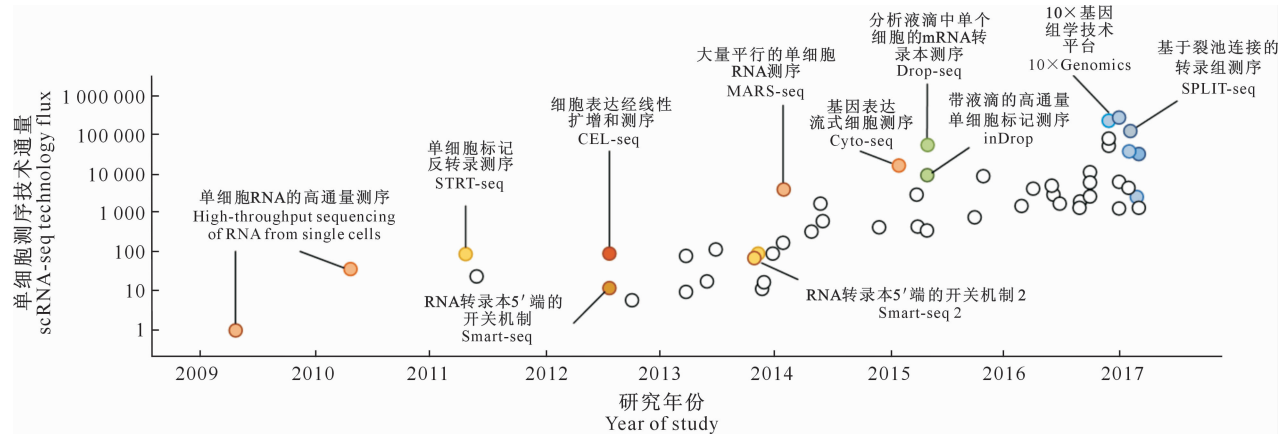


图 1 scRNA-seq 技术的发展历程<sup>[14]</sup>

Fig. 1 Development of single-cell RNA sequencing<sup>[14]</sup>

1.2 scRNA-seq 技术平台

目前研究者已经开发出了多种 scRNA-seq 技术平台(表 1),根据使用范围不同,大致可将其分为两大类:一类是基于全长序列的测序方法,如 Smart-seq 技术<sup>[6]</sup>和 Smart-seq2 技术<sup>[15]</sup>;另一类是基于独特的分子标识符(unique molecular identifiers,UMIs)的测序方法,包括细胞表达经线性扩增和测序 2 (cell expression by linear amplification

and sequencing 2,CEL-seq2)技术<sup>[16]</sup>、Drop-seq 技术<sup>[11]</sup>、MARS-seq 技术<sup>[8]</sup>和 10×Genomics 技术<sup>[17]</sup>等。

目前,10×Genomics 技术是 scRNA-seq 技术中的常用研究手段。单细胞转录组分析的主要步骤包括单细胞捕获/分选、反转录/PCR 扩增和建库测序,其中单细胞的捕获/分选是决定单细胞检测成本和通量的关键步骤。

表 1 不同 scRNA-seq 技术平台的比较

Table 1 Comparison of different scRNA-seq technology platforms

| 技术平台<br>Technology platform | 单细胞分选<br>Single cell sorting                    | 反转录/PCR 扩增<br>Reverse transcription/PCR amplification | 建库测序方法<br>Database construction for sequencing |
|-----------------------------|-------------------------------------------------|-------------------------------------------------------|------------------------------------------------|
| Smart-seq                   | 微流控 Microfluidic                                | 体外转录<br><i>In vitro</i> transcription                 | 全长 Full length                                 |
| Smart-seq2                  | 荧光激活细胞分选<br>Fluorescence-activated cell sorting | 聚合酶链式反应<br>Polymerase chain reaction                  | 全长 Full length                                 |
| CEL-seq2                    | 微流控 Microfluidic                                | 聚合酶链式反应<br>Polymerase chain reaction                  | 独特的分子标识符<br>Unique molecular identifiers       |
| MARS-seq                    | 荧光激活细胞分选<br>Fluorescence-activated cell sorting | 体外转录<br><i>In vitro</i> transcription                 | 独特的分子标识符<br>Unique molecular identifiers       |
| Drop-seq                    | 微液滴 Droplet                                     | 聚合酶链式反应<br>Polymerase chain reaction                  | 独特的分子标识符<br>Unique molecular identifiers       |
| 10×Genomics                 | 微液滴 Droplet                                     | 聚合酶链式反应<br>Polymerase chain reaction                  | 独特的分子标识符<br>Unique molecular identifiers       |

1.2.1 细胞捕获/分选技术 细胞捕获/分选方法主要包括特异性分选和非特异性分选两类<sup>[18]</sup>(表 2)。特异性分选是指从目标样本中,基于荧光标记物、细胞形态等特定标记对特定目标细胞进行筛选,并对目标细胞进行测序的方法。特异性选择的方法包括:(1)通量较低的微吸管分离法(micropipette separation technique,MST)<sup>[3]</sup>或激光捕获显微分离(laser capture microdissection,LCM)<sup>[19]</sup>,该类方法经常与通量较低的单细胞测序技术(如 Smart-seq

技术)配合,针对目标细胞进行扩增和测序。(2)通量较高的荧光激活细胞分选(fluorescence-activated cell sorting,FACS)或免疫磁珠分选(magnetic-activated cell sorting,MACS)<sup>[20]</sup>,常与通量较高的单细胞测序技术相配合,对大量目标细胞亚群进行检测(表 1)。但这 2 类特异性选择方法有很多局限,主要体现在细胞分选效率低、细胞分选依赖已知标记、成本高。

非特异性分选技术检测通量较高,其系统可以

自动对单细胞悬液中的每个细胞分别进行分离和处理,实现对系统中每个细胞进行平行检测。该单细胞悬液可以是来源于某个组织样本的所有细胞,也可以是 FACS、MACS 等预筛选的大量细胞。目前,成熟的高通量非特异性细胞分选技术可以分为微流控芯片、微孔、微液滴 3 类(表 2)。(1)Fluidigm C1 技术<sup>[10]</sup>是基于微流控的代表性技术,单细胞悬液导入芯片后,单个细胞被引入独立的腔室进行反转录和 cDNA 文库构建。Fluidigm C1 芯片共 96 个腔室,所以可以平行进行 96 个细胞的建库测序,但由于芯片格子密度的开发上限有限,故对于目标组织细胞图谱全景扫描研究,采用 Fluidigm C1 技术会大大提高芯片耗材投入成本。(2)BD Rhapsody (BD Rhapsody<sup>TM</sup>)技术是基于微孔的代表性技术,其技术原型为 Cyto-seq 技术<sup>[9]</sup>。该技术应用中单细胞悬液流过蜂窝板(微孔数>20 万,控制细胞数远小于微孔数),使细胞落入微孔中,再在蜂窝板上铺上微磁珠,最终在微孔中形成“磁珠-细胞-反应试剂”的体系,完成细胞裂解、反转录、扩增等过程,但

单次运行费用相对昂贵且耗时。(3)10×Genomics 技术是基于微液滴的代表性技术,其技术原型是 Drop-seq 技术<sup>[11]</sup>,该技术是利用 10×Genomics 芯片,使连接后续反应所需接头、barcode、UMI 等序列的磁珠在流动管道中经过 2 个“十字路口”,第 1 个“十字路口”的液体环境包含后续反应所需的酶,磁珠在该环境中吸附上细胞,然后在第 2 个“十字路口”形成油包水磁珠(gel bead in emulsion,GEM)的微液滴体系,每个单细胞在这些 GEM 中完成细胞裂解、反转录、扩增等步骤。10×Genomics 技术具有测序深度高、细胞通量高以及基于 UMIs 消除 PCR 测序偏差等特点。10×Genomics 芯片检测细胞数一般在 3 000~8 000,大大降低了投入资金,为较少投入开展大规模单细胞转录组的实验室研究提供了可能。Zhang 等<sup>[21]</sup>比较了 Drop-seq 技术、10×Genomics 技术和 inDrop 技术这 3 种基于液滴的高通量 scRNA-seq 技术平台,结果表明,10×Genomics 技术具有较高的灵敏度和精密度,其技术噪音也比 Drop-seq 技术和 inDrop 技术低。

