



空间转录组技术研究进展*

危莹** 张小丹** 胡苗苗 吴忠琴 程 酩 郭 妍***

(上海交通大学生物医学工程学院, 上海 200240)

摘要 空间转录组技术旨在对细胞的基因表达进行定量测量, 同时提供细胞在组织空间的具体位置信息。与传统的转录组技术相比, 空间转录组技术能获得细胞在组织生理环境下真实的基因表达特征, 以及与微环境的关系, 进一步推进对正常和病理状态下细胞特性的理解。近年来, 空间转录组技术的发展取得了重要的进展, 检测的细胞通量、转录本数量和质量不断提高, 空间定位信息更加准确全面。本文总结了四大类不同空间转录组技术的特点和应用, 并比较了目前不同空间转录组技术的优缺点。

关键词 空间转录组, 组学技术, 空间位置, 单细胞

中图分类号 Q33

DOI: 10.16476/j.pibb.2021.0006

细胞作为机体的基本单元, 在特定的空间位置与微环境协同, 发挥其特有的生物学功能。生物学研究的重要目标是理解正常和病理状态下细胞的特点和功能。近十几年转录组技术飞速发展, 已可从全基因组尺度对基因表达进行定量测量, 极大地推进了研究人员对细胞功能的理解。转录组技术的发展主要经历了三个重要的阶段。第一个阶段是对大量混合细胞进行转录组测序 (bulk transcriptome)^[1], 该方法获得的是大量细胞的基因平均表达水平, 虽然推进了对细胞群体的特性认识, 但无法获知群体内特定细胞的基因表达情况。第二个阶段是对单细胞进行转录组测序 (single cell transcriptome), 该方法将对细胞基因表达的认识推进到单细胞的层面, 目前已在发现新细胞类型、揭示细胞异质性等方面展现出优势。但是目前的单细胞转录组测序方法仍存在一定局限性, 由于单细胞悬液的获得需要对组织进行酶解, 悬液中的单细胞失去了组织空间位置信息, 而且已有研究表明, 有些单细胞在酶解过程中或离开特定微环境后, 基因表达谱会发生变化^[2]。除此以外, 某些特殊形态的细胞, 如大脑中结构复杂的神经元和神经胶质细胞, 难以通过组织酶解获取。近年来, 转录组研究进入空间转录组 (spatial transcriptome) 阶段, 该技术可以同时获得细胞的空间位置信息和

基因表达数据, 进一步推进了对组织原位细胞真实基因表达的研究, 为组织细胞功能、微环境互作、发育过程谱系追踪、疾病病理学等多个领域提供了重要的研究手段。多种新的空间转录组技术得到快速发展, 成为生物技术研究领域的前沿热点, 本文就近年来空间转录组技术的发展和應用进行综述。

1 空间转录组技术的发展

1.1 基于原位杂交的空间转录组技术

1.1.1 单分子荧光原位杂交 (smFISH)

空间转录组技术最早可追溯到单分子荧光原位杂交 (single molecule fluorescence *in situ* hybridization, smFISH) 技术^[3]。原位杂交技术于1960年代产生, Gall等^[4]将细胞经过碱处理使得DNA变性, 而后与放射性同位素标记的RNA探针一同孵育, 利用放射自显影技术检测了青蛙卵母细胞中高度扩增的核糖体DNA。随后, Bauman等^[5]于1980年将3'端带有荧光基团标记的RNA用作特

* 国家重点研发计划(2018YFC1003500)和国家自然科学基金(81972909)资助项目。

** 并列第一作者。

*** 通讯联系人。

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

收稿日期: 2021-04-20, 接受日期: 2021-07-13

定DNA序列的探针,首次实现了荧光原位杂交(fluorescence *in situ* hybridization, FISH)技术的应用。FISH对识别目标DNA或RNA具有高分辨率、高灵敏度和高特异性等特点,可直接应用于中期染色体和间期细胞核的目标序列检测^[6-7]。除此以外,FISH还可以在单细胞水平上根据可视化的杂交信号区分不同的转录本^[8]。1998年,Femino等^[9]将数字成像显微镜应用于FISH,并且合成了带有5个荧光基团的寡核苷酸探针,校准了单个探针发出的光,使得检测单个RNA分子成为可能。smFISH一般通过多个标记有荧光团的20 bp短寡核苷酸探针来进行(通常每个探针都标记有一个荧光团),这些荧光探针在组织原位与特定的mRNA进行特异性杂交(图1a),以此反映细胞内转录本丰度并同时进行空间定位^[10]。

1.1.2 组合标记

由于smFISH能够同时使用的荧光团种类有限,一次识别的转录本数量很少,通常为3或4个,因此smFISH还不能实现基因组范围的大通量分析^[4]。2002年,Levsky等^[11]发现可以使用组合标记来增加可识别的转录本数量,并且可以减少单色假阳性信号的出现。实验中,通过将探针用不同条形码(barcode)标记来进行分组,分别与4个荧光基团相结合,在显微镜下识别荧光团的颜色。之后通过检测算法进行图像处理,根据所识别的不同颜色下的像素点的强度,分析出该位点是哪些颜色的结合,以此分析出不同的转录本。当使用的荧光基团的种类为 n 时,可鉴定的转录本的数量为 2^n-1 个。Jakt等^[12]将组合标记方法应用到体外分化的小鼠胚胎干细胞中,结果表明组合标记可定量单个细胞的多个转录本,还可以追踪细胞分化过程中细胞身份的变化。虽然组合标记方法增加了转录本的检测数量,但是对于基因组范围的分析还是远远不够,而且由于受到组织自发荧光的影响,组合标记的方法应用于组织切片还有一定的困难。

2012年,Lubeck等^[13]通过结合超分辨显微镜(super-resolution microscopy, SRM)和组合标记,识别单个mRNA上不同荧光寡核苷酸探针的空间排序,大幅提高了检测的转录本数量。SRM的分辨率高达10~20 nm,20个寡核苷酸探针长度约为7 nm,转录本长度约为100 nm,在SRM下可以看到不同荧光寡核苷酸探针的空间排序,因此可解析出某个感兴趣的mRNA分子的序列。在SRM的帮助下,可以根据不同颜色荧光团的排列组合来区分

转录本,例如某红、黄、蓝顺序的转录本与蓝、黄、红顺序的转录本是不同的转录本(图1b),因此在可使用的荧光团数量有限的情况下,可以通过不同的荧光团顺序来鉴定多个转录本。

