

• 综 述 •

## 单细胞转录组测序技术及其应用\*

邱晓芬<sup>1,2</sup>, 陈洁晶<sup>2</sup>, 薛 雯<sup>2</sup>, 林 华<sup>2</sup>, 杨 明<sup>2△</sup>, 汤冬娥<sup>3</sup>, 戴 勇<sup>2,3▲</sup>

(1. 广西师范大学生命科学学院, 广西桂林 541006; 2. 中国人民解放军联勤保障部队第九二四医院中心实验室/广西代谢性疾病研究重点实验室, 广西桂林, 541002; 3. 暨南大学第二临床医学院/深圳市人民医院临床医学研究中心, 广东深圳 518020)

**摘 要:** 细胞间普遍存在异质性, 传统转录组测序是以大量细胞或组织为研究样本, 反映的是细胞总体上转录组特征, 不能分析单个细胞的基因表达情况。单细胞转录组测序(sc RNA-seq)技术的发展为揭示单个细胞转录组特征提供了有效方法, 进一步发现基因可变剪接和新转录本, 特别是为数量稀少的, 如早期胚胎细胞、干细胞及疾病分子机制等深入研究开辟了新道路。该文就 sc RNA-seq 技术的方法及应用展开综述。

**关键词:** 单细胞转录组测序; 细胞异质性; 高通量测序

**DOI:** 10.3969/j.issn.1673-4130.2020.15.019

**中图法分类号:** S813.3

**文章编号:** 1673-4130(2020)15-1876-04

**文献标识码:** A

### Single cell transcriptome sequencing and its application\*

QIU Xiaofen<sup>1,2</sup>, CHEN Jiejing<sup>2</sup>, XUE Wen<sup>2</sup>, LIN Hua<sup>2</sup>, YANG Ming<sup>2△</sup>, TANG Dong'e<sup>3</sup>, DAI Yong<sup>2,3▲</sup>

(1. College of Life Sciences, Guangxi Normal University, Guilin, Guangxi 541006, China; 2. Guangxi Key Laboratory of Metabolic Diseases Research/Central Laboratory of the 924th Hospital of Joint Logistics Support Force of PLA, Guilin, Guangxi 541002, China; 3. the Second Clinical Medical College of Jinan University/Clinical Medical Research Center of Shenzhen People's Hospital, Shenzhen, Guangdong 518020, China)

**Abstract:** More and more studies have shown that there is widespread heterogeneity between cells. Traditional transcriptome sequencing takes a large number of cells or tissues as research samples, which reflects the overall transcriptome characteristics of cells and cannot analyze the gene expression of a single cell. The development of single cell transcriptome sequencing (sc RNA-seq) provides an effective method for revealing the characteristics of single cell transcriptome, and further discovering gene alternative splicing and new transcripts, especially for the further study of rare numbers, such as early embryonic cells, stem cells and the molecular mechanism of disease. This paper summarizes the methods and applications of scRNA-seq technology.

**Key words:** single cell transcriptome sequencing; cell heterogeneity; high throughput sequencing

由于基因组和表观基因组的重编程及细胞分裂和分化过程中的 DNA 复制错误, 单个细胞水平会呈现不同的基因组、转录组和表观基因组<sup>[1]</sup>。传统转录组测序方法是对群体水平细胞簇的转录组进行测序, 是对组织平均表达水平的剖析, 难以揭示其中单个细胞的基因表达情况, 如肿瘤细胞的异质性、循环肿瘤细胞(CTC)、最早的胚胎干细胞和肿瘤干细胞<sup>[2]</sup>, 如法医鉴定中痕量生物样本, 这些细胞的转录组难以通过常规测序技术进行分析。单细胞转录组测序(sc RNA-seq)技术的发展和改进用于细胞异质性发现, 也可用于微量样品的转录组分析, 可更全面地理解真核生物转录组基因复杂性, 单细胞测序技术可以确定