表 2 不同细胞分选技术比较

Table 2 Comparison of different cell sorting technologies

| 指标<br>Index                                       | 特异性分选 Specific sorting                           |                                                  |                                                        |                                                  | 非特异性分选 Nonspecific sorting         |                                    |                                    |
|---------------------------------------------------|--------------------------------------------------|--------------------------------------------------|--------------------------------------------------------|--------------------------------------------------|------------------------------------|------------------------------------|------------------------------------|
|                                                   | 微吸管分离<br>Micropipette<br>separation<br>technique | 激光捕获<br>显微分离<br>Laser capture<br>microdissection | 荧光激活细胞分选<br>Fluorescence-<br>activated<br>cell sorting | 免疫磁珠分选<br>Magnetic-<br>activated<br>cell sorting | 微流控芯片<br>Microfluidics-            | 微孔<br>Microwell-                   | 微液滴<br>Droplet-                    |
| 样本要求<br>Sample<br>requirements                    | 任何组织<br>Any<br>organization                      | 任何组织<br>Any<br>organization                      | 单细胞悬液<br>Single cell<br>suspension                     | 单细胞悬液<br>Single cell<br>suspension               | 单细胞悬液<br>Single cell<br>suspension | 单细胞悬液<br>Single cell<br>suspension | 单细胞悬液<br>Single cell<br>suspension |
| 单次分选的细胞量<br>Amount of cells<br>sorted at one time | 少<br>Few                                         | 少<br>Few                                         | 数百至上千<br>Hundreds to<br>thousands                      | 不受限<br>Unlimited                                 | 数百<br>Hundreds                     | 数千<br>Thousands                    | 数千<br>Thousands                    |

1.2.2 反转录/PCR 扩增技术 获得单细胞之后,需要对单个细胞进行转录组扩增,构建 cDNA 文库以满足后续测序需求。目前单细胞 cDNA 文库构建多采用 oligo(dT)引物和 Superscript II/Ⅲ 逆转录酶,对带有 poly(A)尾的 RNA(主要 mRNA)进行反转录获得第一链 cDNA,然后以 Smart 方式使用 TSO 引物合成 cDNA 的第二条链。

对逆转录生成的 cDNA 进行全转录组扩增的方法主要包括:(1)PCR 方法(多重扩增方法)。这种扩增方法需要首先在 cDNA 两端加上锚定序列,然后利用锚定序列进行 PCR 扩增。根据合成第二条 cDNA 链方法的不同,PCR 扩增法分为末端加尾法、模板转换法和随机引物法。(2)基于体外转录(*in vitro* transcription,IVT)的方法(使用条形码的多重线性扩增)。体外转录扩增法从单细胞逆转录反应开始,使用 1条结构为 5'-T7 RNA 聚合酶启动

子序列-oligo(dT)-3'的引物逆转录为 cDNA,随后合成第二链 cDNA,在 T7 RNA 聚合酶作用下以第二链 cDNA 为模板进行体外转录<sup>[5]</sup>获得 cDNA。体外转录扩增法以线性扩增法代替 PCR 指数扩增,避免了 PCR 指数扩增带来的试验误差,但扩增效率远低于 PCR 扩增,需要进行 3 轮体外扩增,且扩增片段较短。

1.2.3 建库测序技术 利用 PCR 或 IVT 方法对 cDNA 扩增完成后,通过酶切方式片段化扩增产物,片段筛选后加测序接头 P5 和 P7,构建 cDNA 文库并进行纯化测序。目前 scRNA-seq 平台的测序方法大致可将其分为两类:(1)基于全长序列测序,如 Smart-seq 技术和 Smart-seq2 技术;(2)基于独特的分子标识符(UMIs)测序,包括 CEL-seq2 技术、Drop-seq 技术、MARS-seq 技术和 10×Genomics 技术等(表 1)。

## 2 植物单细胞转录组基础分析

与传统 bulk RNA-seq 技术相比,scRNA-seq 技术获得的数据结构相同,数据量大且存在较高的技术噪音,会对后续分析造成影响,所以需要

对数据进行评估和调整误差,以减少技术或生物来源带来的噪声。scRNA-seq 技术的基础分析步骤主要包括数据预处理、降维和聚类分群、亚群标记基因筛选和可视化、细胞亚群鉴定等,所有植物通用,随后根据不同的植物特点以及研究方向进行定制化分析(图 2)。

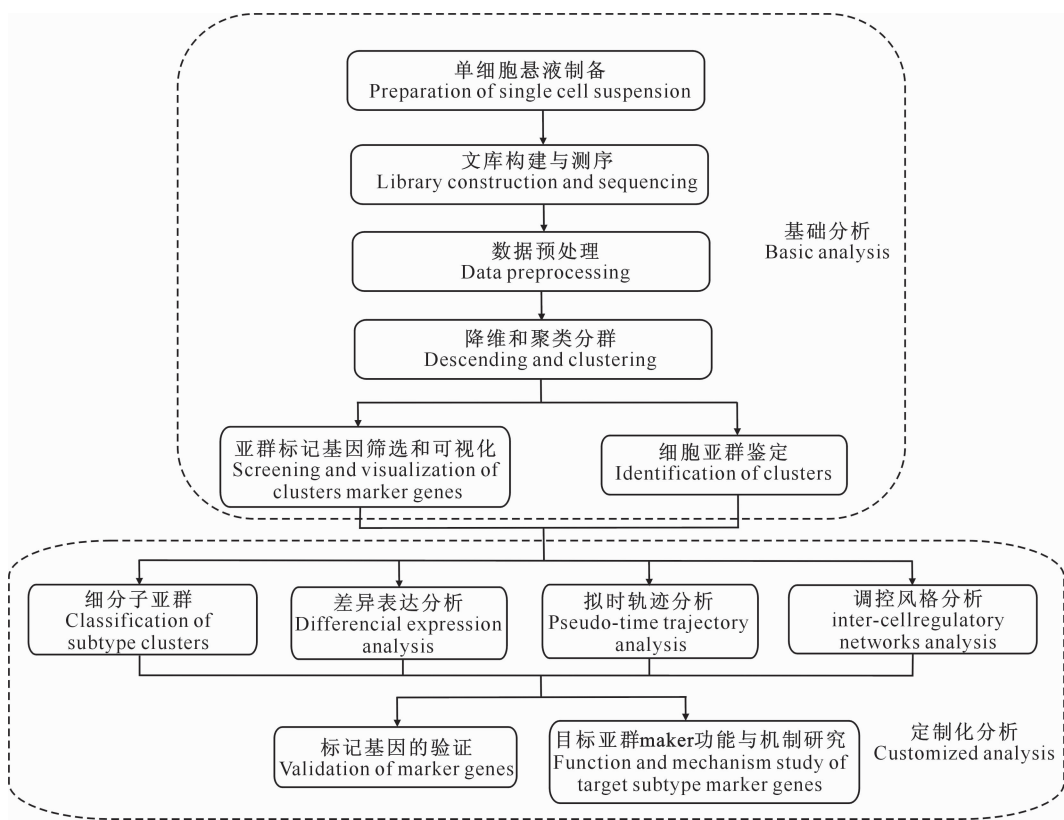


图 2 scRNA-seq 技术基础分析相关步骤

Fig. 2 Basic analysis steps of scRNA-seq technology

### 2.1 scRNA-seq 数据预处理

scRNA-seq 数据的预处理包括质量控制、数据整合和数据标准化 3 个部分<sup>[22-23]</sup>。(1)质量控制。每个样本获得的数据并非全部合格,需要通过检测细胞中的基因数、UMI 总数,过滤掉来源于空载、过载的 GEM 数据和状态异常的细胞信息<sup>[24]</sup>。UMI 是 cDNA 扩增前给每个 cDNA 连上的标签序列,一个细胞中的 UMI 总数代表细胞中初始 mRNA 的总量,故检测基因数或 UMI 总数太少的单细胞数据可能来自于空载的 GEM;检测基因数或 UMI 总数太多的单细胞数据可能来自于过载的 GEM。另外可以通过线粒体基因的超表达去除自噬凋亡的细胞。(2)数据整合。将各样本数据合并以消除试验过程中生物因素、技术因素以及批次效应引起的数据偏差,尽可能展示单个细胞的真实表达情况<sup>[25-26]</sup>。(3)数据标准化。包括由基因特异性产生偏差的样本内标准化及测序深度和转录本捕获率等造成的样本间标准