1.1.3 顺序杂交

虽然组合标记增加了可识别的转录本数量,但是还无法检测全基因组的转录本。2014年,Lubeck等^[14]开发了顺序杂交技术。由于固定的细胞中转录本的位置不会发生改变,即使通过多轮杂交,相同的荧光点还是可以保留在原位,所以能通过顺序杂交进行单细胞原位的RNA分析。该技术在每一轮杂交中,使用一组标记了单一类型荧光团的探针,结合转录本后进行成像。之后用DNA酶处理,清除杂交链上结合的探针,再将转录本与另一组不同荧光染料标记的、但序列相同的探针杂交(图1c)。理论上,这种方法可以检测 F^n 个不同的转录本(F =荧光团数量, n =杂交回合)^[15]。因此,原则上,4种荧光染料和8个杂交回合可以检测一个细胞中的所有转录本。这种方法的好处在于即使只有2种荧光染料,也可以通过多次杂交使得检测规模迅速扩大。但是,如果想要将这种方法的应用范围扩大到整个转录组,价格较昂贵。除此以外,虽然顺序杂交可检测的转录本数量随杂交回合的增加而呈指数增长,但同时检测误差也呈指数增长。

1.1.4 MERFISH

为了克服检测误差,2015年,Chen等^[16]开发了一种名为多重错误鲁棒荧光原位杂交(multiplexed error-robust FISH, MERFISH)的方法。该方法利用二进制条形码,在检测时会将单个检测错误进行自动校正,只有当检测错误为多个时,才会识别为另一个转录本。除此以外,该方法还结合了组合标记、顺序杂交和成像等方法,可检测单个细胞中数千种RNA的拷贝数和空间定位。Moffitt等^[17]通过对MERFISH进行一系列改进,包括使用化学裂解代替光漂白来去除连续两轮smFISH成像之间的荧光信号,增加成像视野并使用多色成像。通过这些改进,在一次18 h的测量中就检测了多达40 000个细胞^[18]。除此以外,Wang等^[19]开发了一种结合MERFISH和扩展显微镜(expansion microscopy)的方法,扩展显微镜可以有效扩大相邻分子之间的距离,显著增加了可测量的RNA的总密度,使得检测效率提高近1倍,单日可测量数以万计的细胞,可测到的转录本密度是以前的10倍以上,并且可与免疫荧光技术相结合。

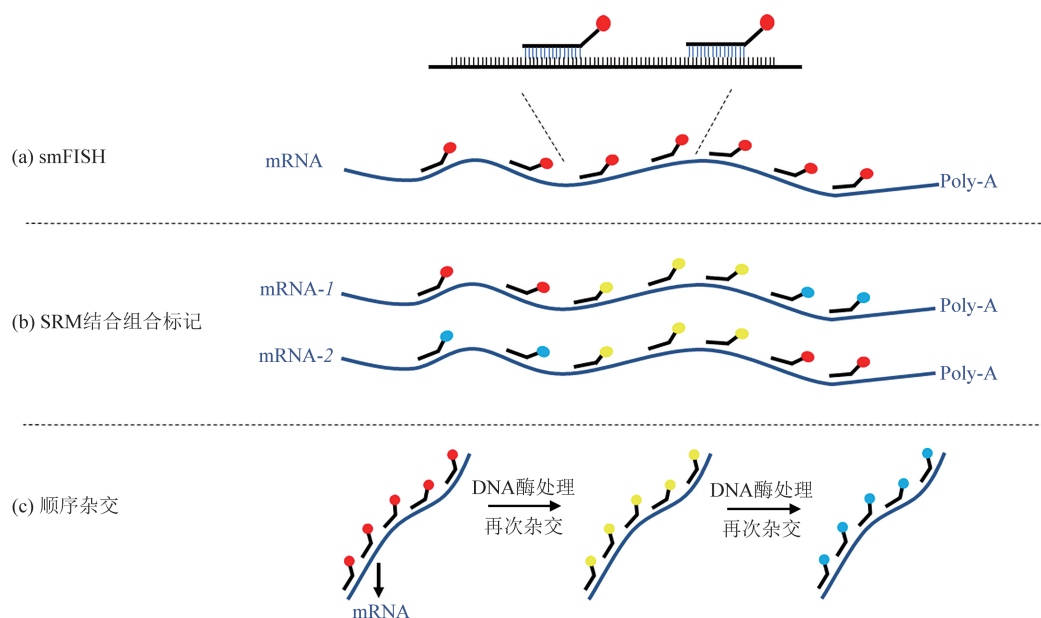


Fig. 1 The principle of spatial transcriptome technologies based on *in situ* hybridization

图1 基于原位杂交的空间转录组技术原理

(a) smFISH: 通过多个标记有荧光团的20 bp的短寡核苷酸探针在组织原位与特定的mRNA进行特异性杂交。(b) SRM结合组合标记: 根据不同颜色荧光团的排列组合来区分转录本, 例如某红、黄、蓝顺序的转录本与蓝、黄、红顺序的转录本是不同的转录本。(c) 顺序杂交: 每一轮杂交过程中, 使用一组标记有单一类型荧光团的探针结合每个转录本后进行成像, 之后用DNA酶处理去除杂交回合之间的结合探针, 再将转录本与不同荧光染料标记的相同探针杂交。

这些改进大大扩展了MERFISH可解决的生物学问题的范畴。

总而言之, smFISH及其衍生方法可应用于组织切片和体外培养细胞的基因表达测量。该类技术已成功应用于细胞间异质性的研究^[20]和疟原虫感染的红细胞基因表达的研究^[21], 还应用于一些模式动物如果蝇、秀丽隐杆线虫和斑马鱼等的相关研究中^[22]。但是该方法在组织中的应用效果不如在培养细胞中的应用效果好, 这是由于组织中的应用受到组织自发荧光的影响。而且随着复用探针次数的增加, 组织背景的信号也增加, 同时背景强度也随着组织复杂程度的增加而增加。Moffitt等^[23]开发了一种降低组织背景的smFISH方法, smFISH实验操作过程中背景的主要来源是探针与非RNA的细胞成分(例如蛋白质)非特异性结合。该方法的关键在于通过将RNA分子锚定在不可膨胀的聚丙烯酰胺基质上, 再去除不需要的非RNA成分(蛋白质和脂质等)。研究人员在成年小鼠下丘脑的冰凍切片上使用MERFISH方法对130种RNA进行检测, 所得到的成像结果证明, 该方法大大降低了成像时的组织背景干扰信号, 能够清晰地观察到代

表单个RNA分子的荧光点。

1.2 基于高通量测序 (high-throughput sequencing) 的空间转录组技术

1.2.1 LCM-RNA-seq技术

1996年, Emmert等^[24]开发了激光捕获显微切割技术 (laser capture microdissection, LCM)。该技术可以在保留空间位置信息的前提下, 可视化地从切片中快速且准确地捕获目标细胞^[25], 现已被广泛用于生物学研究中。该技术原理是将组织切片贴在特殊薄膜覆盖的玻璃载玻片上, 组织切片经过染色后, 在显微镜下可视化地进行特定细胞的选择、激光切割和分选, 收集组织中特定的细胞样品。分选获得的细胞可以进行特定基因表达的分析, 结合基因芯片或测序技术(图2a), 对分选细胞的基因表达谱进行解析^[26]。该技术分选的细胞可来源于活细胞培养物、新鲜冷冻组织和福尔马林固定的石蜡包埋组织等。

