



## 单细胞全基因组扩增技术与应用\*

徐晓丽 吴凌娟 鄢仁祥\*\*

(福州大学生物科学与工程学院, 福州 350100)

**摘要** 同一组织中的细胞往往具有类似的结构和功能, 然而通过对单个细胞进行测序分析后, 发现每个细胞都具有一定异质性. 单细胞全基因组扩增技术是进行单细胞测序的前提, 该技术可用于揭示单细胞基因组结构差异, 同时在肿瘤研究、发育生物学、微生物学等研究中发挥重要作用, 并成为生命科学研究技术的热点之一. 单细胞全基因组扩增技术的难点在于单细胞的分离和全基因组的扩增. 本文介绍了单细胞全基因组扩增技术中常用的单细胞分离技术和单细胞全基因组扩增技术, 并对各技术间的优缺点进行比较, 同时着重讨论该技术在肿瘤研究、发育生物学和微生物学研究中的应用.

**关键词** 单细胞分离技术, 单细胞全基因组扩增技术, 比较, 应用

**中图分类号** Q7, Q3

**DOI:** 10.16476/j.pibb.2018.0259

DNA 测序技术的快速发展使得学术界对包括人类在内的各类物种的基因组有了更全面的认识, 下一代测序的全基因组扩增 (whole-genome amplification, WGA) 在生物学和医学领域都有广泛的应用. 人类对于基因组的普遍认识来源于对一类细胞的全部基因组进行测序, 但问题是同一个体中的单个细胞由于后天的环境变化会拥有一个独特的基因组, 而生殖细胞也会因携带着来自父亲和母亲基因不同的组合而不一样. 例如, 生活中常见的异卵双胞胎和同卵双胞胎之间都有着各自的区别. 随着时间的推移, 基因组位置会发生随机变化, 体细胞相应地就会出现基因组变化, 这种基因组变异包括单核苷酸变异 (single-nucleotide variations, SNVs)、拷贝数变异 (copy-number variations, CNVs) 和结构变化, 常导致癌症和其他疾病. 随着测序技术、单细胞分离技术和全基因组扩增技术的快速发展, 单细胞研究技术被列为未来几年最值得关注的技术之一<sup>[1]</sup>. 在单细胞水平上对基因组进行测序能更好地解决传统技术的局限性问题. Navin 等<sup>[2]</sup> 通过利用流式细胞仪对来自多基因组肿瘤的 100 个单细胞和单株原发肿瘤及其肝转移的 100 个单细胞进行测序分析, 开展肿瘤的进化研究. 单细胞测序技术描述单个细胞的基因组, 能够解析细胞间更加细微的差异, 探索基因组变异的异

构细胞<sup>[3-5]</sup>, 揭示了各种生物过程的细节. 例如, 构建新的微生物进化树和基因组序列, 推动微生物生命活动过程的研究<sup>[6-7]</sup>, 同时对发育生物学、神经科学、免疫、癌症等领域的发展也发挥重要作用, 包括胚胎发育<sup>[8]</sup>、肿瘤进化<sup>[2, 9-10]</sup>等. 不可否认, 单细胞测序技术正成为生命科学研究的焦点之一<sup>[11]</sup>. 如今, 单细胞基因组的表征研究越来越受到关注, 且单细胞基因组学的重要性在珍稀样本中变得更加显著<sup>[12]</sup>, 例如胚胎细胞<sup>[13]</sup>和循环肿瘤细胞<sup>[9]</sup>. 本文从单细胞研究基本技术和应用两方面对单细胞基因组扩增技术的最近研究进展进行了阐述.

## 1 单细胞分离技术

根据所用技术原理, 单细胞分离技术可以划分为不同种类 (表 1). 对各种主要方法我们介绍如下.

### 1.1 梯度稀释法

梯度稀释法是通过将细胞进行梯度稀释, 最终得到单细胞的一种简易方法. 其最大优点是技术简

\* 国家自然科学基金(31500673)资助项目.

\*\* 通讯联系人.