化。数据标准化后能够有效降低技术噪音,获得可靠且可重复的结果<sup>[27]</sup>。常用每千个碱基转录每百万映射读取的 reads (reads per kilobase million, RPKM)、每千个碱基转录每百万映射读取的碎片数 (fragments per kilobase million, FPKM) 和每百万条 reads 的转录本 (transcripts permillion, TPM)<sup>[28]</sup> 作为标准化数值。

### 2.2 scRNA-seq 数据降维和聚类分群

单细胞转录组数据是高维数据,涉及数千个基因和大量细胞,在高维基因表达空间进行比较时,群体间和群体内的差异难以区分。常用的策略是将数据投影到较低维的空间中达到可视化分析,即进行降维。降维分析可最佳地保留原始数据的一些关键属性,从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),比使用所有基因选择 HVGs 进行降维更为有效。常用的降维方法有线性降维方法,如主成分分析

(principal component analysis, PCA); 非线性降维方法, 如  $t$  分布随机邻接嵌入 ( $t$ -distributed stochastic neighbor embedding, t-SNE)<sup>[29]</sup>、均匀流形近似和投影 (uniform manifold approximation and projection, UMAP)<sup>[30]</sup> 等。

将数据降维后再进行聚类分析<sup>[31]</sup>, 通过基因表达谱的差异来区分细胞, 其中基于转录组相似性的无监督聚类分析是 scRNA-seq 聚类的主流方法, 最常采用的是 K-means 算法 (K-means clustering algorithm) 和图聚类算法 (graph-based)。基于图聚类算法的 Seurat 软件是分析单细胞数据的常用软件, 通过计算细胞间的差异性, 可优化细胞间聚类关系距离的权重值, 实现对细胞的聚类。

### 2.3 基于 scRNA-seq 技术的亚群标记基因筛选和可视化

细胞分为若干亚群后, 需要依赖标记基因对细胞亚群进行鉴定。将聚类后的细胞亚群进行差异表达分析能确定差异表达基因 (differential expressed genes, DEGs) 或标志基因, 可反映细胞亚群的特异性<sup>[32]</sup>。scRNA-seq 技术通常关注亚群特异上调表达的基因, 采取一对多的策略, 如要关注特定 2 个子亚群细胞的差异, 则与常规转录组相似, 关注彼此上下调的基因。将标记基因的信息与 scRNA-seq 数据对应, 通过小提琴图、气泡图、热图和基因映射图可视化标记基因的表达量。(1) 小提琴图。此类图可以反映细胞亚群中各个细胞的标记基因表达量分布, 常用于细胞类型的验证, 但对低表达的标记基因不敏感。(2) 气泡图。此类图反映亚群中各个标记基因的平均表达量和表达标记基因的细胞比例, 可以很好地呈现低细胞数量亚群的表达信息, 同时对低丰度的标记基因有一定的检出率, 但会将表达量信息平均化, 掩盖低比例的表达目标基因的细胞信息。(3) 热图。此类图可以综合反映基因表达水平、表达标记基因的细胞亚群分布及细胞亚群中表达标记基因的细胞数量。(4) 基因映射图。此类图可以给予最直观的表达标记基因细胞分布情况, 但一张图只能呈现 1 个基因, 缺乏呈现多因素结果的能力, 更适用于展示基因分布而非亚群鉴定。合理选择数据呈现形式有助于更直观地判断细胞类型, 同时可以采用多图结合分析。

### 2.4 细胞亚群鉴定

参考已有文献和数据库, 筛选样本组织中常见细胞类型的已知标记基因, 通过标记基因与测序数据中各细胞亚群的上调/下调基因进行匹配, 从而鉴

定各细胞亚群所属细胞类型, 但通常需要综合多个标记基因确定其细胞类型。目前, 可通过综合样本信息、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人工鉴定细胞类型。此外还可采用自动化注释工具对细胞类型进行鉴定, 目前已开发出 30 多种自动化鉴定工具, 包括 SingleR、CHETAH、cellassign 等<sup>[33]</sup>。借助这类软件, 可通过 Marker 基因的分布判断细胞亚群所属细胞类型, 提高细胞注释的效率, 但对于参考数据集的依赖性强, 而且对于相似度很高的细胞类型鉴定准确性会降低。人工鉴定和用自动化注释工具鉴定方法各有优劣, 可以相互结合用于细胞类型鉴定。

## 3 scRNA-seq 技术在植物研究中的应用

在多细胞生物中, 细胞存在固有的异质性。单细胞测序技术极大地丰富了对细胞异质性和细胞功能的理解。近年来, 随着植物原生质体制备等难题的突破, 植物单细胞测序技术研究愈发火热, 被广泛应用于拟南芥、水稻、玉米、番茄、杨树、花生等物种, 具备绘制单细胞图谱与挖掘标记基因、解析非生物胁迫响应机制、构建细胞分化发育轨迹图谱、揭示细胞分化发育的转录因子及其调控网络功能。

### 3.1 绘制单细胞图谱与挖掘标记基因

机体变化是一个连续过程, 细胞并非从一个状态跳跃到另一个状态, 而是一个渐变的过程, 涉及很多中间态。为了更精细地呈现细胞亚群间的渐变式关系, 可以通过分化轨迹对细胞亚群分类分析进行补充<sup>[34]</sup>。Liu 等<sup>[35]</sup>通过对水稻表皮细胞和根毛细胞分化轨迹的分析, 发现在分化过程中存在的过渡态细胞近根毛表皮细胞。目前, 多数研究在初步完成细胞分类鉴定后, 一般还会选取较关注的细胞类型进一步细分子亚群, 并对其进行差异分析。例如 Shulze 等<sup>[36]</sup>在拟南芥根中鉴定得到中柱 (4 个簇)、内皮层 (2 个簇)、皮层 (2 个簇)、毛细胞 (1 个簇)、非毛细胞 (4 个簇) 和小柱 (2 个簇) 6 个大亚群, 其中每个亚群又可进一步细分为子亚群, 如中柱细胞亚群, 包括原生木质部 (簇 13)、韧皮部样 (簇 14)、分生木质部 (簇 15) 和周环 (簇 16) 4 个子亚群。

在得到的 scRNA-seq 数据中, 可以挖掘出在某个聚类中高度特异表达的潜在标记基因, 还可采用激光扫描共聚焦免疫荧光 (laser scanning confocal immunofluorescence, LSCIF)<sup>[37]</sup>、荧光原位杂交 (fluorescence in situ hybridization, FISH)<sup>[38]</sup> 等技

术在体内验证这些标记基因。Denyer 等<sup>[39]</sup>用有监督分类方法对拟南芥根组织注释细胞类型后,还用无监督分类方法在其差异基因集中筛选到了数百个自定义的标记基因,并从中挑选出 10 个特异性高的有关根发育的标记基因,为了在体内对这些标记基因进行验证,使用转录启动子  $3 \times \text{Venus-NLS}$  报告系,用包含 *Kpn* I 和 *Xma* I 限制性位点的 PCR 引物扩增,并通过经典克隆引入 JM164,以产生与核定位的三重 Venus 标签的转录融合,将所有构建体转化到 Col-0 背景中,利用激光扫描共聚焦显微镜成像,评估了选定基因的时空表达模式。Liu 等<sup>[35]</sup>通过 scRNA-seq 数据集揭示了水稻根中的 8 个细胞聚类,对应于 8 个主要细胞类型,通过鉴定一组新的水稻特异性标记基因并将其用于原位杂交和报告基因技术来进行验证。

