



单细胞转录组技术及其在人类细胞图谱构建中的应用*

倪健¹⁾ 胡苗苗¹⁾ 危莹¹⁾ 陈蕊²⁾ 郭妍^{1)**}

(¹⁾ 上海交通大学生物医学工程学院, 上海 200240; (²⁾ 上海交通大学医学院附属仁济医院整形外科, 上海 200240)

摘要 单细胞转录组技术在单细胞水平上进行转录组测序, 提供了单个细胞的基因表达差异信息, 使在单细胞尺度下研究个体细胞、相关环境细胞及其相互作用的机理成为可能. 近年来, 单细胞转录组技术在 cDNA 扩增原理上经历了从末端加尾、体外逆转录到模板置换的方法发展, 大大提高了基因检测的数量、基因表达的准确性等. 同时, 在单细胞选取方式上进行了从 96/384 孔板到油包水液滴以及纳米微孔的创新, 在提高通量和重复性的同时降低了整体实验成本. 单细胞转录组技术广泛应用于细胞群体分类和异质性研究, 推动了从发育生物学到正常、病态组织细胞图谱的构建. 本文对单细胞转录组技术近年的技术进展以及在人类细胞图谱构建中的应用进行了综述.

关键词 单细胞, 转录组, 异质性, 细胞图谱
中图分类号 Q7

DOI: 10.16476/j.pibb.2019.0033

生物个体的细胞起源于一套遗传物质, 但却能够分化为形态各异、行使不同功能的细胞类型. 而每一种细胞类型通过遗传物质的调控以及和环境的互作, 最终决定组织和器官等复杂结构的形成和功能的实现. 传统转录组测序方法只能分析群体细胞的基因平均表达值, 难以解析细胞群体内部的异质性. 近年来, 快速发展的转录组技术能在单细胞分辨下, 理解细胞的起源、功能和分化等重要生物现象, 同时可以通过研究组织器官内各个细胞间的互作等, 深入了解组织器官的功能. 近年来, 单细胞转录组建库原理的革新以及单细胞大规模平行处理方法的创新, 极大提高了单细胞转录组技术的灵敏性、通量和准确性. 目前单细胞转录组技术的应用已大大推进了发育生物学、正常和病态组织的细胞图谱的构建和研究, 从而推动癌症生物学等领域的研究和发展. 本文主要综述了近年来单细胞转录组技术的最新发展和在人类细胞图谱构建中的应用.

1 单细胞转录组技术发展

1.1 单细胞转录组原理

1.1.1 末端加尾法

单细胞转录组测序最早可以追溯到 2006 年, Kurimoto 等^[1] 通过 Oligo-dT 引物反转单个细胞中的 mRNA 获得一链 cDNA, 再使用末端加 A 的策略获得双链 cDNA. 由于当时二代测序平台尚未建立, 研究人员使用 T7 启动子逆转录 cDNA, 获得 RNA 后同基因芯片杂交获得转录本信息. 该方法奠定了之后末端加尾单细胞方法的基础, 同时也启发了体

* 国家重点研发计划项目(2018YFC1003501), 国家自然科学基金重大项目(91529302)和国家自然科学基金青年项目(31700762)资助.

** 通讯联系人.

Tel: 021-34206632, E-mail: yanguo@sjtu.edu.cn

收稿日期: 2019-02-25, 接受日期: 2019-06-18

外逆转录方法的建立. 2009年, Tang等^[2]沿用了Kurimoto末端加A获得双链cDNA的策略, 利用Applied Biosystem测序平台SOLiD system, 完成了第一个真正意义上的单细胞转录组测序(图1). 随后, 一大批基于末端加尾的单细胞建库方法被报道. 2013年, Quartz-seq^[3]方法对Tang等^[2]的单细胞转录组技术进行优化, 在表征细胞异质性上更有优势, 该方法将反应体系始终控制在一个反应管中, 同时运用抑制性PCR的方法让多余引物形成锅柄结构来极大降低副产物, 此外还优化了反转及合成二链cDNA的条件. 2015年, Nakamura等^[4]认为逆转录合成全长cDNA是一个低效的过程, 同时扩增时会产生偏差, 而UMI^[5](unique molecular identifiers)的精确定量需要较高的测序深度. 为此他们建立了SC3-seq(single-cell mRNA 3-prime end sequencing)技术, 采用了富集3'端的建库方法. 然而随着测序成本的逐渐降低, 以及UMI对转录本绝对定量的优势, UMI的使用在单细胞转录组测序方法中已经成为常规, 该方法在后来也没有被推广. 同年, SUPeR-seq^[6]方法使用半随机引物来抓取polyA(+)(-)的RNA(包括CircRNA), 该方法减小了3'端偏好并意外减小了对rRNA扩增的偏好, 由于操作步骤较多且通量较难提高, 因此也没有推广开. 2017年, Sheng等^[7]将MATQ-seq方法和MALBAC方法结合, 通过反转polyA(+)(-)RNA, 同时结合末端加C获得双链cDNA, 该策略是一种高效的单细胞转录组方法, 和单细胞转录组的金标准方法SMART-seq2^[8]具有同样优势. 2018年, Quartz-seq2^[9]方法结合细胞标签(Barcode)和UMI策略, Barcode的使用在提高建库细胞通量的同时也降低了成本, UMI的使用可对RNA的Transcripts数进行绝对定量分析. 该方法同时优化了末端加A的条件, 大大提高cDNA的产量. Barcode和UMI结合的策略也被后来的几乎所有单细胞高通量转录组测序方法所使用.