单核苷酸多态性(SNP)、拷贝数变异、单细胞基因组的基因组结构变异、基因表达水平、基因融合、可变剪接及单细胞表观基因组中的 DNA 甲基化状态<sup>[3]</sup>, 有助于充分了解疾病与生命活动进程, 进而找到治疗疾病的方法。

#### 1 单细胞分离技术

sc RNA-seq 技术首先要从微生物群落或细胞团中分离出单个细胞。目前已开发出多种适合不同情况的单细胞分离技术, 主要包括连续稀释法、显微操作法、激光捕获显微切割技术(LCM)、荧光激活细胞分选技术(FACS)、微流体控技术、拉曼镊子技术(ROT)<sup>[4]</sup>。单个细胞易受外界环境的影响, 应注意避

\* 基金项目: 广西重点实验室建设项目(17-259-57); 广西壮族自治区桂林市科学研究与技术开发计划项目(2016012702-1)。

△ 通信作者, E-mail: yangming181@yeah.net. ▲ 共同通信作者, E-mail: daiyong22@aliyun.com

本文引用格式: 邱晓芬, 陈洁晶, 薛雯, 等. 单细胞转录组测序技术及其应用[J]. 国际检验医学杂志, 2020, 41(15): 1876-1879.

免机械力与化学作用对单个细胞的破坏。连续稀释法是将细胞置于特定的细胞悬液中梯度稀释,最终获得单个细胞的方法,该方法对设备依赖小,只需手工操作就能实现单细胞分离,成本低廉、操作简单,缺点是不能特异性筛选靶细胞,在操作过程中容易发生分离错误,甚至丢失细胞,因此,更适合分离均一性好的样本。显微操作法是低通量方法<sup>[3]</sup>,在高倍显微镜下用显微操作镊子分离单个细胞,基于细胞形态和着色特征的视觉辨别中在显微镜下观察和拍摄每个独立的单细胞,实现无偏差分离,但耗费人力,因此,常用于从人工培养的细胞系或早期胚胎中分离单个细胞。LCM是在显微镜下观察组织细胞,直接从固定组织的切片中切割和分离单个细胞的方法,可特异性获得目的细胞,同时可以将不需要的细胞切除而获得单个细胞<sup>[5]</sup>。LCM的缺点是分离的单个细胞的完整性可能会被破坏,并且激光能量进行微切割也可能对DNA造成直接损害而导致遗传物质损失,且通量低、成本高。FACS是流式细胞仪基于荧光标记物的细胞特异性特征,如不同的光散射信号实现自动化高通量方式分离单个细胞的技术<sup>[3]</sup>,FACS需要大量悬浮细胞作为起始材料,因此,不适合从数量低的细胞群中分离出单个细胞。ROT是光学镊子和拉曼光谱相组合的技术<sup>[6]</sup>,前者通过轮廓区分细胞而无需染色,后者利用激光捕获细胞,具有同时操作和表征单个细胞而无标签或物理接触的优点,在肿瘤和微生物细胞的异质性研究中起重要作用。微流体控技术广泛应用于DNA测序、药物筛选、化学合成等领域,具有代表性的有两种:液滴式微控和芯片式微流控,这两种方法都是将细胞分离至纳升级的微容器中,提高单细胞制备效率,其中芯片式微流控技术可利用不同特性材料的芯片作为反应容器,将传统生化分析实验室操作集成于微小芯片上,集采样、反应与分离一体,几乎囊括全部化验室功能。可利用少量样品和化学试剂实现单个细胞的识别和分离技术,具有灵敏度高,精确度好,少污染,灵活组合,精确操控、高效筛选目的细胞的优势<sup>[7]</sup>。微流体控技术是一种较新的单细胞分离技术,被认为是一种较合适的方法,在分离单细胞的应用中具有广阔的前景。