### 3.2 解析非生物胁迫响应机制

非生物胁迫是植物生长发育过程中重要的环境影响因素。通过 scRNA-seq 技术,探索在不同处理条件下植物组织中细胞类型和细胞群体的组成变化,有利于了解单细胞水平上植物非生物胁迫反应机制和发育生物机理。2020 年,Wang 等<sup>[40]</sup>对野生型水稻和高盐胁迫下水稻幼苗进行单细胞测序,共获得 4 580 个细胞并鉴定出 5 种细胞类型,对野生型与对照组差异基因分析发现,非生物胁迫诱导的转录组变化因细胞类型而异,非生物胁迫处理也改变了细胞群体的组成,减缓了向叶肉细胞的分化。同年,Wendrich 等<sup>[41]</sup>对 6 日龄拟南芥幼苗根尖进行单细胞转录组测序发现,TMO5/LHW 复合体(target of monopteros 5/lonesome highway)的靶基因主要在根毛细胞中富集,在低磷条件下,TMO5/LHW 复合体会诱导维管细胞中可移动细胞分裂素的合成,从而通过改变表皮细胞的长度和细胞命运来增加根毛密度。此外,Jean-Baptiste 等<sup>[42]</sup>对拟南芥幼苗进行热胁迫处理,构建了热激响应基因调控网络,发现热激响应基因在不同细胞类型中都有表达,但仍有一些其他基因在不同细胞类型中存在显著表达差异。

基于单细胞水平,比较植物同一物种不同亚种间或不同物种间的单细胞转录组图谱,不仅可以揭示发育过程中分化轨迹的差异性和保守性,也有助于深入解析不同亚种或不同物种在应对外部环境刺激时的响应机制。Liu 等<sup>[35]</sup>分别构建了 2 个重要水稻栽培亚种(Nip 和 93-11)根尖的转录组图谱,且 2 个水稻亚种的发育轨迹显示出高度一致的拟时间顺

序,揭示了不同亚种之间发育轨迹的保守性;功能富集分析发现,2 个水稻亚种每种细胞类型的差异表达基因多与环境响应有关,但不同水稻亚种在受到外部环境刺激时响应机制存在差异。

### 3.3 构建细胞分化发育轨迹图谱

在生物学过程中,机体为响应各种环境刺激,细胞在不同状态下进行转变。通过构建植物细胞的发育轨迹,可以深入挖掘随着时间变化细胞类型的改变,还可以进一步解析细胞分化路径,了解植物细胞的动态发育过程,揭示复杂组织中关键的细胞生物学进程。对于单细胞转录组数据,可使用所有集群进行拟时轨迹分析,也可以选取具有潜在分化关系的集群进行拟时轨迹分析。Monocle 是常用的拟时分析软件,该软件采用了机器学习算法,能自动化优化分析参数,找出最优轨迹类型,运行效率高。研究者在拟南芥的根冠<sup>[43]</sup>、表皮<sup>[44]</sup>、分生组织<sup>[39]</sup>、内皮层<sup>[36]</sup>、气孔<sup>[45]</sup>、木质部<sup>[41]</sup>,玉米生殖细胞<sup>[46]</sup>以及杨树木质部<sup>[47]</sup>等都采用拟时轨迹分析揭示细胞分化轨迹。2019 年,Zhang 等<sup>[43]</sup>对拟南芥根尖进行单细胞转录组测序,共获得由 7 695 个根尖细胞聚类的 24 个细胞亚群,通过细胞注释鉴定出较为稀有的细胞类型,并找到了一批细胞类型标记基因。2020 年,Liu 等<sup>[45]</sup>对 5 日龄拟南芥幼苗子叶进行单细胞测序,共获得由 12 844 个细胞聚类的 11 个细胞亚群,用 Monocle2 进行拟时发育轨迹分析,构建了从拟分生组织母细胞(meristemoid mother cell, MMC)到保卫细胞(guard cells, GC)的气孔谱系细胞的不同发育过程,揭示了它们之间潜在的相互作用。

### 3.4 揭示细胞分化发育的转录因子及其调控网络功能

植物协调生长发育过程中,转录因子调节不同的信号转导因子以响应不同变化,在细胞命运决定中起到核心作用,对不同细胞类型的转录因子调控网络分析,有助于深入理解细胞发育的生物学功能。

目前,有研究对在特定细胞类型中显著上调的基因进行了功能预测,并通过实时荧光定量 PCR 和基因敲除等方法进行验证;采用分化轨迹分析可以找出驱动亚群分化的关键基因,一般会特别关注分支差异的转录因子。Denyer 等<sup>[39]</sup>通过拟时轨迹分析追踪拟南芥的发育轨迹,利用高度相互关联的转录因子表达波动描绘了一个从干细胞到逐步分化过程中相关细胞的精细级联。Liu 等<sup>[48]</sup>根据轨迹分析描述了花生中从原基和薄壁组织向叶肉细胞类型分

化的假设模式,以及原始叶原基在叶片大小扩张过程中直接进化为表皮细胞的模式,通过鉴定 8 个核心转录因子在每个细胞中的表达分布,发现转录因子潜在调节细胞簇位置 and 分化方向之间的关系,如总体基因表达分布表明 *NAC002* 可能决定叶细胞生长的多方面机制,一些转录因子表达的抑制可能决定花生植物叶片海绵细胞的分化状态。Jean-Baptiste 等<sup>[42]</sup>通过鉴定拟南芥中随着发育轨迹而表达的转录因子基序,得到了一些与特定细胞类型高度相关的转录因子,评估并解释了总 RNA 表达沿着发育轨迹的变化,并展示出轨迹分支点标记发育决策。

差异基因分析可以说明在细胞类型中哪些基因上调或下调,对于标记基因之间的调控关系、哪些基因处于调控网络的核心位置则需要进行基因调控网络分析。加权基因共表达网络分析(weighted gene co-expression network analysis, WGCNA)<sup>[49]</sup>是最

常见的一种基于大样本转录组数据的基因调控网络构建方法。Liu 等<sup>[45]</sup>构建了拟南芥气孔分生母细胞中基因表达的转录因子调控网络,通过进一步分析所鉴定的标记基因和转录因子的拟南芥突变体的气孔发育表型,确定了这些基因在调控气孔发育中所扮演的重要角色。Denyer 等<sup>[39]</sup>通过构建拟时分析中 239 个转录因子的基因调控网络,发现了拟南芥根系生毛细胞分化发育阶段相关的核心基因及其相互调控作用,其中大部分核心基因与根发育相关研究尚未见报道。对其中 25 个核心组件进行分析,发现了一系列在根生长和分化中起到广泛作用的负反馈调控的转录因子,并通过对皮层细胞的类似分析得到了肯定结果。

scRNA-seq 技术在植物 scRNA-seq 中的研究进展见表 3。

表 3 scRNA-seq 技术在植物 scRNA-seq 中的研究进展  
Table 3 Research progress of scRNA-seq technology in plant scRNA-seq