1.1.2 体外逆转录线性扩增法

传统转录组研究中扩增cDNA的方法主要依赖PCR. 在UMI尚未应用以前, PCR带来的非线性扩增对于RNA表达的定量分析造成困难, 而体外逆转录扩增方法则大大降低了扩增偏差. CEL-seq^[10]方法使用Oligo-dT引物反转得到一链cDNA, 使用切口平移法获得二链cDNA, 通过T7启动子在体外反转形成ssRNA, 再通过带已知序列接头的ssDNA同逆转录出的RNA连接, 然后反转获得

DNA来引入测序接头. 2014年, MARS-seq^[11]方法在反转引物前加Barcode和UMI, 并通过流式分选仪将细胞分到96孔板中, Barcode的使用使得单细胞可以合并进行cDNA扩增和建库, 而UMI则可以定量计算转录本个数. 2016年, CEL-seq2^[12]方法优化了前一代的反转引物, 在建库时通过带测序接头的半随机引物反转逆转录的RNA, 来引入测序接头引物片段, 大大提高了建库效率. InDrops^[13]方法采用同样的原理, 但将反应腔室从PCR管换到了油包水滴中, 将入口细胞通量提高到上千甚至上万数量级. 总体而言, 体外逆转录扩增法虽然可以降低cDNA的扩增偏差, 但是由于该方法步骤繁琐, RNA的长时间暴露也增大了其降解的风险. 随着UMI的使用, PCR扩增中转录本难以精确定量的问题得以解决, 因此该类方法也没有得到大规模的应用.

1.1.3 模板置换法

模板置换法(template-switching oligonucleotide)是目前应用最广泛的单细胞转录组文库构建方法. 2011年, STAT-seq^[14]方法率先应用了模板置换法, 即使用Oligo-dT反转mRNA获得一链cDNA, 同时在末端形成3~6个C碱基, 再利用模板转换引物TSO(template-switching oligonucleotide)末端的RNA碱基G, 同第一链cDNA末端多加的C碱基杂交, 将模板置换成第一链cDNA来合成二链cDNA. 该高通量单细胞转录组方法在TSO引物中加了细胞Barcode序列, 配合96孔板收集细胞达到5'端富集. 2012年, SAMRT-seq^[15]方法和2009年3'端偏好的方法^[2]和富集5'端的STAT-seq方法不同, 该方法扩增全长cDNA进行建库. 2014年, SMART-seq2^[8]技术在反应体系中加入甜菜碱以提高酶的热稳定性和反转效率, 同时加入Mg²⁺以结合甜菜碱酸盐阴离子形成离子对, 成为DNA不稳定剂, 并使用额外的热循环(50°C 2 min; 42°C 2 min)来解开RNA二级结构, 此外其对TSO的末端做锁核酸修饰, 以增加核酸双链热稳定性并提高杂交的特异性, 促进模板转换. 在建库部分, 使用Tn5转座酶系统, 将DNA打断和加接头一步完成. 目前通过优化达到高灵敏度、高基因覆盖率的SMART-seq2已经成为单细胞转录组测序的金标准, 同时其便捷高效的二链合成配合Illumina转座酶建库, 已经广泛应用于生物学各领域.