## 2 RNA-seq cDNA 文库构建

RNA-seq是以单个细胞为研究对象进行转录组水平测序,并对所得数据进行生物信息学统计分析的技术。单个哺乳动物细胞总RNA水平仅约为10 pg,通常mRNA最多只有0.5 pg<sup>[8]</sup>。但一般高通量测序建文库至少需要几十纳克DNA,细胞内核酸扩增需十万倍以上构建cDNA文库达到测序水平,首要解决的问题是用微量RNA准确建库,避免扩增偏倚,提高灵敏度和可重复性,并获得细胞追溯来源。cDNA文库构建的核心技术是全面捕获目标样本内的总RNA,其中关键步骤是全转录组扩增,将RNA反

转录成cDNA,然后再进行cDNA扩增,必须将RNA完全反转录成cDNA进行线性扩增才能表征单个细胞的转录组特征。目前主要的方法可分为同聚物加尾唐氏法、模板转换法 Smart-seq/Smart-seq2、STRT-seq<sup>[9]</sup>、体外转录扩增 CEL-Seq、Quartz-seq<sup>[3,10-11]</sup>等,以下简要介绍其中几种。

**2.1 同聚物加尾唐氏法** TANG等<sup>[12]</sup>首次使用Tang's RNA-seq在单个小鼠卵母细胞和卵裂球进行转录组分析,是利用同聚物加尾扩增的一种方法。分离单个细胞并通过转移到裂解物缓冲液中并直接在全细胞裂解物上进行反转录。通过核酸外切酶除去游离引物,用末端脱氧核苷酸转移酶将poly(A)尾添加到第一链cDNA的3'末端,PCR循环扩增单细胞cDNA,得到100~200 ng扩增的cDNA用于构建测序文库,文库用于SOLiD系统进行深度测序。

**2.2 体外转录扩增 CEL-Seq** CEL-Seq是体线性(IVT)扩增单细胞转录组的方法。利用含有oligo-dT序列、独特条形码、Illumina测序接头和T7启动子的引物将单细胞RNA反转录成一链cDNA。在合成第2条链后,合并几个样品的cDNA用于IVT。然后将T7转录的RNA片段化并与Illumina衔接子连接,反转录为cDNA,从而可以选择含有Illumina衔接头与条形码的30个最大片段。由于IVT是线性扩增方法,与其他基于PCR的扩增方法比较,可减少扩增偏倚,同时与基于PCR的扩增方法比较,CEL-Seq具有更高的可重复性和灵敏度。HASHIMSHONY等<sup>[10]</sup>通过研究早期秀丽隐杆线虫胚胎发育表明,早在胚胎时期两细胞之间已存在转录产物的差异,CEL-Seq转录组定量方法可用于含有不同细胞类型群体的复杂组织的转录组学分析。

**2.3 模板转换法 Smart-seq/Smart-seq2** DANIEL等<sup>[2]</sup>通过Smart-seq捕获来自黑素瘤的潜在循环肿瘤细胞,筛选黑素瘤循环肿瘤细胞的生物标志物,Smart-seq明显增加长度超过1 kb的所有转录本的全长覆盖率。测定前列腺(PC3和LNCaP)和膀胱(T24)癌细胞系细胞的12个单个细胞的基因表达情况,在这3个细胞系中鉴定数百个差异表达的基因,对数十个起始细胞总量少的RNA检测灵敏度较高。PICELLI等<sup>[13]</sup>在开发的Smart-seq2在Smart-seq的基础上进一步完善反转录,模板转换和预扩增等流程已获得高产量cDNA。Smart-seq2可以增加单细胞的cDNA文库产量和平均长度,并且可以实现更高的灵敏度,降低扩增偏倚,可同时分析数百个细胞,可回收全长cDNA,可分析每个转录物的所有外显子并检测不同的剪接变体,可实现全面的SNP和突变分析,应用领域宽广<sup>[14]</sup>。

## 3 sc RNA-seq 技术的应用

sc RNA-seq其主要应用是分析稀有细胞类型及揭示细胞间异质性<sup>[15]</sup>,在免疫系统<sup>[16]</sup>、神经细

胞<sup>[17-18]</sup>、胚胎干细胞<sup>[19-20]</sup>、细胞图谱构建<sup>[21]</sup>、肿瘤及微环境<sup>[22]</sup>等领域应用。