目前基于激光显微切割的空间转录组技术, 已应用于非常少量的细胞, 甚至单细胞的分选。2016年, Peng等^[27]通过整合与优化LCM和单细胞转录组测序 (single cell RNA-seq) 技术, 构建了一

种能解析少量细胞转录组信息，同时保留细胞原有位置信息的方法。通过LCM分区捕获，可以得到小鼠胚胎E7.0时期的转录组信息，以及各个区域的标志基因，将这些标志基因作为“邮政编码(zipcode)”，用于推断空间坐标，从而将未知来源的胚胎细胞“映射”到体内胚胎的特定位置。

最近，Moor等^[28]利用基于激光显微切割的空间转录组技术获得组织空间区域特异性的标志基因，通过标志基因的表达将单细胞转录组数据定位到组织空间，获得单细胞分辨率的空间转录组信息。他们使用激光显微切割技术将小鼠小肠绒毛底部到顶部之间5个等间隔的区域进行切割，分别捕获各个区域的上皮细胞并进行转录组测序。他们通过差异表达分析获得每个区域的肠上皮细胞标志性基因，建立不同空间的基因坐标，并利用这些坐标将单细胞“映射”到小肠绒毛的对应区域。该方法也在其他器官的空间单细胞转录组研究中得到成功应用，如肝小叶组织^[29]。

目前基于激光显微切割的空间转录组技术已在肿瘤生物学、发育生物学以及各类疾病生物学研究领域发挥积极的作用。通过激光显微切割获得细胞的具体空间位置，根据研究目的将分选的细胞构建不同的测序文库，通过一次高通量转录组测序，就

能获得大量已知和未知基因的表达信息，甚至可以得到可变剪切的基因信息。随着测序成本的不断降低，该方法用于研究空间转录组具有很高的性价比。

1.2.2 HDST和Slide-Seq

2019年Vickovic等^[30]提出了高分辨空间转录组(high-definition spatial transcriptomics, HDST)测序方法，通过在载玻片表面雕刻2.05 μm 的大量小孔，将直径为2 μm 的硅胶磁珠分配到这些孔里，通过在磁珠表面连接带有特定barcode序列以及多种唯一分子标识符(unique molecular identifiers, UMIs)和poly-dT的寡核苷酸链，来捕获对应位置细胞的mRNA进行反转录，并进行文库构建和转录组测序。同一时期，Rodriques等^[31]发明了类似的Slide-Seq方法，该方法用带已知barcode修饰的直径为10 μm 磁珠覆盖的玻片(磁珠同时连有不同的UMIs以及poly-dT)，捕获带poly(A)尾的mRNA，进行cDNA合成后建库测序(图2b)。这两种方法通过结合测序数据和图像处理，都可以将基因表达情况可视化地比对到原始的组织切片中，得到非常直观的结果，通量大且分辨率高。但其存在一定比例的磁珠会同时比对到两种细胞类型的情况，且磁珠的捕获效率仍然有待提高。

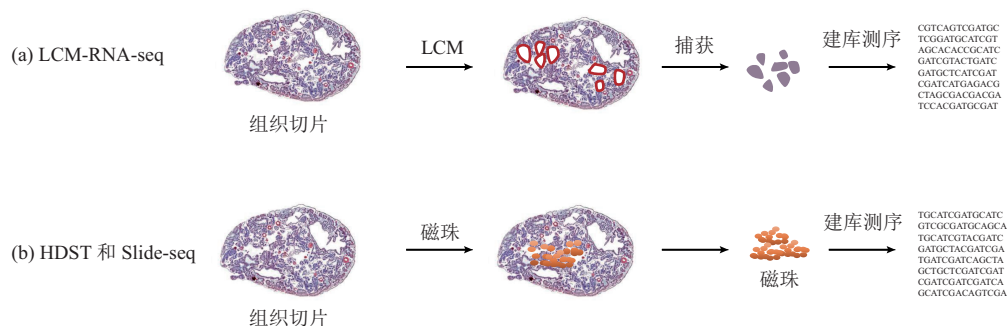


Fig. 2 The principles of space transcriptome technologies based on high-throughput sequencing

图2 基于高通量转录组测序的空间转录组技术原理

(a) LCM-RNA-seq: 将组织切片贴在特殊薄膜覆盖的玻璃载玻片上，染色或者细胞标记后将玻片在显微镜下可视化地进行激光切割和分选，收集组织中特定的细胞样品进行建库测序，对分选细胞的基因表达谱进行解析。(b) HDST和Slide-seq: 将组织切片贴在磁珠覆盖的玻片上，磁珠带有多种UMIs和poly-dT来捕获对应位置细胞的mRNA进行反转录，之后收集磁珠可进行文库构建和转录组测序。

1.2.3 10× Genomics Visium空间转录组测序技术

2019年，10× Genomics公司推出了Visium空间转录组技术，结合显微成像技术和RNA测序技术，将染色后的组织切片与不同空间位置细胞的mRNA表达数据相结合。利用Visium技术，可以

从整个组织切片获得高通量的转录组数据，研究人员可以通过此技术研究正常和疾病组织的细胞组成及其与基因表达的关系，可用于发现新的组织生物标志物。基于此技术，研究人员还开发了整合可视化基因表达和形态学数据的软件，可以在一个工作

日之内完成从对样品切片到建好测序文库。

Visium 技术的原理是利用载有 poly(A) 和空间 barcode 标记的载玻片, 原位捕获对应细胞的 mRNA 并进行反转录, 通过高通量测序结合原位的 barcode 信息, 绘制复杂组织样品的空间基因表达图谱。该技术所使用的一张玻片可容纳 4 片组织切片, 将组织切片贴在具有 5 000 个空间捕获区域的载玻片上, 每个空间捕获区域中包含数百万个含有 poly-dT 的寡核苷酸链, 这些寡核苷酸链可以捕获 mRNA, 寡核苷酸链连接有唯一的分子标识符 UMI (计数 mRNA 分子) 和 barcode (每个 barcode 对应特定的捕获区域)。每个空间区域的直径为 55 μm , 可以检测到 1~10 个细胞^[32-33]。Visium 技术的操作流程是首先将样品进行冰冻切片, 贴在载玻片上的捕获区域。利用标准的固定和染色技术对切片进行染色, 然后对切片进行透化, 释放 mRNA, 让其与空间区域中带有 poly-dT 序列的引物进行结合, 然后合成 cDNA。之后, 将每个空间区域反转形成的 cDNA 合并回收, 进行最终文库的构建以及测序, 通过数据分析软件即可获得细胞在组织原始的空间位置所发生的基因转录情况^[34]。