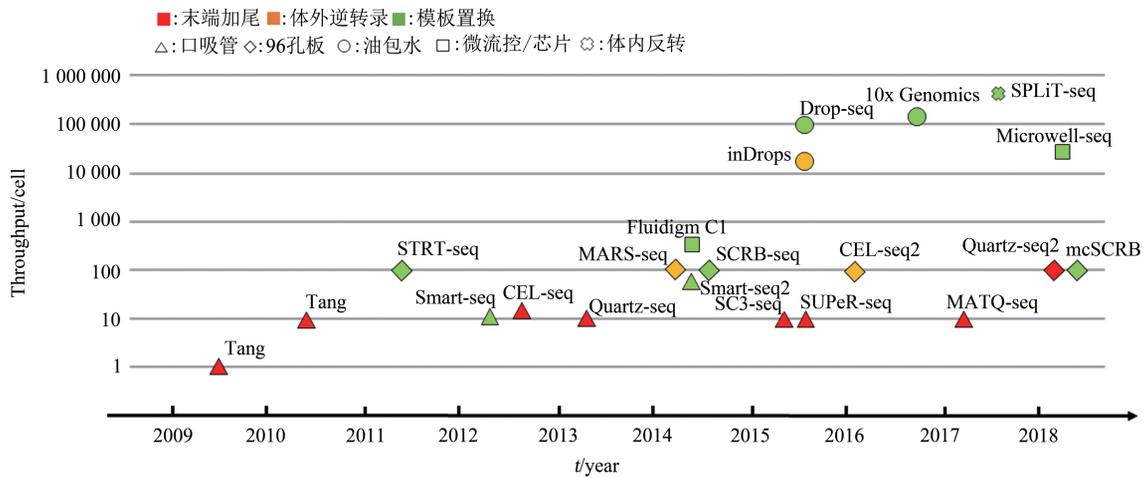


Fig. 1 The development of single cell RNA-seq

图1 单细胞转录组技术发展路径

1.2 单细胞转录组技术的高通量策略

单细胞中转录组建库的过程中, 通过在每一个反转引物或者 TSO 引物上加上细胞身份识别的 Barcode, 从而实现了大量细胞的同时建库, 提高了该技术的通量以及反应的重复性. 在该技术中, 单细胞隔离的模式决定了各方法的通量大小 (表 1). STAT-seq 方法中细胞 Barcode 加在 TSO 上, 并通过磁珠抓取 5' 端片段进行扩增建库. 当 Barcode 加在反转引物上, 获得的 cDNA 在其 3' 端具有细胞来源信息. 高通量单细胞建库主要是群体细胞的 cDNA 建库, 首先需要将 cDNA 打断, 由于只有 3' 端携带有细胞信息, 因此其他的 cDNA 片段将无法区分细胞来源信息. 因此, 高通量单细胞转录组主要是富集 3' 端建库.

2014 年, SCRB-seq^[16] (single-cell RNA barcoding sequencing) 作为高通量转录组方法, 基于较高通量的 96/384 孔板, 使用了含有 Barcode 和 UMI 的富集 3' 端策略. 2018 年, mcSCRB-seq^[17] 方法在反转体系中加入聚乙二醇, 并优化了各步反应条件, 提高了反应效率, 并获得了更多的基因信息. 以上两种方法均基于 96/384 孔板进行批量操作, 具有一定的通量, 但是要达到万级细胞数量则较为困难. Drop-seq^[18] 方法将前期的反应腔室换到了油包水滴中, 大大提高了通量. 在油包水滴中, 特制的带有 Barcode 和 UMI 的磁珠通过 T 碱基抓取 mRNA 的 polyA 尾巴, 磁珠上每条引物都抓取单个

独立细胞的 mRNA, 再将这些磁珠收集进行反转、模板置换和扩增建库. 虽然 Drop-seq 将细胞隔离在油包水的液滴中, 其通量可以达到上万个细胞, 但由于其需要额外搭建液滴泵系统, 成本较高. 2018 年, Microwell-seq^[19] 方法将磁珠抓取 mRNA 步骤放在了纳米级的微孔中, 该微孔板通过模具灌胶的方式获得, 通过梯度稀释方法使每个微孔包含一个细胞和一个磁珠, 在微孔中裂解细胞, 抓取其 mRNA, 该方法材料易获得, 且通量达到上万细胞. Indrop、Drop-seq 和 Microwell-seq 的磁珠均通过 Split-and-pool 的原理合成, 虽然该类磁珠的生产已经商业化, 但是价格较为昂贵. SPLiT-seq^[20] 方法将细胞固定打孔, 将每个固定后的细胞作为独立的反应腔室, mRNA 与 Barcode 在细胞内反转并进行 3 次 split-and-pool 过程, 给细胞加标签, 通过排列组合, 使每个细胞都有特异性的 Barcode 组合标签, 在提高通量的同时降低了费用. 单细胞转录组方法在高通量的发展过程中, 隔离腔室经历了从 96/384 孔板到油包水滴, 再到纳米微孔以及固定细胞内腔的发展, 在通量和可行性提高的同时, 成本也逐渐下降. 目前基于油包水的商业化高通量单细胞测序平台 10x Genomics, 由于其低成本的建库费用及标准化的操作和分析流程, 已被广泛推广.