**3.1 揭示细胞新亚型** JAITIN 等<sup>[23]</sup>对来自超过 4 000 个小鼠脾单细胞的 RNA 进行测序,筛选 1 575 个差异基因,鉴定出细胞类型特异性应答基因组,如巨噬细胞中的 Tnf 和 Marco,自然杀伤细胞(NK 细胞)中的 Xcl1 和 Gzmb,以及大量 I 型干扰素应答基因(Irf7、Stat2、Ifit1、Cxcl10)。

POWELL 等<sup>[24]</sup>第一次直接测量了单个 CTC 中的高维基因表达,与正常细胞系比较,转移 NPTN、S100A4、S100A9 和上皮间充质转变相关的基因升高,VIM、TGFβ1、ZEB2、FOXC1、CXCR4 的转录水平较高。CTC 转录分析受白细胞污染的限制,克服这个问题的方法是单细胞分析。晚期癌症患者的药物选择,血液含有具有这些不同表型的 CTC 的患者可以在优化的多种药物治疗方案中得到更好的治疗。FREDRIK 等<sup>[25]</sup>在单细胞转录组水平上分析了从一只成年大鼠海马体 CA1 亚区捕获的单细胞,表达谱分析揭示,CA1 中至少有两种不同的神经元细胞类型。DOMINIC 等<sup>[26]</sup>随机选择小鼠肠类器官的数百个细胞进行 sc RNA-seq,并开发 RaceID 算法与之结合,将 Reg4 鉴定为肠内分泌细胞的新标记物,肠内分泌细胞是一种罕见的激素产生的肠细胞群,揭示了罕见肠细胞类型的存在。

**3.2 肿瘤微环境研究** 有研究表明,基因突变是肿瘤发生的根本原因,肿瘤微环境研究有助于揭示肿瘤发生的具体机制,CTC 是在实体瘤患者的血液中发现的罕见瘤细胞,实体瘤和 CTC 明显是不同克隆起源<sup>[27]</sup>,可能在癌症传播中起关键作用,揭示 CTC 表型为治疗提供了可能的途径,对于推测分析来自不同肿瘤区域的细胞能阐明肿瘤的进展。TIROSH 等<sup>[28]</sup>从 19 例患者分离的 445 个单细胞进行单细胞测序,分析肿瘤微环境中恶性肿瘤细胞、免疫细胞、基质细胞等在细胞个体、空间、功能和转录组的异质性,结果显示,黑色素瘤细胞系可能存在潜在的耐药群体。

**3.3 免疫系统** MONIKA 等<sup>[29]</sup>已经通过 PCR 技术在多发性肌炎患者的肌肉和血液中鉴定了 CD8 T 细胞的克隆扩增,包括 T 细胞受体(TCR)互补决定区(CDR)3 长度分析(分型)。将 CDR3 光谱分型、激光显微切割和单细胞 PCR 结合来检测这些克隆扩增的 T 细胞可能的致病作用,接触、侵入和破坏骨骼肌纤维的单个细胞毒性 T 细胞,通过 CDR3 光谱类型筛选鉴定的寡克隆峰与形态学表征的显微切割的 T 细胞相关联。1 例患者中部分显微切割的 T 细胞携带共同的 TCR-BV 氨基酸 CDR3 基序和 CDR3 区域中的保守核苷酸交换,结果表明,CDR3 谱型和单细胞分析可以结合起来识别、跟踪血液和靶组织中的自身攻击性 T 细胞克隆起源。这种方法适合其他炎症和自身免疫性疾病。