2019 年, Maniatis 等^[35] 使用 10 $\times$ Genomics Visium 空间转录组技术, 研究肌萎缩侧索硬化 (amyotrophic lateral sclerosis, ALS) 模型小鼠以及 ALS 患者死后脊髓组织的基因表达。ALS 导致的瘫痪是由于运动神经元退化造成骨骼肌神经支配失调, 但是对完整的脊髓组织中分子发生过程的时空顺序仍然知之甚少。该团队通过 10 $\times$ Genomics Visium 空间转录组技术表征了鞘脂信号在多种细胞类型、脊髓区域和疾病阶段的动力学, 为之后药物的设计及治疗策略提供了方向。

空间转录组技术此前已成功应用于小鼠大脑复杂的结构与功能关系研究。大脑的空间结构与其功能是紧密相关的, 研究大脑的空间结构有助于更好地理解神经发育、神经精神病学和神经退行性疾病。2020 年, Maynard 等^[36] 首次使用 10 $\times$ Genomics Visium 技术获得了人类背外侧前额叶皮层 (dorsolateral prefrontal cortex, DLPFC) 的空间基因表达图谱。根据组织学和细胞结构, DLPFC 可以分为 6 层 (L1~L6), Maynard 等通过 Visium 技术构建了 DLPFC 空间基因表达图谱, 发现了很多以前未得到充分认识的、但可能具有更高保真度的空间分层特有标记基因, 包括 HPCAL1 (L2)、KRT17 (L6) 和 TRABD2A (L5), 这些基因可用于人

类大脑单细胞 RNA 测序的分层注释。

10 $\times$ Genomics Visium 空间转录组测序可以对测序数据进行可视化分析, 与组织切片形貌进行结合, 深入探究细胞在特定空间位置上的生物学功能。下一步如果将免疫荧光标记和 10 $\times$ Genomics Visium 技术结合起来, 可以把蛋白质检测和整个转录组分析相结合, 在空间范围下探索转录组和蛋白质的关系。

1.3 基于原位测序的空间转录组技术

1.3.1 原位测序 (ISS)

2013 年, Ke 等^[37] 开发了原位测序 (*in situ* sequencing, ISS) 方法, 该方法在固定的组织或细胞中先将 mRNA 在原位逆转录成 cDNA, 然后用 RNA 酶 H 使 cDNA 上的 RNA 链降解, 再通过锁式探针 (padlock probe) 与 cDNA 杂交。由于探针末端之间有缺口, 可以通过 DNA 聚合和 DNA 连接使锁式探针形成一个完整的环, 称为 “DNA 球”, 通过滚环扩增 (rolling-circle amplification, RCA) 形成滚环产物 (rolling-circle product, RCP) (图 3a)。之后, 在显微镜下观察荧光团偶联的寡核苷酸杂交, 这是一个循环。在进行下一个循环之前, 要将之前的探针冲洗干净, 再重复连接、成像、冲洗的步骤来进行下一个寡核苷酸的测序, 直至几个循环完成 (图 3b)。通过图像分析获得的碱基顺序来实现原位测序 (图 3c)。组织中可分辨的基因数量与荧光团的数量以及测序周期数有关, F 个荧光团和 n 个测序周期可解码 F^n 个不同的序列条码, 如 4 个荧光团和 4 个测序周期 (得到 4 个核苷酸长度的序列), 则可以解码 $4^4 = 256$ 个不同的序列条码。除此以外, ISS 首次实现了对小 RNA 片段的组织原位测序。

利用该技术, Ke 等^[37] 对人类乳腺癌组织切片中的突变位点进行了原位测序, 并揭示了其基因表达特征。2014 年, Pacureanu 等^[38] 开发了一种基于图像的方法, 每个文库都用荧光染料标记, 该染料为每个 RCP 生成特定的颜色 (对应于匹配的碱基), 并使用染料专用滤光片通过荧光显微镜获取不同焦平面的图像, 可对细胞和组织中的大量 RNA 片段进行测序。他们对 3 个 HER2 阳性新鲜冷冻乳腺癌组织切片进行了原位 RNA 表达和定位分析, 通过成像区分肿瘤与非肿瘤之间的界限。

ISS 适用于复杂结构的组织或细胞, 它可以对单个 RNA 分子进行测序, 分辨率高, 为研究复杂的生物学事件提供了新的途径, 例如原位异质细胞

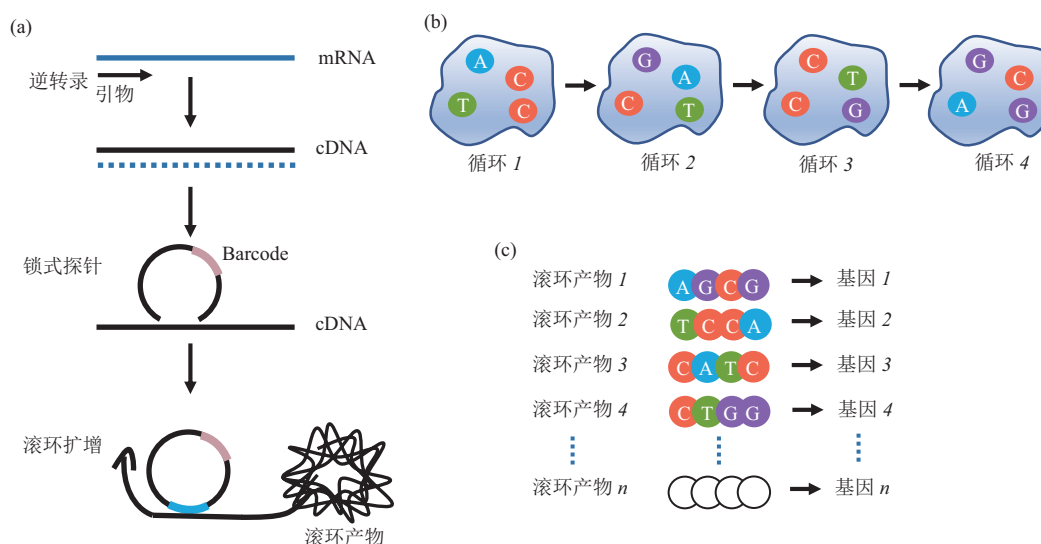
Fig. 3 The principle of *in situ* sequencing

图3 靶向原位测序原理

(a) mRNA在原位逆转录成cDNA, 然后用RNA酶使RNA降解, 再通过锁式探针与cDNA杂交, 探针末端之间有缺口, 通过DNA聚合和DNA连接可以使锁式探针形成一个完整的环, 之后通过滚环扩增(RCA)形成滚环产物(RCP)。(b) 在显微镜下观察荧光团偶联的寡核苷酸杂交, 在进行下一个循环之前, 要将之前的探针冲洗干净, 再重复连接、成像、冲洗的步骤来进行下一个寡核苷酸的测序, 直至几个循环完成。(c) 通过图像分析所获得的碱基顺序, 以此实现原位测序。