目前实验室应用单细胞转录组技术, 主要依赖商业公司的 10x Chromium 平台或者 SMART-seq2 试剂盒, 技术建立门槛低, 同时通量高的单细胞转

录组技术亟待发展. 而单细胞转录组正在向更高通量发展, 但是高通量方法主要为富集3'端测序, 因此将丢失细胞中的部分信息. 就针对构建器官图谱这一目标, 均应用高通量的单细胞转录组技术, 并以10x Chromium系统为主. 然而由于成本的限制, 通常牺牲细胞基因的表达丰富度来获取更高的通量, 而当下的策略将综合两类技术来达到研究目

的. 即使用10x技术先进行表面探索性实验, 以提供细胞群内异质性的估计, 然后在较低的细胞数量上引导更深入的测序方法SMRT-seq2, 以捕获分析中的更多细节, 未来更高灵敏度的全长cDNA捕获技术加之三代测序的长读长优势, 将有可能得到单个细胞的完整信息.

Table 1 The principle and method of single-cell RNA-seq

表1 单细胞转录组测序原理及方法

方法	Barcode	UMI	捕捉RNA类型	cDNA富集位置	细胞隔离方式	评价
末端加尾原理						
Kurimoto2006	无	无	polyA (+)	3'端偏好的全长	单个PCR管	效率低
Tang2009	无	无	polyA (+)	3'端偏好的全长	单个PCR管	效率低
Quartz-seq	无	无	polyA (+)	3'端偏好的全长	单个PCR管	低副产物
SC3-seq	无	无	polyA (+)	3'端	单个PCR管	扩增偏倚
SUPeR-seq	无	无	polyA (+) (-)	全长	单个PCR管	捕获总RNA但通量低
MATQ-seq	无	无	polyA (+) (-)	全长	单个PCR管	灵敏度高
Quartz-seq2	有	有	polyA (+)	3'端	96/384孔板	通量高
体外逆转录原理						
CEL-seq	有	有	polyA (+)	3'端	96/384孔板	线性扩增
CEL-seq2	有	有	polyA (+)	3'端	96/384孔板	建库效率高
MARS-seq	有	有	polyA (+)	3'端	96/384孔板	提高通量
inDrops	有	有	polyA (+)	3'端	油包水滴	高通量设备要求高
模板置换原理						
STAT-seq	有	无	polyA (+)	5'端	96/384孔板	一定通量, 非线性扩增
SMART-seq	无	无	polyA (+)	全长	单个PCR管	灵敏度高通量低
SMART-seq2	无	无	polyA (+)	全长	单个PCR管	
SCRB-seq	有	有	polyA (+)	3'端	96/384孔板	一定通量和灵敏度
Drop-seq	有	有	polyA (+)	3'端	油包水滴	通量高, 步骤繁琐, 设备要求高
Microwell-seq	有	有	polyA (+)	3'端	纳米微孔	
SPLiT-seq	有	有	polyA (+)	3'端	固定细胞	

2 单细胞转录组技术在人类细胞图谱构建中的应用

2.1 正常组织细胞图谱

2.1.1 发育生物学图谱

从单个受精卵起始的个体发育过程是在严格的时间和空间序列中的精密调控过程, 理解发育过程的精确基因转录调控图谱, 以及寻找关键的驱动基因, 将有助于更好地理解器官发育的过程, 并推动细胞和组织工程的发展. 理解个体发育过程最直接的方法是对每个细胞的基因转录图谱进行测量, 根

据细胞之间的相互关系和空间位置, 构建出整体的三维调控关系. 然而, 由于技术的限制, 传统的转录组分析只能将发育过程的群体细胞作为研究对象, 所获得的基因表达数据是所有细胞的平均, 无法反映单个细胞的基因表达情况, 对细胞间的相互作用无法分析. 而单细胞转录组学技术的快速发展和应用, 不仅发现了胚胎发育最早期的一些关键基因事件, 而且构建了器官发育的高分辨率基因表达二维甚至三维图谱, 并确定不同发育阶段的每种细胞类型中关键的转录因子、生长因子和受体, 同时阐明细胞分化时各个细胞类型间的关系, 描述驱动发育的基因表达程序, 探索新的生物学现象.

利用单细胞转录组技术, 研究人员已经对胚胎发育早期的关键事件进行了详细的研究. 2016年, Ancelin等^[21]通过对受精卵早期合子的单细胞转录组测序, 讨论了KDM1A在确定受精卵受精后的表观遗传、保护基因组完整性和启动小鼠早期发育的基因组表达方面发挥着关键作用. 同年, Petropoulos等^[22]分离出处于8细胞阶段到着床前时间段(E3-E7)来自88个人类着床前胚胎的1 529个单细胞, 绘制了一个完整的人类胚胎着床前发育转录图谱, 揭示了细胞处于谱系特异性基因共表达的中间状态. Peng等^[23]对激光显微切割的单细胞进行转录组测量, 绘制了小鼠早期发育原肠运动中期(E7)胚胎三维分子图谱, 同时呈现小鼠细胞谱系建立过程中的空间转录组特征以及转录因子和信号通路调控网络. Scialdone等^[24]收集了胚胎期第6.5天(E6.5)的早期原肠胚形成时期, 到E7.75的原始血细胞产生时期的1 205个细胞, 进行单细胞转录组分析, 构建了健康哺乳动物早期发育阶段的基因表达图谱, 并研究了关键的造血转录因子Tall1的功能.