**3.4 细胞图谱构建** AIZARANI 等<sup>[30]</sup>用 sc RNA-seq 技术对来自 9 例人类供体正常肝脏组织约 10 000 个细胞进行测序分析,创建了所有重要肝脏细胞类型的人类肝脏细胞图谱,其中包含肝细胞、肝脏主要代谢细胞、血管内皮细胞、肝脏巨噬细胞等,鉴定一个胆管细胞亚群之前未知的特性,很可能是肝前体细胞或祖细胞。AIZARANI 等<sup>[30]</sup>解释这类祖细胞在肝脏再生中起重要作用,并且可能参与肝脏疾病或肿瘤产生。VENTO-TORMO 等<sup>[31]</sup>对妊娠早期(6~14 周)胎盘约 7 万个细胞进行单细胞转录组测序并绘制胎盘细胞图谱,鉴定出蜕膜 dNK 细胞的 3 个主要亚群。初次妊娠者 dNK1 细胞亚群细胞与特定的胎盘细胞之间相互作用可能使 dNK1 细胞能够更加有效地应答再次妊娠时的胎盘植入,这一新发现对理解早期妊娠有指导意义,给诊治妊娠相关疾病提供了新思路。

#### 4 小 结

自然界 99% 的微生物还不能人工培养<sup>[32]</sup>,只有少数 CTC 可以从癌症患者的外周血中分离,这些还难以研究的目标都可以通过单细胞测序技术来解决。利用单细胞水平上的高通量 RNA 测序已经在生命研究中有大量新发现,从新细胞类型的鉴定到随机基因表达全局模式的研究,还可以同时应用多组学来共同揭示生物生长发展机制。目前,scRNA-seq 除了要克服检测技术上的通量、效率、灵敏度、成本等问题外,还要克服具体计算策略和分析的挑战。尽管用于分析来自大量细胞群体的 RNA-seq 数据的一些工具可以容易地应用于 scRNA-seq 数据,但是相信随着生物信息学的发展,许多新的计算策略可以充分利用起来解读数据<sup>[33]</sup>,并且能够在基因表达的基础上进行全面详细的分析。

#### 参考文献

- [1] KALISKY T, BLAINEY P, QUAKE S R. Genomic analysis at the single-cell level[J]. *Annu Rev Genet*, 2011, 45 (1): 431-445.
- [2] DANIEL R L, SHUJUN L, YU-CHIEH W, et al. Full-length mRNA-seq from single-cell levels of RNA and individual circulating tumor cells[J]. *Nature Biotechnol*, 2012, 30(8): 777-782.
- [3] LIANG J, CAI W, SUN Z. Single-cell sequencing technologies: current and future[J]. *J Genet Genomics*, 2014, 41 (10): 513-528.
- [4] LOVETT M. The applications of single-cell genomics[J]. *Hum Mol Genet*, 2013, 22(R1): R22-R26.
- [5] DAN F, WASSERSTROM A, ITZKOVITZ S, et al. Amplification of multiple genomic loci from single cells isolated by laser micro-dissection of tissues[J]. *BMC Biotechnol*, 2008, 8(1): 1-16.
- [6] LI M, BOARDMAN D G, WARD A, et al. Single-cell random sorting[J]. *Methods Mol Biol*, 2014, 1096(1): 147-