群体中的细胞分化^[37]。然而, ISS的主要缺点在于一次实验所能识别到的RNA数量较少, 且使用的锁式探针可能会产生大量的探针特定偏差, 即使结合了图像的方法, 增加了识别到的RNA数量, 但还需要处理大量图像数据, 且需要高分辨率的显微成像才能够准确获得碱基对信息, 扫描出来的图像数据通常非常大, 这对图像分析有不小的挑战。

1.3.2 荧光原位测序(FISSEQ)

ISS技术中, 锁式探针可能会产生大量的探针特定偏差, 而且ISS技术通常仅限于有明确注释基因的检测。针对这些问题, 2015年, Lee等^[39]开发了荧光原位测序(fluorescence *in situ* sequencing, FISSEQ)技术, 利用随机六聚体引物在固定细胞中逆转录RNA, 然后通过cDNA的环化以非靶向方式生成测序文库。该方法进一步提高了原位测序技术的检测通量, 能获得全基因组范围的基因表达图谱, 包括基因表达, RNA剪接和转录后修饰等, 同时保留了它们的空间位置信息。

FISSEQ技术首先将细胞固定在载玻片上, 然后原位进行逆转录。逆转录后, 将残留的RNA降解以防止其竞争性抑制双链DNA环化连接酶(CircLigase)。cDNA片段在60°C环化, CircLigase将环化后的cDNA连接起来, 之后进行滚环扩增

(rolling-circle amplification, RCA), 将每个cDNA环形扩增为包含多个原始cDNA序列的滚环产物(RCP)(图4)。该过程中, 在第k个循环引入荧光标签, 对应于该cDNA链的第k个碱基, 然后再进行成像。之后冲洗掉荧光标签, 再进行下一个循环。FISSEQ技术适合有细胞显微操作经验的研究

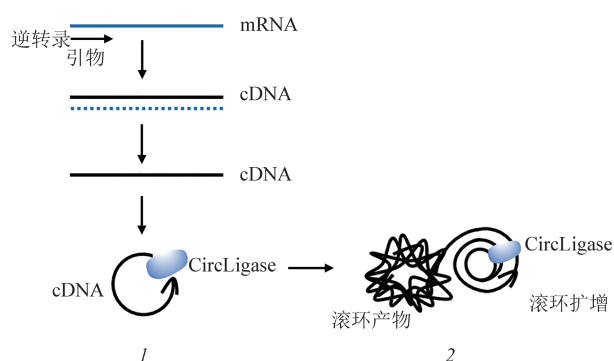


Fig. 4 The principle of rolling-circle amplification in FISSEQ

图4 荧光原位测序的滚环扩增原理

1: 原位进行逆转录, 之后将RNA降解, cDNA片段在60°C环化, CircLigase将环化后的cDNA连接起来。2: 进行滚环扩增(RCA), 将每个cDNA环线性扩增为包含多个原始cDNA序列的滚环产物(RCP), 在第k个循环引入荧光标签, 对应于该cDNA链的第k个碱基。

人员, 文库构建和测序可以在 14 d 内完成, 该技术对计算处理能力的要求较低, 图像分析大约只需 2 d^[39]。

FISSEQ 技术的优势在于, 对于功能重要的转录本, 其富集程度是单细胞 RNA-seq 的 10 倍以上。FISSEQ 方法的局限性在于获得的基因数远少于 RNA-seq, 每个细胞仅获得约 200 个 mRNA 片段, 而单细胞 RNA-seq 则约为 40 000 个 mRNA 片段。该方法测序深度低, 即使最初去除了 rRNA, 但 rRNA 读数仍占总读数的 40%~80%。除此以外, RCP 的大小又会限制每个细胞所能得到的读数, 从而限制了该方法的灵敏度^[34]。FISSEQ 无法提供单个细胞内 RNA 的完整信息, 会丢失低丰度的转录本^[40]。

FISSEQ 技术适用于大多数培养细胞和组织切片的样品 (包括石蜡切片和冰冻切片), 也适用于果蝇胚胎、诱导的多能干细胞及类器官等组织。FISSEQ 测序得到的每个 read 均有空间坐标, 可以用于分析转录本在原位的分布情况。FISSEQ 技术还可以与锁式探针技术结合, 来检测癌组织中特异的 RNA 标志物。除此以外, FISSEQ 还可以对单个细胞的转录本进行测序^[34]。

1.3.3 STARmap 技术

2018 年, Wang 等^[41]开发了 STARmap (spatially-resolved transcript amplicon readout mapping) 方法, 将水凝胶的组织化学方法和原位测序相结合, 可应用于完整的组织。STARmap 技术首先去除组织中的蛋白质和脂质, 将组织转化为 3D 水凝胶, 降低成像背景, 再用锁式探针与 cDNA 结合, 进行酶促扩增, 形成“DNA 球”, 这一过程相当于原位测序的滚环扩增。接着进行成像, 成像一轮后使用 60% 甲酰胺将探针从组织水凝胶中洗脱下来, 再开始下一个循环, 从而实现对完整组织的 3D 原位测序。

1.4 基于活细胞标记的空间转录组技术

1.4.1 TIVA

上述几类空间转录组技术, 需要对组织进行切片, 而组织一旦切片, 就无法获得完整单细胞的基因表达信息, 当前的捕获方法难以从单个细胞中分离出 mRNA 而不损害相邻组织。2014 年, Lovatt 等^[42]开发了一种新的空间转录组方法——TIVA (transcriptome *in vivo* analysis), 第一次非破坏性地从活细胞中捕获 mRNA。该方法设计了一种光激活的标签 (TIVA-tag), TIVA-tag 包括一段 poly(U)

序列和两个 poly(A) 片段, 该标签通过二硫键连接的细胞穿透肽穿透细胞膜, 以使 TIVA-tag 进入细胞, 之后在细胞中进行光裂解, 使得 poly(U) 序列暴露出来, 并与细胞中 mRNA 的 poly(A) 尾退火连接。组织裂解后, 用链霉亲和素和磁珠纯化得到 TIVA-tag-mRNA 结合产物。之后, 将 mRNA 洗脱下来, 进一步用于转录组分析 (图 5)。