| 分析技术<br>Analysis<br>technique | 物种<br>Species                          | 样本类型<br>Sample type            | 细胞数<br>Cells<br>number  | 总基因数<br>Total<br>number<br>of genes | 每个细胞中的<br>基因数<br>Genes number<br>of per cell | 主要结果<br>Main<br>results                                                                                                                                                                         | 文献来源<br>Refer-<br>ences |
|-------------------------------|----------------------------------------|--------------------------------|-------------------------|-------------------------------------|----------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------|
| CEL-seq2                      | 玉米<br>Maize                            | 雄性生殖细胞<br>Male germ<br>cells   | 144                     | 101 245                             | —                                            | 重建了玉米雄性减数分裂的发育程序<br>Reconstruct the developmental<br>program into meiosis in maize                                                                                                              | [46]                    |
| Smart-seq2                    | 拟南芥<br><i>Arabidopsis<br/>thaliana</i> | 侧根<br>Lateral root             | 573                     | 22 557                              | 4 415                                        | 揭示了 LR 形成的本体层次结构,主要根组织和干细胞的早期和连续分裂<br>Reveals an ontological hierarchy for LR formation with an early and sequential split of main root tissues and stem cells                                  | [50]                    |
| Smart-seq2                    | 拟南芥<br><i>Arabidopsis<br/>thaliana</i> | 雌性配子体<br>Female<br>gametophyte | —                       | —                                   | 5 456<br>14 619<br>5 460                     | 揭示了拟南芥雌性配子体中倍性依赖和细胞特异性转录组的变化<br>Revealing the changes of ploidy-dependent and cell-specific transcriptome in female gametophytes of <i>Arabidopsis thaliana</i>                                 | [51]                    |
| Drop-seq                      | 拟南芥<br><i>Arabidopsis<br/>thaliana</i> | 根-木质部<br>Root-xylem            | 374                     | —                                   | —                                            | 揭示了响应 VND7 表达的木质部细胞的身份转变<br>Revealing the identity transition of xylem cells responding to VND7 expression                                                                                      | [52]                    |
| Drop-seq                      | 拟南芥<br><i>Arabidopsis<br/>thaliana</i> | 根<br>Root                      | 12 198                  | —                                   | 1 000                                        | 揭示了细胞类型对蔗糖反应的相对频率变化<br>Revealing the changes in the relative frequencies of cell types in response to sucrose                                                                                   | [36]                    |
| 10× Genomics                  | 拟南芥<br><i>Arabidopsis<br/>thaliana</i> | 根<br>Root                      | >110 000                | —                                   | —                                            | 揭示了拟南芥野生型和突变体根系的发育轨迹<br>Revealing developmental trajectories in wild type and cell identity mutants                                                                                             | [53]                    |
| 10× Genomics                  | 拟南芥<br><i>Arabidopsis<br/>thaliana</i> | 根-维管细胞<br>Root-vascular cell   | 5 145                   | 21 492                              | 6 781                                        | 揭示了植物表皮对土壤中低磷酸盐环境的响应<br>Revealing plant epidermis responds to low phosphate environment in soil                                                                                                 | [41]                    |
| 10× Genomics                  | 拟南芥<br><i>Arabidopsis<br/>thaliana</i> | 根<br>Root                      | 3 121<br>1 009<br>1 079 | 22 419<br>21 237<br>22 971          | 6 152<br>1 009<br>4 079                      | 构建了热激响应基因调控网络<br>Revealing a regulatory network of heat-responsive genes                                                                                                                        | [42]                    |
| 10× Genomics                  | 拟南芥<br><i>Arabidopsis<br/>thaliana</i> | 根<br>Root                      | 7 522                   | 22 000                              | 5 000                                        | 揭示了单个细胞从分生组织向根毛和非根毛细胞分化成熟阶段的发展轨迹<br>Revealing developmental trajectories for individual cells progressing from meristematic through mature stages of root-hair and nonhair cell differentiation | [44]                    |

表 3(续) Continued table 3

| 分析技术<br>Analysis<br>technique | 物种<br>Species                          | 样本类型<br>Sample type                           | 细胞数<br>Cells<br>number | 总基因数<br>Total<br>number<br>of genes | 每个细胞中的<br>基因数<br>Genes number<br>of per cell | 主要结果<br>Main results                                                                                                                                                                                                                                                       | 文献<br>来源<br>Refere-nces |
|-------------------------------|----------------------------------------|-----------------------------------------------|------------------------|-------------------------------------|----------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------|
| 10×Genomics                   | 拟南芥<br><i>Arabidopsis<br/>thaliana</i> | 根<br>Root                                     | 4 727                  | 16 975                              | —                                            | 揭示了拟南芥根的时空发育轨迹,确定了所有主要细胞类型的表达特征<br>Revealing the developmental trajectories of root, identifying defining expression features for all major cell types                                                                                                                     | [39]                    |
| 10×Genomics                   | 拟南芥<br><i>Arabidopsis<br/>thaliana</i> | 根<br>Root                                     | 7 695                  | 23 161                              | —                                            | 揭示了拟南芥根尖细胞的高度异质性,并重构了根尖分生组织细胞的发育轨迹<br>Revealing a high degree of heterogeneity of root cells and reconstructing the developmen-<br>tal trajectories of root apical meristem cells                                                                                          | [43]                    |
| 10×Genomics                   | 拟南芥<br><i>Arabidopsis<br/>thaliana</i> | 叶-维管细胞<br>Leaf-vascular cell                  | 5 230                  | —                                   | 3 342                                        | 揭示了叶维管系统以及叶细胞类型的作用及关系<br>Revealing the leaf vasculature and the role and relationship of the leaf cell types                                                                                                                                                               | [54]                    |
| 10×Genomics                   | 拟南芥<br><i>Arabidopsis<br/>thaliana</i> | 叶-气孔谱系细胞<br>Leaf- stomatal line-<br>age cells | 12 844                 | —                                   | —                                            | 揭示了气孔分生母细胞中基因表达的转录因子调控网络<br>Revealing transcription factor regulatory net-<br>work of gene expression in meristemoid moth-<br>er cells                                                                                                                                     | [45]                    |
| 10×Genomics                   | 拟南芥<br><i>Arabidopsis<br/>thaliana</i> | 营养茎尖<br>Vegetative shoot<br>apex              | 36 643                 | —                                   | —                                            | 在单细胞水平上揭示了拟南芥茎尖细胞的异质性,以及茎尖干细胞分化成为不同细胞类型的动态连续过程<br>Revealing the heterogeneity of stem tip cells in <i>Arabidopsis thaliana</i> at the single cell level and describing the dynamic continuous process of stem tip stem cell differentiation into differ-<br>ent cell types | [55]                    |
| 10×Genomics                   | 水稻<br>Rice                             | 根<br>Root                                     | 10 968<br>12 564       | —                                   | 2 592<br>2 636                               | 在单细胞水平上鉴定了水稻根尖主要细胞类型的转录图谱,揭示了水稻不同亚种之间发育轨迹的保守性<br>Revealing the transcriptomic landscape of ma-<br>jor cell types of the rice root tip at single-cell resolution and conservativeness of develop-<br>mental trajectory among different subspecies of rice                   | [35]                    |
| 10×Genomics                   | 水稻<br>Rice                             | 根<br>Root                                     | 27 469                 | 29 919                              | —                                            | 重建了表皮细胞和基础组织的连续发育轨迹,揭示了双子叶植物和单子叶植物之间保守和不同的根发育途径<br>Reconstruct continuous developmental trajec-<br>tories of epidermal cells and reveal conserved and divergent root developmental pathways between dicots and monocots                                                    | [56]                    |
| 10×Genomics                   | 水稻<br>Rice                             | 茎和鞘<br>Stem and sheath                        | 4 580                  | —                                   | 2 480                                        | 概述了水稻对非生物胁迫的细胞和发育反应<br>Cellular and developmental responses of rice to abiotic stresses                                                                                                                                                                                    | [40]                    |
| 10×Genomics                   | 玉米<br>Maize                            | 茎尖<br>Stem                                    | 327<br>12 967          | —                                   | 2 221<br>4 965                               | 解析了玉米幼苗中干细胞的功能和细胞命运的获得过程<br>Function and fate of stem cells in maize seed-<br>lings                                                                                                                                                                                        | [57]                    |
| 10×Genomics                   | 玉米<br>Maize                            | 叶片<br>Leaf                                    | 3 763                  | —                                   | 4 874                                        | 揭示了玉米叶片韧皮部蔗糖转运机制<br>Revealing sucrose transport mechanism in phloem of maize leaves                                                                                                                                                                                        | [58]                    |
| 10×Genomics                   | 花生<br>Peanut                           | 叶<br>Leaf                                     | 6 815                  | 46 004                              | 1 807                                        | 揭示了花生叶片细胞分化发育轨迹<br>Revealing the developmental trajectory of pea-<br>nut leaf cell differentiation                                                                                                                                                                         | [48]                    |
| 10×Genomics                   | 番茄<br>Tomato                           | 茎尖<br>Stem                                    | 13 377                 | 21 402                              | —                                            | 通过 snRNA-seq 鉴定了主要的茎尖组织和发育阶段的不同细胞类型,描绘了细胞分化发育轨迹<br>Different cell types of main stem tip tissues and developmental stages were identified by snRNA-seq, and the trajectory of cell differ-<br>entiation and development was depicted                                       | [59]                    |
| 10×<br>Genomics               | 杨树<br><i>Populus</i>                   | 木质部<br>Xylem                                  | 9 798                  | 50 762                              | 3 673                                        | 揭示了杨树木质部高分辨率表达图谱<br>Revealing a high-resolution cell atlas of xylem in <i>Populus</i>                                                                                                                                                                                      | [47]                    |