除了对早期胚胎发育进行单细胞转录组的图谱研究, 2017年Tsang等^[25]随机选取了剖宫产的正常胎盘和子痫前期患者胎盘的24 000个细胞, 进行微流体单细胞转录组分析, 获得了多种胎盘细胞亚群图谱, 重建了滋养层细胞的发育轨迹, 解决了妊娠过程中造血和胎盘细胞的纵向细胞动力学问题, 并能够无创发现早期子痫前期胎盘绒毛外滋养细胞的细胞功能障碍. 2018年, Vento-Tormo等^[26]分析了来自妊娠早期胎盘的约70 000个单细胞的转录组, 其具有匹配的母体血液和蜕膜细胞. 揭示了位于不同蜕膜层的基质和血管周围细胞亚群, 以及具有独特的免疫调节和趋化因子谱的蜕膜自然杀伤细胞有三个主要子集, 构建了早期母胎界面的细胞图谱.

此外, 在单细胞尺度下对器官发育的基因转录图谱描绘也成为研究的热点. 2016年, Bakken等^[27]结合激光显微切割和单细胞转录组技术, 绘制了恒河猴从胚胎发育至成年的大脑转录组全图. 2018年, Gao等^[28]采集了人类6到25周胚胎中食道、胃、小肠和大肠以及成年人的大肠总计5 277个单细胞, 通过单细胞转录组分析揭示了这4种器官的精准发育路径, 基因表达特征以及重要的生物

学特性, 包括发育过程、信号通路、细胞周期、营养物质消化吸收代谢和转录因子网络. 同年该研究组^[29]还对人类胚胎期7周到25周肾脏的3 000多个单细胞进行了转录组测序分析, 探索了肾单位形成的前体细胞的异质性, 获得了肾小球系膜的代谢基因及细胞外基质组成.

2.1.2 成熟器官图谱

人体细胞图谱的构建是理解器官功能的基础. 传统的人体细胞图谱构建是基于各类的显微观察, 近年单细胞转录组技术的发展, 推动了2017年10月人类细胞图谱计划的启动, 使得在单细胞水平了解复杂器官的组成、功能和病变成为可能. 近期来自15个机构的多学科研究人员对大脑的发育分子机制以及自闭症、精神分裂症、双向情感障碍等神经精神疾病与基因的关联进行了深入分析^[30-39], 成功绘制了人脑基因组图谱. 2018年, Rhie等^[40]通过对13 000个单细胞转录组的聚类分析绘制了肾小球细胞基因表达图谱, 并鉴定了三种已知的肾小球细胞类型. 同年Rheaume等^[41]对6万多个神经节细胞的单细胞转录组分析, 聚类得到了40个神经节细胞亚群, 探讨了神经节细胞亚群的功能、发育进程, 同时鉴定了神经节视网膜细胞的细胞特异性基因.

2.2 疾病生物学图谱

单细胞转录组技术不仅提供了正常器官的细胞图谱、发现新的细胞类型, 同样单细胞转录组技术也推动了对于器官疾病状态细胞图谱的全局理解. 例如, 2018年Ordovas-Montanes^[42]对原发性慢性鼻窦炎患者的手术样本(18 036个细胞, $n=12$)进行大规模平行单细胞RNA测序的疾病谱分析, 获得了II型炎症疾病中人类呼吸系统上皮细胞、免疫细胞和基质细胞类型及亚群的转录组图谱. Liu等^[43]借助单细胞转录组测序对不同损伤阶段的巨噬细胞和内皮细胞进行分析, 揭示了巨噬细胞在血管修复中的作用.

单细胞转录组技术对于癌组织的细胞图谱构建, 推动了对于癌症发生发展的理解, 以及癌细胞和微环境互动的理解. 癌组织是一个高度异质细胞的组合, 除了癌细胞本身具有高度的异质性, 微环境中的成纤维细胞、巨噬细胞、淋巴细胞、内皮细胞等也存在相当的异质性. 对于癌组织异质性的理解是理解癌细胞功能、演化、转移等重要问题的关

键. 单细胞转录组技术的发展和运用, 以及和其他单细胞组学技术的结合, 使得从单细胞尺度对癌细胞、微环境细胞以及细胞相互作用的理解成为可能, 并对理解癌细胞产生的抗药性和后续治疗有重要参考意义.