- 153.
- [7] STREETS A M, ZHANG X, CAO C, et al. Microfluidic single-cell whole-transcriptome sequencing[J]. Proc Natl Acad Sci U S A, 2014, 111(19):7048-7053.
- [8] YONG W, NAVIN N. Advances and applications of single-cell sequencing technologies [J]. Mol Cell, 2015, 58(4):598-609.
- [9] ISLAM S, KJÄLLQUIST 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.
- [10] HASHIMSHONY T, WAGNER F, SHER N, et al. CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification[J]. Cell Rep, 2012, 2(3):666-673.
- [11] SHALEK A K, RAHUL S, JOE S, et al. Single-cell RNA-seq reveals dynamic paracrine control of cellular variation[J]. Nat, 2014, 510(7505):363-369.
- [12] TANG F, BARBACIORU C, NORDMAN E, et al. RNA-Seq analysis to capture the transcriptome landscape of a single cell[J]. Nat Protoc, 2010, 5(3):516-535.
- [13] PICELLI S, BJÖRKLUND Å 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.
- [14] SIMONE P, FARIDANI O R, BJÖRKLUND A K, et al. Full-length RNA-seq from single cells using Smart-seq2 [J]. Nat Protoc, 2014, 9(1):171-181.
- [15] ZHENG H, POMYEN Y, HERNANDEZ M O, et al. Single cell analysis reveals cancer stem cell heterogeneity in hepatocellular carcinoma [J]. Hepatology, 2018, 68(1):127-140.
- [16] BJÖRKLUND A K, FORKEL M, PICELLI S, et al. The heterogeneity of human CD127(+) innate lymphoid cells revealed by single-cell RNA sequencing [J]. Nat Immunol, 2016, 17(4):451-460.
- [17] ZEISEL A, MUÑOZMANCHADO A B, CODELUPPI S, et al. Brain structure, cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq [J]. Science, 2015, 347(6226):1138-1142.
- [18] TREUTLEIN B, LEE Q Y, CAMP J G, et al. Dissecting direct reprogramming from fibroblast to neuron using single-cell RNA-seq [J]. Nature, 2016, 534(7607):391-395.
- [19] ISLAM S, ZEISEL A, JOOST S, et al. Quantitative single-cell RNA-seq with unique molecular identifiers [J]. Nat Methods, 2014, 11(2):163-166.
- [20] FAN X, ZHANG X, WU X, et al. Single-cell RNA-seq transcriptome analysis of linear and circular RNAs in mouse preimplantation embryos [J]. Genome Biol, 2015, 16(1):148-149.
- [21] WINKELS H, EHINGER E, VASSALLO M, et al. Atlas of the immune cell repertoire in mouse atherosclerosis defined by single-cell RNA-sequencing and mass cytometry [J]. Circ Res, 2018, 122(12):1675-1688.
- [22] CIUCCI T, VACCHIO M S, GAO Y, et al. The emergence and functional fitness of memory CD4<sup>+</sup> T cells require the transcription factor thpok [J]. Immunity, 2019, 50(1):91-105.
- [23] JAITIN D A, KENIGSBURG E, KERENSHAUL 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.
- [24] POWELL A A, TALASAZ A H, HAIYU Z, et al. Single cell profiling of circulating tumor cells: transcriptional heterogeneity and diversity from breast cancer cell lines [J]. PLoS One, 2012, 7(5):e33788.
- [25] FREDRIK K, RANELLE S, JINGXUE Y, et al. Single-cell microarray analysis in hippocampus CA1: demonstration and validation of cellular heterogeneity [J]. J Neurosci, 2003, 23(9):3607-3615.
- [26] DOMINIC G, ANNA L, LENNART K, et al. Single-cell messenger RNA sequencing reveals rare intestinal cell types [J]. Nature, 2015, 525(7568):251-255.
- [27] RAMIN N, HUBING S, QI W, et al. Melanomas acquire resistance to B-RAF(V600E) inhibition by RTK or N-RAS upregulation [J]. Nature, 2010, 468(7326):973-977.
- [28] TIROSH I, IZAR B, PRAKADAN S M, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq [J]. Science, 2016, 352(6282):189-196.
- [29] MONIKA H, SOLVEIGH W, HOLGER B, et al. Clonal tracking of autoaggressive T cells in polymyositis by combining laser microdissection, single-cell PCR, and CDR3-spectratype analysis [J]. Proc Natl Acad Sci U S A, 2003, 100(7):4090-4095.
- [30] AIZARANI N, SAVIANO A, MAILLY L, et al. A human liver cell atlas reveals heterogeneity and epithelial progenitors [J]. Nature, 2019, 572(7768):199-204.
- [31] VENTO-TORMO R, EFREMOVA M, BOTTING R A, et al. Single-cell reconstruction of the early maternal-fetal interface in humans [J]. Nature, 2018, 563(7731):347-353.
- [32] RAPPE M S, GIOVANNONI S J. The uncultured microbial majority [J]. Annu Rev Microbiol, 2003, 57(1):369-394.
- [33] OLIVER S, TEICHMANN S A, MARIONI J C. Computational and analytical challenges in single-cell transcriptomics [J]. Nature Reviews Genetics, 2015, 16(3):133-145.