TIVA 适用于活细胞或活组织, Lovatt 等^[42]对比了 TIVA-tag 方法和移液管吸取方法捕获的单个神经元细胞的转录组相关性, 结果显示无显著差异 ($P = 0.2396$, t 检验)。除此以外, Lovatt 等^[42]还在小鼠大脑的海马体活组织切片中使用了 TIVA-tag 方法, 证明了光活化的 TIVA-tag 可以保留在目标细胞内, 不会迁移到邻近细胞。TIVA-tag 可从活组织切片中的单个细胞中捕获大小和丰度不同的 mRNA, 而不会造成明显的细胞损伤。利用 TIVA-tag 方法, 他们还鉴定了人脑组织切片中单个细胞的基因表达, 证明 TIVA 可用于表征人类细胞在原位组织环境中的转录组。2017 年, Yeldell 等^[43]对原始 TIVA 探针进行改进, 增强了光解前的热稳定性, 提高了光解后对 mRNA 的捕获亲和力。

TIVA 方法与 RNA-seq 的结合能获得具有空间信息的完整单细胞基因表达信息, 而不是某个切片或者部分单细胞的基因表达信息, 从而为研究复杂的组织微环境对完整单细胞转录组的影响提供了一种可靠途径。但是该方法目前不适用于大量细胞的分析, 且因其只能穿透活细胞, 对已处理过的组织样品 (例如固定后或冰冻后的组织样品) 不适用。

1.4.2 ZipSeq

2020 年, Hu 等^[44]发明了 ZipSeq 方法, 用于活细胞和活组织切片的空间转录组研究。该方法利用光控空间编码进行实时的活组织单细胞空间转录组研究。该技术将连有双链 DNA 片段的高亲和力抗体或者脂质修饰过的寡核苷酸 (lipid-modified oligonucleotide, LMO) 锚定到细胞膜表面, 用特定模式照明, 释放特定组织区域细胞上锚定的双链 DNA 片段上所连接的笼锁序列。再利用与该笼锁序列相匹配的, 且带有不同 zipcode, 以 poly(A) 序列以及 Illumina 读取序列结尾的单链杂交。该 zipcode 作为组织细胞空间位置的识别序列, poly(A) 序列作为反转录时 poly-dT 引物的捕获对象, 对其进行放大作为细胞空间位置鉴定的依据。该方法对活组织切片进行位置标记后, 将细胞解离成单个细胞, 再结合 10×Genomics 单细胞测序的方

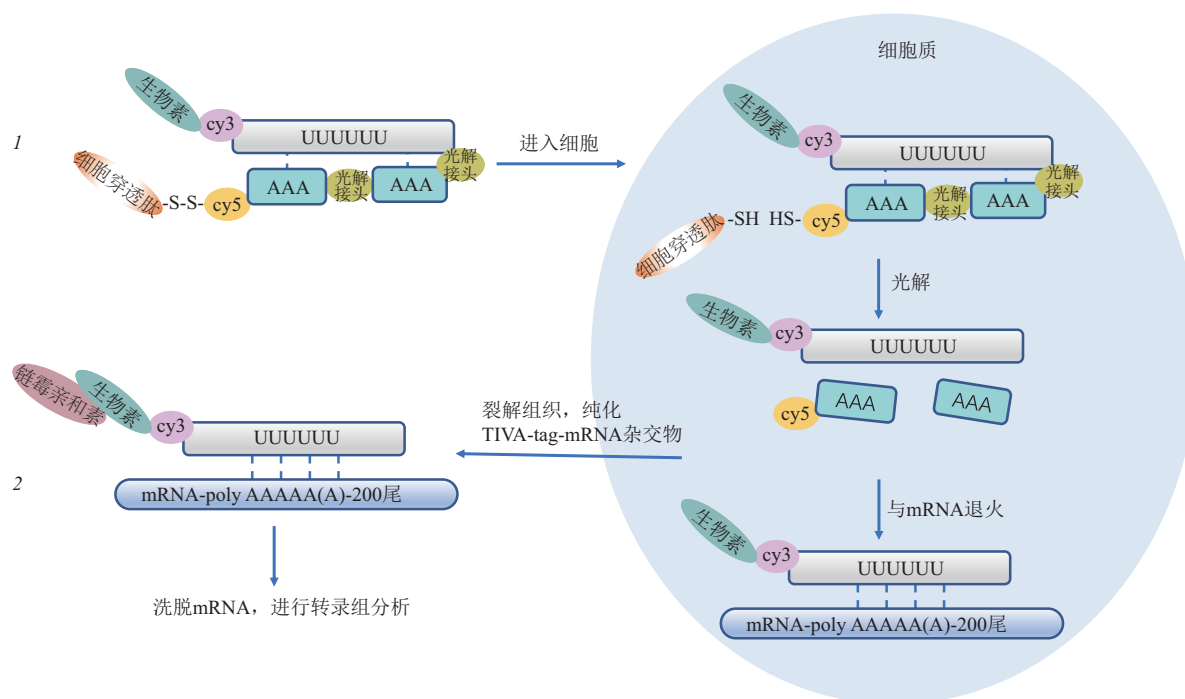


Fig. 5 The principle of TIVA

图5 TIVA原理

1: TIVA-tag通过二硫键连接的细胞穿透肽穿透细胞膜进入细胞, 之后在细胞中进行光裂解, 使得poly(U)序列暴露出来, 并与细胞中的mRNA的poly(A)尾退火连接。2: 组织裂解后, 用链霉亲和素和磁珠纯化得到TIVA-tag-mRNA结合物。之后, 将mRNA洗脱下来, 进一步用于转录组分析。

法进行转录组测序。

利用 ZipSeq 技术, 他们检测了体外伤口愈合过程中细胞基因表达的变化情况、淋巴结活组织切片的空间转录组情况以及肿瘤活组织内空间转录组表达情况。每种模型中都发现了一些与组织结构紧密相关的新的基因表达模式。在肿瘤微环境中, 基因表达模式表明了骨髓和T细胞由外围向内的分化轨迹。该方法还提供了两种 ZipSeq 的组合变体, 可以有效缩放所定义区域的数量, 从而为实时的或干扰之后的活组织基因表达模式的完整绘制提供了有效的途径, 但该方法在空间分辨率上仍然有待进一步提高。

2 展 望

目前发展的各类空间转录组技术, 都存在各自的优缺点(表1)。总的来说, 基于原位杂交技术的 smFISH 及其衍生的空间转录组方法在不断地提高检测的转录本数目。但是检测的通量和灵敏度还受到组织空间的限制, 所获得的基因表达信息只在一个组织层面上, 还无法反映组织内完整单细胞的基因表达情况, 价格也相对昂贵。基于激光显微切

割的空间转录组方法, 可以直观地分选出任何形态的细胞或细胞群, 通过高通量测序获得已知和未知的基因表达信息, 具有较好的性价比。但是由于组织需要切片, 同样无法获得组织内完整单细胞的基因表达信息。基于组织原位测序的空间转录组方法, 空间分辨率较高, 可以直观地在较大组织切片尺度观察基因表达的丰度情况, 但是获得的也是某个组织层面的细胞基因表达信息, 检测通量不高, 且捕获效率目前仍然非常有限。最近发展的活细胞空间转录组方法, 可以获得具有空间信息的完整单细胞的基因表达情况, 但对于细胞的精准定位以及细胞通量还需要进一步提高。