## 4 scRNA-seq 在甘薯上的应用前景

甘薯块根是甘薯的主要收获器官,根的发育和膨大直接影响甘薯产量。随着甘薯参考基因组以及三浅裂野牵牛(*I. trifida*)、牵牛花(*I. nil*)和三裂叶薯(*I. triloba*)等近缘野生种基因组的解析,研究者挖掘了一些参与甘薯块根形成及膨大的转录因子和基因<sup>[60]</sup>。在甘薯及其近缘野生种基因组注释的基础上,可以采用 scRNA-seq 技术来探索甘薯不同组织器官中细胞的功能及细胞群体间的协作关系,确定不同类型细胞中的基因表达模式和调控网络,对甘薯中一些基因功能及关系、甘薯块根膨大的系统调控机制等问题进行深入研究。此外,甘薯可以选取茎段或顶端嫩茎直接扦插种植,且甘薯离体叶片能形成根,生根的甘薯叶片先在叶柄出现愈伤组织,然后形成贮藏根与纤维根。利用 scRNA-seq 技术可以在单细胞分辨率下识别甘薯不同组织的单细胞转录组信息,针对每个细胞信息得到从一种细胞类型分化为另一种细胞类型的过程,也可能会寻找到该分化过程中的过渡态细胞等,这是 bulk RNA-seq 技术不能做到的,因此具有重要意义。

在甘薯品质、抗逆机制研究方面,bulk RNA-seq 技术已有广泛应用。例如,Yang 等<sup>[61]</sup>对高、低淀粉含量甘薯品种进行转录组比较,来挖掘一些与淀粉合成、激素调控及信号转导相关的基因及转录因子。张毅等<sup>[62]</sup>利用盐胁迫后甘薯的多转录组数据与 WGCNA 方法相结合,鉴定甘薯耐盐相关共表达网络及核心基因。2016 年,Cao 等<sup>[63]</sup>通过对甘薯栽培种与野生种三浅裂野牵牛进行比较转录组分析,挖掘了对植物响应干旱、盐碱等逆境胁迫有关的包含 bZIP、bHLH、MYB 等在内的 51 个转录因子家族。但 bulk RNA-seq 技术不能解决细胞之间的异质性,而 scRNA-seq 技术可以分析品质、抗逆相关的细胞类型及基因表达模式,结合差异表达分析和调控网络分析,可以多层次揭示甘薯关键品质形成和抗逆机制,因此具有较好的应用前景。

## 5 展 望

目前,植物单细胞转录组测序研究的难点在于单细胞悬液的制备,需要先将植物细胞制备成原生质体,但模式植物原生质体提取方法不适用于非模式植物,故需要根据植物组织的特性不断优化酶解条件,从而分离出数量多、活性高的原生质体。但目前对于甘薯原生质体提取的研究报道较少,导致包

括亚细胞定位、遗传转化和蛋白质互作验证等原生质体相关研究无法开展,而将 scRNA-seq 技术用于甘薯研究可以解决上述问题。针对不能制备成单细胞悬液或细胞过大的植物,可以采用单核转录组测序(single-nuclei RNA sequencing, snRNA-seq)技术,将细胞核从细胞质裂解物中分离出来进行建库测序<sup>[59]</sup>。此外,有些细胞类型数量较少,无法在 scRNA-seq 数据集中识别相应的细胞簇,通过基于细胞类型标记的细胞分选技术可以有效筛选稀有细胞。因此综合采用 scRNA-seq 技术与基于细胞类型标记的细胞分选技术可能有助于研究植物中稀有细胞类型的维持和分化。

scRNA-seq 技术是通过收集离体的细胞进行测序,在建库过程中导致细胞在组织中原始位置的不可还原性,组织微环境、胞间联系也无法还原,这些局限导致部分问题无法获得确切回答,如细胞微环境和空间异质性。如今,很多研究者都认为细胞微环境对分子机理研究有至关重要的推进作用,比如植物根系的发育受到内源性共生菌构建的微环境影响;组织层面,如表皮细胞、木质部细胞、气孔等的空间异质性很容易观察到,但在组织内部存在的空间异质性却鲜有研究。现有的空间基因表达图谱主要是通过原位杂交这种低通量方法创建的,但方法复杂,限制了多样本评估。空间转录组(spatial transcriptomics-seq, ST-seq)技术为空间异质性研究提供了解决方案,实现了转录本所属组织位置的还原<sup>[64]</sup>,但仅知道转录本在什么位置,而不能确定每个位置由哪些细胞组成。因此,ST-seq 技术和 scRNA-seq 技术可通过关联分析实现优势互补,可以预见植物单细胞研究的 scRNA-seq 技术、利用转座酶研究染色质可进入性的高通量测序(assay for targeting accessible-chromatin with high-throughput sequencing, ATAC-seq)技术<sup>[65]</sup>、高通量染色体构象捕获(High-through chromosome conformation capture, Hi-C)技术<sup>[66]</sup>和 ST-seq 技术的结合,能够从分子、亚细胞、细胞和组织水平了解植物,使绘制高分辨的植物细胞分子时空图谱成为可能<sup>[67]</sup>,从而推动植物科学基础问题的研究。

## [参考文献]

- [1] Zheng C, Zheng L, Yoo J K, et al. Landscape of infiltrating T cells in liver cancer revealed by single-cell sequencing [J]. Cell, 2017, 169(7): 1342-1356.
- [2] Han X P, Zhou Z M, Fei L J, et al. Construction of a human cell landscape at single-cell level [J]. Nature, 2020, 581(7808):