2016年Tirosh等^[44]对19名黑色素瘤患者的肿瘤细胞, 包括癌细胞、免疫细胞、间质细胞和内皮细胞, 进行单细胞转录组分析, 发现每个患者的恶性细胞都是聚类在一起的, 来自于同一肿块的癌细胞呈现出的转录特异性与细胞周期、空间组织和耐药程序有关. 这表明癌细胞间存在高度的异质性. 此外该研究探讨了黑色素瘤微环境中各种细胞类型如免疫细胞、成纤维细胞和内皮细胞在个体之间的异质性. 2018年Lambrechts等^[45]呈现了人肺癌微环境52 698个细胞的转录组信息, 并独立验证了40 250个肺癌细胞, 通过与匹配的非恶性肺部样本进行比较, 共鉴定了52种基质细胞亚型, 绘制了完整的肺癌细胞图谱, 其中包括迄今仍被认为是同质细胞类型中的新亚群, 以及其异质性的转录因子. Bian等^[46]利用单细胞多组学方法, 从单细胞分辨率、多组学水平深入解析了人类结直肠癌在发生和转移过程中, 基因组拷贝数变异、DNA甲基化异常及基因表达改变的特点和相互关系.

2018年, Zhang等^[47-48]多篇报道研究了肿瘤免疫微环境图谱. 他们对14例非小细胞癌的12 346个T细胞做单细胞转录组测序分析. 这些细胞来自患者的外周血、癌组织及癌旁组织. 在单细胞水平绘制了与非小细胞肺癌相关的T细胞免疫图谱. 在另一项研究中, 对来自12例结直肠癌初治患者的外周血、癌及癌旁组织的大量T细胞进行了单细胞全长转录组测序和分析, 描述了结直肠癌相关T细胞的分布、克隆、迁移、转化等特点并揭示了结直肠癌中20类不同功能特性的T细胞亚类的动态变化关系. 此外, 该研究还报道了从6例肝细胞癌(HCC)患者取血液、癌旁正常组织、肿瘤组织和癌旁与肿瘤交界处组织中分离得到的T细胞样品, 并在单细胞水平对其转录组及受体(TCR)序列进行了综合分析, 通过探索不同亚群之间的关系, 鉴定每个亚群特异的基因表达, 揭示了肿瘤中的T细胞在功能、分布和发展状况方面和其他部位的T细胞的异质性, 并揭示了肿瘤细胞逃逸免疫监视的缘由.

3 展望和前瞻

近年来, 单细胞转录组技术已经在提高细胞分析通量, 基因表达定量测量和提高测量灵敏度上有了质的飞跃, 但是目前大多数单细胞转录组研究都集中于检测溶液里的单细胞. 对于组织样本, 包括低温保存和固定的样本, 如果能在结合细胞空间位置基础上, 进行单细胞转录组研究, 将进一步为组织和器官的三维转录组图谱构建打开新的大门. 同时, 不少组织和器官里的RNA酶含量很高, 在进行组织原位的单细胞转录组研究中存在RNA容易降解、获得的单细胞转录组数据质量低等问题, 仍需要进一步探索新的通用方法. 此外由于单细胞转录组技术成本的不断下降, 细胞通量可以达到 10^4 级, 如何在高通量情况下依然保持高灵敏度和高效的转录组捕获也值得继续探索. 同时对大量单细胞采集到的如此庞大的数据集, 加之人类细胞图谱计划的最终完成, 如何发展生物信息工具, 挖掘关键信息和充分整合和利用这些数据集也是未来发展的关键. 这一系列单细胞转录组学技术问题的解决, 以及和其他技术的结合, 将极大推动人类对生命科学各领域的理解以及对疾病的防治.

参 考 文 献

- [1] Kurimoto K, Yabuta Y, Ohinata Y, *et al.* An improved single-cell cDNA amplification method for efficient high-density oligonucleotide microarray analysis. *Nucleic Acids Res*, 2006, **34**(5): e42
- [2] Tang F, Barbacioru C, Wang Y, *et al.* mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods*, 2009, **6**(5): 377-382
- [3] Sasagawa Y, Nikaido I, Hayashi T, *et al.* Quartz-Seq: a highly reproducible and sensitive single-cell RNA sequencing method, reveals non-genetic gene-expression heterogeneity. *Genome Biol*, 2013, **14**(4): R31
- [4] Nakamura T, Yabuta Y, Okamoto I, *et al.* SC3-seq: a method for highly parallel and quantitative measurement of single-cell gene expression. *Nucleic Acids Res*, 2015, **43**(9): e60
- [5] Kivioja T, Vaharautio A, Karlsson K, *et al.* Counting absolute numbers of molecules using unique molecular identifiers. *Nat Methods*, 2011, **9**(1): 72-74
- [6] Fan X, Zhang X, Wu X, *et al.* Single-cell RNA-seq transcriptome analysis of linear and circular RNAs in mouse preimplantation embryos. *Genome Biol*, 2015, **16**: 148
- [7] Sheng K, Cao W, Niu Y, *et al.* Effective detection of variation in single-cell transcriptomes using MATQ-seq. *Nat Methods*, 2017,