可以预见, 这些空间转录组的方法将不断克服各自的不足, 包括防止RNA的降解、提高检测的通量和效率、降低成本和获取完整空间单细胞转录组等。同时, 在获得细胞空间信息的基础上, 未来空间转录组技术将加入时间的维度, 从时空转录组的角度进一步提升转录组的研究层次, 不断加深对组织细胞真实特性的理解, 从而推进对发育过程和癌症等恶性疾病发生机制的理解和新治疗方案的研究。

Table 1 Comparison of spatial transcriptome technologies

表1 空间转录组技术比较

空间转录组方法		可检测转录本数	优势	局限性
基于原位杂交的空间转录组技术	smFISH	3~4个	检测效率相对较高	检测的RNA数量有限; 提高检测数量需杂交易产生检测误差, 且价格昂贵; 要求来自单个RNA的信号非重叠, 可测量的RNA的密度有限
	组合标记	2^n-1 (n =荧光团数量)		
	组合标记与SRM结合	$>2^n-1$ (n =荧光团数量)		
	顺序杂交	F^n 个 (F =荧光团数量, n =杂交回合)		
	MERFISH	$>1\ 000$		
基于高通量测序的空间转录组技术	LCM-RNA-seq	高通量	可用于分离难以分离的组织, 较为直观, 应用广泛	由于组织切片会将细胞切破, 可能捕获的是不完整细胞
	HDST和Slide-Seq	高通量	将基因表达情况可视化地比对到原始的组织切片中, 结果直观, 分辨率高	磁珠的捕获效率仍有待提高
	10× Genomics Visium	高通量	操作简便, 周期短, 具有高灵敏度和高细胞分辨率	无法获得全长转录本信息
基于原位测序的空间转录组技术	ISS	F^n 个 (F =荧光团数量, n =测序周期数)	可以在组织原位和细胞中测序单个RNA分子, 且适用于小RNA片段	一次实验所能识别到的RNA数量较少, 锁式探针易产生大量的探针特定偏差
	FISSEQ	每个细胞约200个mRNA	对于功能重要的转录本富集程度高	rRNA读数占总读数的40%~80%, 检测效率低
	STARmap	F^n 个 (F =荧光团数量, n =测序周期数)	组织经水凝胶化处理, 自发荧光背景较低, 可应用于完整组织	一次实验所能识别到的RNA数量较少, 锁式探针易产生大量的探针特定偏差
基于活细胞标记的空间转录组技术	TIVA	高通量	可以从活细胞中获取mRNA, 且不会破坏周围结构	不适用于大量细胞的分析, 只能穿透活细胞, 对已处理过的组织样品不适用
	ZipSeq	高通量		空间分辨率仍有待提高

参 考 文 献

[1] Hrdlickova R, Toloue M, Tian B. RNA-Seq methods for transcriptome analysis. Wiley Interdiscip Rev RNA, 2017, 8(1): 10

[2] Van Den Brink S C, Sage F, Vártesy Á, et al. Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. NatMethods, 2017, 14(10): 935-936

[3] Raj A, Van Den Bogaard P, Rifkin S A, et al. Imaging individual mRNA molecules using multiple singly labeled probes. Nat Methods, 2008, 5(10): 877-879

[4] Gall J G, Pardue M L. Formation and detection of RNA-DNA hybrid molecules in cytological preparations. Proc Natl Acad Sci USA, 1969, 63(2): 378-383

[5] Bauman J, Wiegant J, Borst P, et al. A new method for fluorescence microscopical localization of specific DNA sequences by in situ hybridization of fluorochrome-labelled RNA. Exp Cell Res, 1980, 128(2): 485-490

[6] Knoll J H, Lichter P. In situ hybridization to metaphase chromosomes and interphase nuclei. Curr Protoc Hum Genet, 2005, 45(1): 4.3.1-4.3.31

[7] Kallioniemi A, Kallioniemi O P, Waldman F, et al. Detection of retinoblastoma gene copy number in metaphase chromosomes and interphase nuclei by fluorescence in situ hybridization. Cytogenet Cell Genet, 1992, 60(3-4): 190-193

[8] Mark H F L, Rehan J, Mark S, et al. Fluorescence in situ hybridization analysis of single-cell trisomies for determination of clonality. Cancer Genet Cytogenet, 1998, 102(1): 1-5