Tel: 0591-22866376, E-mail: yanrenxiang@fzu.edu.cn

收稿日期: 2018-12-28, 接受日期: 2019-03-05

Table 1 Comparison of different single-cell isolation methods

表1 不同单细胞分离方法的比较

分离方法	原理	优点	缺点	应用
梯度稀释法	梯度稀释	技术操作简单, 成本低	易分离错误或丢失	主要用于人工培养样品分离研究
微流控技术	微流体芯片平行分离	灵活选择, 密闭操作空间	需要专门空间, 操作较复杂	用于样品量少的分离研究
荧光激活细胞分选技术	荧光标记特异分选	分选准确度高	分选机制复杂, 设备昂贵, 有气溶胶危险	用于对细胞活性和形态要求较低的研究
显微操作技术	倒置显微镜下直接分离	灵活取样, 操作简单, 成本低	可能损坏细胞, 出错率高	适用于小细胞群体的分离研究
激光捕获显微切割技术	激光定位切割	保证组织完整性	成本高, 出错率高, 细胞损坏严重	用于冰冻或石蜡包埋样品分离研究

单、成本低, 对于可以进行培养的样本研究均可用该方法得到单细胞, 缺点在于在细胞稀释过程中容易出现分离错误或者丢失. 这种方法在大批量细胞分析中也有一定的应用. 例如, Zhang 等<sup>[14]</sup>在2006年通过 Affymetrix 芯片杂交分析了两种大肠杆菌的分布情况.

### 1.2 微流控技术

微流控技术是使用具有专用微流体芯片进行平行单细胞分离, 捕获单细胞<sup>[15]</sup>. 微流体装置具有集成的细胞分选器, 用于分离选定的单个细胞, 从而允许灵活的样品选择且封闭的操作空间可以有效地避免污染. Streets 等<sup>[16]</sup>利用微流控芯片实现单细胞测序, 提高了反应效率, 减少了操作误差, 提高了可靠性和平行性.

### 1.3 荧光激活细胞分选

荧光激活细胞分选 (fluorescence activated cell sorting, FACS) 技术<sup>[17]</sup>借助流式细胞仪, 单个细胞有序地通过激光束, 利用激光激发荧光素标记的细胞, 根据激发的荧光信号对细胞进行高度特异性分选, 含有靶细胞的液滴被选择性地偏转到收集器中, 是比较常用的单细胞分离技术<sup>[18]</sup>. FACS 系统是在单细胞基础上分离靶细胞, 具有更高的分选准确度. 虽然非常有效和准确, 但复杂的分拣机制使得 FACS 系统显得昂贵且笨重, 并且潜在的气溶胶也有一定的危险性. 此外仪器中快速流动的液体可能会对细胞的活性和形态状态产生一定的影响<sup>[19]</sup>, 但是对于基因组 DNA 的研究没有影响.

### 1.4 显微操作技术

显微操作技术是利用显微镜直接进行单细胞的分离, 是一种较常用的灵活获取单细胞的技术, 显微操作系统以倒置显微镜为主. 其关键技术要点在于对培养的细胞进行梯度稀释, 如果细胞浓度过高

会影响分离和挑取单细胞的准确性. 通过细胞计数板计数将细胞稀释到大约 100 个/ml, 或者每个细胞在缓冲液中是独立分开状态即可, 在倒置显微镜的观察下可借助移液器直接挑取单个细胞. 其优点是操作容易进行、可以直接通过可见视野直观地分离单细胞、成本低、主要是用于较小细胞群体的目标细胞的分离. 缺点是通量低、分离过程可能会损伤细胞、而且细胞识别出错率高.

### 1.5 激光捕获显微切割

激光显微切割可以在不破坏组织结构, 并保证周围组织完整性的前提下将特定的单个细胞从不同的组织中分离出来, 此方法常应用于与人类健康相关的单细胞基因组学研究<sup>[20-21]</sup>. 其缺点是成本较高、切割不精确, 可能会造成遗传物质的丢失.