- 14(3):267-270
- [8] Picelli S, Faridani O R, Bjorklund A K, *et al.* Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc*, 2014, **9**(1): 171-181
- [9] Sasagawa Y, Danno H, Takada H, *et al.* Quartz-Seq2: a high-throughput single-cell RNA-sequencing method that effectively uses limited sequence reads. *Genome Biol*, 2018, **19**(1): 29
- [10] Hashimshony T, Wagner F, Sher N, *et al.* CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. *Cell Rep*, 2012, **2**(3): 666-673
- [11] Jaitin DA, Kenigsberg E, Keren-Shaul H, *et al.* Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science*, 2014, **343**(6172): 776-779
- [12] Hashimshony T, Senderovich N, Avital G, *et al.* CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. *Genome Biol*, 2016, **17**: 77
- [13] Klein A M, Mazutis L, Akartuna I, *et al.* Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell*, 2015, **161**(5): 1187-1201
- [14] Islam S, Kjallquist U, Moliner A, *et al.* Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. *Genome Res*, 2011, **21**(7): 1160-1167
- [15] Ramskold D, Luo S, Wang Y C, *et al.* Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat Biotechnol*, 2012, **30**(8): 777-782
- [16] Soumillon M, Cacchiarelli D, Semrau S, *et al.* Characterization of directed differentiation by high-throughput single-cell RNA-Seq [M]. bioRxiv. doi: <http://doi.org/10.1101/003236>
- [17] Bagnoli J W, Ziegenhain C, Janjic A, *et al.* Sensitive and powerful single-cell RNA sequencing using mcSCR-seq. *Nat Commun*, 2018, **9**(1): 2937
- [18] Macosko E Z, Basu A, Satija R, *et al.* Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell*, 2015, **161**(5): 1202-1214
- [19] Han X, Wang R, Zhou Y, *et al.* Mapping the mouse cell atlas by microwell-Seq. *Cell*, 2018, **172**(5): 1091-1107
- [20] 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. *Science*, 2018, **360**(6385): 176-182
- [21] Ancelin K, Syx L, Borensztein M, *et al.* Maternal LSD1/KDM1A is an essential regulator of chromatin and transcription landscapes during zygotic genome activation. *Elife*, 2016, **5**. pii: e08851
- [22] Petropoulos S, Edsgard D, Reinius B, *et al.* Single-cell RNA-Seq reveals lineage and X chromosome dynamics in human preimplantation embryos. *Cell*, 2016, **165**(4): 1012-1026
- [23] Peng G, Suo S, Chen J, *et al.* Spatial Transcriptome for the molecular annotation of lineage fates and cell identity in mid-gastrula mouse embryo. *Developmental Cell*, 2016, **36**(6): 681-697
- [24] Scialdone A, Tanaka Y, Jawaid W, *et al.* Resolving early mesoderm diversification through single-cell expression profiling. *Nature*, 2016, **535**(7611): 289-293
- [25] Tsang J C H, Vong J S L, Ji L, *et al.* Integrative single-cell and cell-free plasma RNA transcriptomics elucidates placental cellular dynamics. *Proc Natl Acad Sci USA*, 2017, **114**(37): E7786-E7795
- [26] Vento-Tormo R, Efremova M, Botting R A, *et al.* Single-cell reconstruction of the early maternal-fetal interface in humans. *Nature*, 2018, **563**(7731): 347-353
- [27] Bakken T E, Miller J A, Ding S L, *et al.* A comprehensive transcriptional map of primate brain development. *Nature*, 2016, **535**(7612): 367-375
- [28] Gao S, Yan L, Wang R, *et al.* Tracing the temporal-spatial transcriptome landscapes of the human fetal digestive tract using single-cell RNA-sequencing. *Nature Cell Biology*, 2018, **20**(6): 721-734
- [29] Wang P, Chen Y, Yong J, *et al.* Dissecting the global dynamic molecular profiles of human fetal kidney development by single-cell RNA sequencing. *Cell Rep*, 2018, **24**(13): 3554-3567
- [30] Rajarajan P, Borrmann T, Liao W, *et al.* Neuron-specific signatures in the chromosomal connectome associated with schizophrenia risk. *Science*, 2018, **362**(6420). pii: eaat4311
- [31] An J Y, Lin K, Zhu L, *et al.* Genome-wide *de novo* risk score implicates promoter variation in autism spectrum disorder. *Science*, 2018, **362**(6420). pii: eaat6576
- [32] Amiri A, Coppola G, Scuderi S, *et al.* Transcriptome and epigenome landscape of human cortical development modeled in organoids. *Science*, 2018, **362**(6420). pii: eaat6720
- [33] Wang D, Liu S, Warrell J, *et al.* Comprehensive functional genomic resource and integrative model for the human brain. *Science*, 2018, **362**(6420). pii: eaat8464
- [34] Li M, Santpere G, Imamura Kawasawa Y, *et al.* Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science*, 2018, **362**(6420). pii: eaat7615
- [35] Zhu Y, Sousa A M M, Gao T, *et al.* Spatiotemporal transcriptomic divergence across human and macaque brain development. *Science*, 2018, **362**(6420). pii: eaat8077
- [36] Gandal M J, Zhang P, Hadjimichael E, *et al.* Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. *Science*, 2018, **362**(6420). pii: eaat8127
- [37] Chen C, Meng Q, Xia Y, *et al.* The transcription factor POU3F2 regulates a gene coexpression network in brain tissue from patients with psychiatric disorders. *Science Translational Medicine*, 2018, **10**(472). pii: eaat8178
- [38] Meng Q, Wang K, Brunetti T, *et al.* The DGCR5 long noncoding RNA may regulate expression of several schizophrenia-related genes. *Science Translational Medicine*, 2018, **10**(472). pii: eaat6912
- [39] Rhie S K, Schreiner S, Witt H, *et al.* Using 3D epigenomic maps of primary olfactory neuronal cells from living individuals to understand gene regulation. *Science Advances*, 2018, **4**(12): eaav8550
- [40] Karaiskos N, Rahmatollahi M, Boltengagen A, *et al.* A single-cell transcriptome atlas of the mouse glomerulus. *Journal of the American Society of Nephrology: JASN*, 2018, **29**(8): 2060-2068
- [41] Rheaume B A, Jereen A, Bolisetty M, *et al.* Single cell transcriptome profiling of retinal ganglion cells identifies cellular