- [9] Femino A M, Fay F S, Fogarty K, *et al.* Visualization of single RNA transcripts in situ. *Science*, 1998, **280**(5363): 585-590
- [10] Raj A, Van Den Bogaard P, Rifkin S A, *et al.* Imaging individual mRNA molecules using multiple singly labeled probes. *Nat Methods*, 2008, **5**(10): 877-879
- [11] Levsky J M, Shenoy S M, Pezo R C, *et al.* Single-cell gene expression profiling. *Science*, 2002, **297**(5582): 836-840
- [12] Jakt L M, Moriwaki S, Nishikawa S. A continuum of transcriptional identities visualized by combinatorial fluorescent *in situ* hybridization. *Development*, 2013, **140**(1): 216-225
- [13] Lubeck E, Cai L. Single-cell systems biology by super-resolution imaging and combinatorial labeling. *Nat Methods*, 2012, **9**(7): 743-748
- [14] Lubeck E, Coskun A F, Zhiyentayev T, *et al.* Single-cell *in situ* RNA profiling by sequential hybridization. *Nat Methods*, 2014, **11**(4): 360-361
- [15] Cai L. Turning single cells into microarrays by super-resolution barcoding. *Brief Funct Genomics*, 2013, **12**(2): 75-80
- [16] Chen K H, Boettiger A N, Moffitt J R, *et al.* Spatially resolved, highly multiplexed RNA profiling in single cells. *Science*, 2015, **348**(6233): aaa6090
- [17] Moffitt J R, Hao J, Wang G, *et al.* High-throughput single-cell gene-expression profiling with multiplexed error-robust fluorescence *in situ* hybridization. *Proc Natl Acad Sci USA*, 2016, **113**(39): 11046-11051
- [18] Moffitt J R, Zhuang X. RNA imaging with multiplexed error-robust fluorescence *in situ* hybridization (MERFISH). *Methods Enzymol.* 2016, **572**: 1-49
- [19] Wang G, Moffitt J R, Zhuang X. Multiplexed imaging of high-density libraries of RNAs with MERFISH and expansion microscopy. *Sci Rep*, 2018, **8**(1): 4847
- [20] Lalmansingh A S, Arora K, Demarco R A, *et al.* High-throughput RNA FISH analysis by imaging flow cytometry reveals that pioneer factor Foxa1 reduces transcriptional stochasticity. *PLoS One*, 2013, **8**(9): e76043
- [21] Ronander E, Bengtsson D C, Joergensen L, *et al.* Analysis of single-cell gene transcription by RNA fluorescent in situ hybridization (FISH). *J Vis Exp*, 2012, **68**: e4073
- [22] Cui C, Shu W, Li P. Fluorescence *in situ* hybridization: cell-based genetic diagnostic and research applications. *Front Cell Dev Biol*, 2016, **4**: 89
- [23] Moffitt J R, Hao J, Bambah-Mukku D, *et al.* High-performance multiplexed fluorescence in situ hybridization in culture and tissue with matrix imprinting and clearing. *Proc Natl Acad Sci USA*, 2016, **113**(50): 14456-14461
- [24] Emmert-Buck M R, Bonner R F, Smith P D, *et al.* Laser capture microdissection. *Science*, 1996, **274**(5289): 998-1001
- [25] Curran S, McKay J, Mcleod H, *et al.* Laser capture microscopy. *Mol Pathol*, 2000, **53**(2): 64
- [26] Davies J A, Unbekandt M, Ineson J, *et al.* Dissociation of embryonic kidney followed by re-aggregation as a method for chimeric analysis. *Methods Mol Biol*, 2012, **886**: 135-146
- [27] 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. *Dev Cell*, 2016, **36**(6): 681-697
- [28] Moor A E, Harnik Y, Ben-Moshe S, *et al.* Spatial reconstruction of single enterocytes uncovers broad zonation along the intestinal villus axis. *Cell*, 2018, **175**(4): 1156-1167.e15
- [29] Halpern K B, Shenhav R, Matcovitch-Natan O, *et al.* Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. *Nature*, 2017, **542**(7641): 352-356
- [30] Vickovic S, Eraslan G, Salmén F, *et al.* High-definition spatial transcriptomics for *in situ* tissue profiling. *Nat Methods*, 2019, **16**(10): 987-990
- [31] Rodriques S G, Stickels R R, Goeva A, *et al.* Slide-seq: a scalable technology for measuring genome-wide expression at high spatial resolution. *Science*, 2019, **363**(6434): 1463-1467
- [32] Larsson L, Frisén J, Lundeberg J. Spatially resolved transcriptomics adds a new dimension to genomics. *Nat Methods*, 2021, **18**(1): 15-18
- [33] Rao N, Clark S, Habern O. Bridging genomics and tissue pathology: 10× Genomics explores new frontiers with the visium spatial gene expression solution. *Genet Eng Biotechnol News*, 2020, **40**(2): 50-51
- [34] Moncada R, Barkley D, Wagner F, *et al.* Integrating microarray-based spatial transcriptomics and single-cell RNA-seq reveals tissue architecture in pancreatic ductal adenocarcinomas. *Nat Biotechnol*, 2020, **38**(3): 333-342
- [35] Maniatis S, Äijö T, Vickovic S, *et al.* Spatiotemporal dynamics of molecular pathology in amyotrophic lateral sclerosis. *Science*, 2019, **364**(6435): 89-93
- [36] Maynard K E, Collado-Torres L, Weber L M, *et al.* Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. *Nat Neurosci*, 2021, **24**(3): 425-436
- [37] Ke R, Mignardi M, Pacureanu A, *et al.* *In situ* sequencing for RNA analysis in preserved tissue and cells. *Nat Methods*, 2013, **10**(9): 857-860
- [38] Pacureanu A, Ke R, Mignardi M, *et al.* Image based *in situ* sequencing for RNA analysis in tissue//IEEE. Proceedings of the 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI). Beijing: IEEE, 2014: 286-289
- [39] Lee J H, Daugharthy E R, Scheiman J, *et al.* Fluorescent in situ sequencing (FISSEQ) of RNA for gene expression profiling in intact cells and tissues. *Nat Protoc*, 2015, **10**(3): 442-458
- [40] Ginart P, Raj A. RNA sequencing *in situ*. *Nat Biotechnol*, 2014, **32**(6): 543-544
- [41] Wang X, Allen W E, Wright M A, *et al.* Three-dimensional intact-tissue sequencing of single-cell transcriptional states. *Science*, 2018, **361**(6400): eaat5691
- [42] Lovatt D, Ruble B K, Lee J, *et al.* Transcriptome *in vivo* analysis (TIVA) of spatially defined single cells in live tissue. *Nat Methods*, 2014, **11**(2): 190-196
- [43] Yeldell S, Ruble B, Dmochowski I. Oligonucleotide modifications enhance probe stability for single cell transcriptome *in vivo* analysis (TIVA). *Org Biomol Chem*, 2017, **15**(47): 10001-10009
- [44] Hu K H, Eichorst J P, McGinnis C S, *et al.* ZipSeq: barcoding for real-time mapping of single cell transcriptomes. *Nat Method*, 2020, **17**(8): 833-843

Advances in Spatial Transcriptome Technologies*

WEI Ying**, ZHANG Xiao-Dan**, HU Miao-Miao, WU Zhong-Qin, CHENG Ming, GUO Yan***

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

Abstract Spatial transcriptome technologies aim to quantitatively measure the gene expression of cells and provide information on the specific location of cells in tissue space. Compared with traditional transcriptome technologies, the spatial transcriptome technologies can obtain the true gene expression characteristics of cells in the tissue in physiological environment and its relationship with the microenvironment, further advancing the understanding of cell characteristics in normal and pathological states. In recent years, significant progress has been made in the development of spatial transcriptome technologies. The cell throughput, detected quantity and quality of transcripts have been continuously improved, and the spatial location information has become more accurate and comprehensive. This paper reviews the development and applications of spatial transcriptome technologies, which were classified into 4 major categories based on *in situ* hybridization, high-throughput sequencing, *in situ* sequencing, and live cell barcodes, respectively. Each of them has its advantages and disadvantages so that should be applied in different situations, and it is foreseeable that these spatial transcriptome technologies will continue to be improved, including preventing RNA degradation, improving detection throughput and efficiency, reducing costs, and obtaining complete spatial single-cell transcriptomes. At the same time, based on the acquisition of cell spatial information, future spatial transcriptome technologies will be combined with the dimension of time, further improve the level of transcriptome research from the perspective of spatiotemporal transcriptome, and continuously deepen the understanding of the true characteristics of tissue cells, so as to advance the understanding of developmental processes, cancer and other malignant diseases and the development of new treatments.

Key words spatial transcriptome, omics technologies, spatial location, single cell

DOI: 10.16476/j.pibb.2021.0006

* This work was supported by grants from The National Key Research and Development Program of China (2018YFC1003500) and The National Natural Science Foundation of China (81972909).

** These authors contributed equally to this work.

*** Corresponding author.

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

Received: April 20, 2021 Accepted: July 13, 2021