- 303-309.
- [3] Tang F, Barbacioru C, Wang Y, et al. mRNA-seq whole-transcriptome analysis of a single cell [J]. *Nat Methods*, 2009, 6(5):377-382.
- [4] Islam S, Kjallquist U, Moliner A, et al. Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq [J]. *Genome Res*, 2011, 21(7):1160-1167.
- [5] Hashimshony T, Wagner F, Sher N, et al. CEL-seq: single-cell RNA-seq by multiplexed linear amplification [J]. *Cell Reports* (Cambridge), 2012, 2(3):666-673.
- [6] Ramsköld D, Luo S, Wang Y C, et al. Full-length mRNA-seq from single-cell levels of RNA and individual circulating tumor cells [J]. *Nature Biotechnology*, 2012, 30(8):777-782.
- [7] Picelli S, Bjorklund A K, Faridani O R, et al. Smart-seq2 for sensitive full-length transcriptome profiling in single cells [J]. *Nat Methods*, 2013, 10(11):1096-1098.
- [8] Jaitin D A, Keren-Shaul H, 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.
- [9] Fan H C, Fu G K, Fodor S P. Combinatorial labeling of single cells for gene expression cytometry [J]. *Science*, 2015, 347(6222):1258367.
- [10] Xu Y, Mizuno T, Sridharan A, et al. Single-cell RNA sequencing identifies diverse roles of epithelial cells in idiopathic pulmonary fibrosis [J]. *JCI Insight*, 2016, 1(20):e90558.
- [11] Macosko E Z, 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.
- [12] Klein A 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.
- [13] Rosenberg A B, Roco C M, Muscat R A, et al. Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding [J]. *Science*, 2018, 360(6385):176-182.
- [14] Svensson V, Vento-Tormo R, Teichmann S A. Exponential scaling of single-cell RNA-seq in the past decade [J]. *Nat Protoc*, 2018, 13(4):599-604.
- [15] Wu A R, Neff N F, Kalisky T, et al. Quantitative assessment of single-cell RNA-sequencing methods [J]. *Nat Methods*, 2014, 11(1):41-46.
- [16] Hashimshony T, Senderovich N, Avital G, et al. CEL-seq2: sensitive highly-multiplexed single-cell RNA-seq [J]. *Genome Biol*, 2016, 17:77.
- [17] Prashant N M, Liu H, Bousounis P, et al. Estimating the allele-specific expression of SNVs from  $10 \times$  genomics single-cell RNA-sequencing data [J]. *Genes (Basel)*, 2020, 11(3):240.
- [18] Kolodziejczyk A A, Kim J K, Svensson V, et al. The technology and biology of single-cell RNA sequencing [J]. *Molecular Cell*, 2015, 58(4):610-620.
- [19] Chandran D, Scanlon M J, Ohtsu K, et al. Laser microdissection-mediated isolation and *in vitro* transcriptional amplification of plant RNA [J]. *Current Protocols in Molecular Biology*, 2015, 112(1):23A-25A.
- [20] Galbraith D W, Elumalai R, Gong F C. Integrative flow cytometric and microarray approaches for use in transcriptional profiling [J]. *Methods Mol Biol*, 2004, 263:259-280.
- [21] Zhang X N, Li T Q, Liu F, et al. Comparative analysis of droplet-based ultra-high-throughput single-cell RNA-seq systems [J]. *Mol Cell*, 2019, 73(1):130-142.
- [22] Aljanahi A A, Danielsen M, Dunbar C E. An introduction to the analysis of single-cell RNA-sequencing data [J]. *Molecular Therapy Methods & Clinical Development*, 2018, 10:189-196.
- [23] Ziegenhain C, Vieth B, Parekh S, et al. Quantitative single-cell transcriptomics [J]. *Briefings in Functional Genomics*, 2018, 17(4):220-232.
- [24] Ding B, Zheng L N, Zhu Y, et al. Normalization and noise reduction for single cell RNA-seq experiments [J]. *Bioinformatics*, 2015, 31(13):2225-2227.
- [25] Butler A, Hoffman P, Smibert P, et al. Integrating single-cell transcriptomic data across different conditions, technologies, and species [J]. *Nat Biotechnol*, 2018, 36(5):411-420.
- [26] Tran H, Ang K S, Chevrier M, et al. A benchmark of batch-effect correction methods for single-cell RNA sequencing data [J]. *Genome Biol*, 2020, 21(1):12.
- [27] Chen G, Ning B T, Shi T L. Single-cell RNA-Seq technologies and related computational data analysis [J]. *Front Genet*, 2019, 10:317.
- [28] Han Henry, 刘文斌. Single-cell RNA sequencing 的正则化方法 [J]. *广州大学学报(自然科学版)*, 2019, 18(2):44-55.  
Han H, Liu W B. A survey on normalization of single-cell RNA sequencing [J]. *Journal of Guangzhou University (Natural Science Edition)*, 2019, 18(2):44-55.
- [29] Linderman G C, Rachh M, Hoskins J G, et al. Fast interpolation-based t-SNE for improved visualization of single-cell RNA-seq data [J]. *Nat Methods*, 2019, 16(3):243-245.
- [30] Becht E, McInnes L, Healy J, et al. Dimensionality reduction for visualizing single-cell data using UMAP [J]. *Nature Biotechnology*, 2018, 37(1):38-44.
- [31] 何睿, 余娜, 李森, 等. 基于单细胞 RNA 测序数据的细胞类型聚类算法 [J]. *智能计算机与应用*, 2020, 10(7):104-108.  
He R, Yu N, Li M, et al. Cell type clustering algorithm based on single-cell RNA sequencing data [J]. *Intelligent Computer and Applications*, 2020, 10(7):104-108.
- [32] Soneson C, Robinson M D. Bias, robustness and scalability in single-cell differential expression analysis [J]. *Nat Methods*, 2018, 15(4):255-261.
- [33] Abdelaal T R M, Michielsen L C M, Cats D, et al. A comparison of automatic cell identification methods for single-cell RNA sequencing data [J]. *Genome Biology*, 2019, 20(1):194.
- [34] Haghverdi L, Buttner M, Wolf F A, et al. Diffusion pseudotime robustly reconstructs lineage branching [J]. *Nat Meth*

- ods, 2016, 13(10): 845-848.
- [35] Liu Q, Liang Z, Feng D, et al. Transcriptional landscape of rice roots at the single-cell resolution [J]. *Molecular Plant*, 2021, 14(3): 384-394.
- [36] Shulze C N, Cole B J, Ciobanu D, et al. High-throughput single-cell transcriptome profiling of plant cell types [J]. *Cell Reports*, 2019, 27(7): 2241-2247.
- [37] Bagci I S, Aoki R, Vladimirova G, et al. Simultaneous immunofluorescence and histology in pemphigus vulgaris using ex vivo confocal laser scanning microscopy [J]. *Journal of Biophotonics*, 2021, 14(5): e202000509.
- [38] 何世斌, 柴连琴, 谭珺隽, 等. 荧光原位杂交技术的研究进展 [J]. *植物科学学报*, 2014, 32(2): 199-204.
- He S B, Chai L Q, Tan J J, et al. Recent advances in fluorescence in situ hybridization [J]. *Plant Science Journal*, 2014, 32(2): 199-204.
- [39] 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.
- [40] Wang Y, Hua Q, Chu X, et al. Single-cell transcriptome analyses recapitulate the cellular and developmental responses to abiotic stresses in rice [J/OL]. *Cold Spring Harbor Laboratory*, 2020. [2022-01-30]. <https://doi.org/10.1101/2020.01.30.926329>.
- [41] Wendrich J R, Yang B, Vandamme N, et al. Vascular transcription factors guide plant epidermal responses to limiting phosphate conditions [J]. *Science*, 2020, 370(6518): eaay4970.
- [42] Jean-Baptiste K, McFaline-Figueroa J L, Alexandre C M, et al. Dynamics of gene expression in single root cells of *Arabidopsis thaliana* [J]. *Plant Cell*, 2019, 31(5): 993-1011.
- [43] Zhang T Q, Xu Z G, Shang G D, et al. A single-cell RNA sequencing profiles the developmental landscape of *Arabidopsis* root [J]. *Molecular Plant*, 2019, 12(5): 648-660.
- [44] Ryu K H, Huang L, Kang H M, et al. Single-cell RNA sequencing resolves molecular relationships among individual plant cells [J]. *Plant Physiol*, 2019, 179(4): 1444-1456.
- [45] Liu Z X, Zhou Y P, Guo J G, et al. Global dynamic molecular profiling of stomatal lineage cell development by single-cell RNA sequencing [J]. *Molecular Plant*, 2020, 13(8): 1178-1193.
- [46] Nelms B, Walbot V. Defining the developmental program leading to meiosis in maize [J]. *Science*, 2019, 364(6435): 52-56.
- [47] Li H, Dai X R, Huang X, et al. Single-cell RNA sequencing reveals a high-resolution cell atlas of xylem in *Populus* [J]. *Journal of Integrative Plant Biology*, 2021, 63(11): 1906-1921.
- [48] Liu H, Hu D X, Du P X, et al. Single-cell RNA-seq describes the transcriptome landscape and identifies critical transcription factors in the leaf blade of the allotetraploid peanut (*Arachis hypogaea* L.) [J]. *Plant Biotechnol J*, 2021, 19(11): 2261-2276.
- [49] 宋长新, 雷 萍, 王 婷. 基于 WGCNA 算法的基因共表达网络构建理论及其 R 软件实现 [J]. *基因组学与应用生物学*, 2013, 32(1): 135-141.
- Song C X, Lei P, Wang T. Gene co-expression network analysis based on WGCNA algorithm-theory and implementation in R software [J]. *Genomics and Applied Biology*, 2013, 32(1): 135-141.
- [50] Laura S R, Pablo P G, Alvaro S C, et al. Reconstruction of lateral root formation through single-cell RNA-seq reveals order of tissue initiation [J]. *Molecular Plant*, 2021, 14(8): 1362-1378.
- [51] Song Q, Ando A, Jiang N, et al. Single-cell RNA-seq analysis reveals ploidy-dependent and cell-specific transcriptome changes in *Arabidopsis female* gametophytes [J]. *Genome Biol*, 2020, 21(1): 178.
- [52] Turco G M, Rodriguez-Medina J, Siebert S, et al. Molecular mechanisms driving switch behavior in xylem cell differentiation [J]. *Cell Rep*, 2019, 28(2): 342-351.
- [53] Shahan R, Hsu C W, Nolan T M, et al. A single cell *Arabidopsis* root atlas reveals developmental trajectories in wild type and cell identity mutants [J/OL]. *Cold Spring Harbor Laboratory*, 2020. [2022-01-30]. <https://doi.org/10.1101/2020.06.29.178863>.
- [54] Kim J Y, Symeonidi E, Pang T Y, et al. Distinct identities of leaf phloem cells revealed by single cell transcriptomics [J]. *Plant Cell*, 2021, 33(3): 511-530.
- [55] Zhang T Q, Chen Y, Liu Y, et al. Single-cell transcriptome atlas and chromatin accessibility landscape reveal differentiation trajectories in the rice root [J]. *Nat Commun*, 2021, 12(1): 2053.
- [56] Zhang T Q, Chen Y, Wang J W. A single-cell analysis of the *Arabidopsis* vegetative shoot apex [J]. *Developmental Cell*, 2021, 56(7): 1056-1074.
- [57] Satterlee J W, Strable J, Scanlon M J. Plant stem-cell organization and differentiation at single-cell resolution [J]. *Proc Natl Acad Sci USA*, 2020, 117(52): 33689-33699.
- [58] Bezruczyk M, Zollner N R, Kruse C, et al. Evidence for phloem loading via the abaxial bundle sheath cells in maize leaves [J]. *Plant Cell*, 2021, 33(3): 531-547.
- [59] Tian C H, Du Q W, Xu M, et al. Single-nucleus RNA-seq resolves spatiotemporal developmental trajectories in the tomato shoot apex [J/OL]. *Cold Spring Harbor Laboratory*, 2020. [2022-01-30]. <https://doi.org/10.1101/2020.09.20.305029>.
- [60] 宋炜涵, 后 猛, 张允刚, 等. 甘薯块根膨大机制研究进展 [J/OL]. *分子植物育种*, 2021. [2022-01-30]. <http://kns.cnki.net/kcms/detail/46.1068.S.20210825.0858.002.html>.
- Song W H, Hou M, Zhang Y G, et al. Research advance on mechanism of expansion in sweetpotato storage roots [J/OL]. *Molecular Plant Breeding*, 2021. [2022-01-30]. <http://kns.cnki.net/kcms/detail/46.1068.S.20210825.0858.002.html>.