- subtypes. *Nat Commun*, 2018, **9**(1): 2759
- [42] Ordovas-Montanes J, Dwyer D F, Nyquist S K, *et al.* Allergic inflammatory memory in human respiratory epithelial progenitor cells. *Nature*, 2018, **560**(7720): 649-654
- [43] Liu C, Wu C, Yang Q, *et al.* Macrophages mediate the repair of brain vascular rupture through direct physical adhesion and mechanical traction. *Immunity*, 2016, **44**(5): 1162-1176
- [44] Tirosh I, Izar B, Prakadan S M, *et al.* Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science*, 2016, **352**(6282): 189-196
- [45] Lambrechts D, Wauters E, Boeckx B, *et al.* Phenotype molding of stromal cells in the lung tumor microenvironment. *Nature Medicine*, 2018, **24**(8): 1277-1289
- [46] Bian S, Hou Y, Zhou X, *et al.* Single-cell multiomics sequencing and analyses of human colorectal cancer. *Science*, 2018, **362**(6418): 1060-1063
- [47] Guo X, Zhang Y, Zheng L, *et al.* Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. *Nature Medicine*, 2018, **24**(7): 978-985
- [48] Zhang L, Yu X, Zheng L, *et al.* Lineage tracking reveals dynamic relationships of T cells in colorectal cancer. *Nature*, 2018, **564**(7735): 268-272

Recent Progress of Single-cell RNA-seq and Its Application in Human Cell Atlas Construction*

NI Jian¹⁾, Hu Miao-Miao¹⁾, WEI Ying¹⁾, CHEN Rui²⁾, GUO Yan^{1)**}

¹⁾School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China;

²⁾Department of Plastic Surgery, Renji Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200240, China)

Abstract Single-cell RNA-seq provides information on the differences in gene expression, making it possible to study the cells and the mechanism of their interactions at the single-cell level. In recent years, single-cell transcriptome technology has experienced the development at the principle of cDNA amplification from homopolymer tailing-based, *in vitro* reverse transcription to template-switch, and greatly improved the sensitivity. At the same time, cell isolation is promoted from 96/384 well plates, droplets to nano-pores, which greatly reduced the experimental cost while increasing the throughput. The single-cell RNA-seq, with its advantages in cell heterogeneity and cell classification, has promoted the construction of cell atlas from developmental biology to normal tissue and pathological cell biology. This paper reviews the recent progress of single-cell RNA-seq and its application in human cell atlas.

Key words single cell, transcriptome, heterogeneity, cell atlas

DOI: 10.16476/j.pibb.2019.0033

* This work was supported by The National Research and Development Program (2018YFC1003501), The National Natural Science Foundation of China (91529302) and The National Natural Science Foundation of Youth Program (31700762).

** Corresponding author.

Tel: 86-21-34206632, E-mail: yanguo@sjtu.edu.cn

Received: February 25, 2019 Accepted: June 18, 2019