Shi D M, Lü G, Jiang G Y, et al. Study on anti-erodibility and change of soil physical property in *Pinus massoniana* woodland [J]. Journal of Soil and Water Conservation, 2005, 19 (6): 35-39.

[39] Barthes B, Roose E. Aggregate stability as an indicator of soil susceptibility to run off and erosion; validation at several levels [J]. Catena, 2002, 47(2): 133-149.

[40] 王 冰, 周 扬, 张秋良. 兴安落叶松林龄对土壤团聚体分布及其有机碳含量的影响 [J]. 生态学杂志, 2021, 40(6): 1618-1628.

Wang B, Zhou Y, Zhang Q L. Effects of *Larix gmelinii* stand age on composition and organic carbon content of soil aggregates [J]. Chinese Journal of Ecology, 2021, 40(6): 1618-1628.

[41] 李鉴霖, 江长胜, 郝庆菊. 土地利用方式对缙云山土壤团聚体稳定性及其有机碳的影响 [J]. 环境科学, 2014, 35(12): 4695-4704.

Li J L, Jiang C S, Hao Q J. Impact of land use type on stability and organic carbon of soil aggregates in Jinyun Mountain [J]. Environmental Science, 2014, 35(12): 4695-4704.

[42] 吕 刚, 吕向楠, 王 磊, 等. 辽西低山丘陵区不同密度荆条下表层土壤抗蚀性特征 [J]. 沈阳农业大学学报, 2019, 50(4): 487-493.

Lü G, Lü X N, Wang L, et al. Soil anti-erosion characteristics of surface soil under *Vitex negundo* L. with different densities in low mountainous upland region of western Liaoning Province [J]. Journal of Shenyang Agricultural University, 2019, 50(4): 487-493.

[43] 王春晓, 谢世友, 王 灿. 重庆南川岩溶山区土壤抗蚀性变化及预测模型研究 [J]. 农业现代化研究, 2009, 30(6): 756-760.

Wang C X, Xie S Y, Wang C. Research on changes of soil anti-erodibility and prediction model in Karst mountain areas in Chongqing Nanchuan [J]. Research of Agricultural Modernization, 2009, 30(6): 756-760.

[44] 霍小鹏, 李贤伟, 张 健, 等. 川西亚高山暗针叶林土壤渗透性能研究 [J]. 水土保持研究, 2009, 16(3): 192-195.

Huo X P, Li X W, Zhang J, et al. Soil permeability capability of subalpine coniferous forests in western Sichuan, China [J]. Research of Soil and Water Conservation, 2009, 16(3): 192-195.

[45] 赵锦梅, 徐长林, 马 瑞, 等. 东祁连山不同高寒灌丛草地土壤抗蚀性研究 [J]. 水土保持学报, 2016, 30(5): 119-123.

Zhao J M, Xu C L, Ma R, et al. Study on soil anti-erodibility of different alpine shrub grassland in Eastern Qilian Mountain [J]. Journal of Soil and Water Conservation, 2016, 30(5): 119-123.

[46] 姜兴雪, 查瑞波, 毛兰花, 等. 花岗岩红壤侵蚀退化地不同治理模式土壤抗蚀性评价 [J]. 水土保持学报, 2021, 35(6): 31-37.

Jiang X X, Zha R B, Mao L H, et al. Evaluation of soil erosion resistance under different treatment modes in erosion degraded land of granite red soil [J]. Journal of Soil and Water Conservation, 2021, 35(6): 31-37.

[47] Chaplot V, Cooper M. Soil aggregate stability to predict organic carbon outputs from soils [J]. Geoderma, 2015, 243: 205-213.

[48] 闫思宇, 王景燕, 龚 伟, 等. 川南山地林分变化对土壤物理性质和抗蚀性的影响 [J]. 长江流域资源与环境, 2016, 25(7): 1112-1120.

Yan S Y, Wang J Y, Gong W, et al. Effects of forest change on soil physical properties and anti-erodibility in southern Sichuan mountains [J]. Resources and Environment in the Yangtze Basin, 2016, 25(7): 1112-1120.

[49] 王 燕, 宫渊波, 尹艳杰, 等. 不同林龄马尾松人工林水土保持功能 [J]. 水土保持学报, 2013, 27(5): 23-27, 31.

Wang Y, Gong Y B, Yin Y J, et al. Soil and water conservation functions of *Pinus massoniana* plantation at different forest ages [J]. Journal of Soil and Water Conservation, 2013, 27(5): 23-27, 31.

(上接第 50 页)

[61] Yang S T, Liu X J, Qiao S, et al. Starch content differences between two sweet potato accessions are associated with specific changes in gene expression [J]. Functional & Integrative Genomics, 2018, 18(6): 613-625.

[62] 张 毅, 吴万亿, 刘霞宇, 等. 利用 WGCNA 鉴定甘薯耐盐相关共表达网络及核心基因 [J]. 河南农业科学, 2021, 50(6): 16-27.

Zhang Y, Wu W Y, Liu X Y, et al. Identification of salt tolerance co-expression modules and hub genes in *Ipomoea batatas* by WGCNA [J]. Journal of Henan Agricultural Sciences, 2021, 50(6): 16-27.

[63] Cao Q H, Li A, Chen J Y, et al. Transcriptome sequencing of the sweet potato progenitor (*Ipomoea trifida* (H. B. K.) G. Don.) and discovery of drought tolerance genes [J]. Tropical Plant Biology, 2016, 9(2): 63-72.

[64] Liao J, Lu X Y, Shao X, et al. Uncovering an organ's molecular architecture at single-cell resolution by spatially resolved transcriptomics [J]. Trends in Biotechnology, 2021, 39(1): 43-58.

[65] Buenrostro J D, Giresi P G, Zaba L C, et al. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position [J]. Nat Methods, 2013, 10(12): 1213-1218.

[66] Dong Q L, Li N, Li X C, et al. Genome-wide Hi-C analysis reveals extensive hierarchical chromatin interactions in rice [J]. Plant J, 2018, 94(6): 1141-1156.

[67] Rhee S Y, Birnbaum K D, Ehrhardt D W. Towards building a plant cell atlas [J]. Trends Plant Sci, 2019, 24(4): 303-310.