## 2 单细胞基因组扩增技术

### 2.1 引物延伸预扩增法

引物延伸预扩增法 (primer-extension preamplification, PEP) 是一种出现较早的 WGA 技术. 1992 年应用 PEP 方法实现单个单倍体细胞基因组的全扩增<sup>[22]</sup>. 这是一种基于 PCR 技术的 WGA 方法, 主要是使用了含 15 个碱基的随机引物, 在 37°C 下退火, 并在 55°C 进行延伸, 如此循环复制.

### 2.2 简并寡核苷酸引物 PCR 技术

基于热循环以 PCR 为基础的简并寡核苷酸引物 PCR 技术 (degenerate oligonucleotide-primed PCR, DOP-PCR)<sup>[23]</sup>, 其采用部分简并序列的寡核苷酸引物 (引物中间部分含有 6 个随机碱基), 用于基因组作图研究. 该技术在最初几个循环, 利用低初始退火温度 (~25°C) 的 PCR 方案确保引物与模板结合, 从给定基因组内的多个均匀分散的位点引发. 此后进行较高退火温度 (~55°C) 的常规多

循环PCR反应. Van等<sup>[24]</sup>在传统DOP-PCR技术上提高了其引物的简并性,同时使用4条简并引物进行扩增,每条引物含有8个随机碱基,并将扩增产物与二代测序技术联合使用,使得WGA方法与未知样品研究实现结合.相对于一般的DNA扩增技术,DOP-PCR是一种快速、高效、独立于一般DNA扩增的技术,但低覆盖率与PCR错误率高常导致单核苷酸变异(SNVs). Blagodatskikh等<sup>[25]</sup>在此基础上通过修改引物设计和DNA聚合酶,改进了DOP-PCR(improved degenerate oligonucleotide-primed PCR, iDOP-PCR)技术,其不仅提高了扩增效率,也提高了扩增质量.

PEP和DOP-PCR都是早期常用的经典技术,不同的序列PCR扩增的效率存在很大的偏差,其中对PCR的扩增方法影响较大的因素是聚合酶和引物浓度,在扩增过程中模板的大小、GC含量、DNA二级结构等都会对聚合酶的扩增效率有影响,不能完全覆盖整个基因组,将导致扩增产物不均匀或者出现非特异性扩增产物、出现扩增偏好性<sup>[26-28]</sup>,给下一步实验分析造成另一种挑战.

### 2.3 多重置换扩增

多重置换扩增(multiple displacement amplification, MDA)是等温的链置换扩增<sup>[29]</sup>,在

恒温条件下,由6个随机碱基构成的随机引物与模板随机退火,紧接着在phi29 DNA聚合酶作用下发生链置换反应,置换后的单链又可与引物发生随机结合、退火、延伸,最终形成分支扩增.与基于PCR的WGA方法相比<sup>[27]</sup>,MDA通过对染色体8个基因位点的扩增偏差小3倍,而DOP-PCR的扩增偏差可达4~6倍.MDA可从1~10基因组拷贝扩增产生约20~30 μg的产物,平均产物长度>10 kb.由于MDA是一种等温扩增方法,对模板质量要求相对较高,同时也存在一定的非特异性扩增,在不含模板的情况下也会有扩增产物<sup>[15, 30]</sup>. Leung等<sup>[31]</sup>用129个正常二倍体细胞的大数据集检测液滴多重置换扩增(droplet multiple displacement amplification, dMDA)技术的性能,结果显示其在扩增均匀性、基因组覆盖度和稳健性方面超过先前报道的单细胞WGA方法.实现了高达80%的单细胞基因组覆盖率,且利用液滴MDA扩增得到的产物在靶向测序中可有效地进行单核苷酸变异(SNV)检测.因此,MDA扩增的人类DNA常用于遗传分析,包括单核苷酸多态性的基因分型、染色体绘制、DNA印迹和限制性片段长度多态性分析、亚克隆和DNA测序等.一般MDA单细胞基因组测序流程如图1所示.

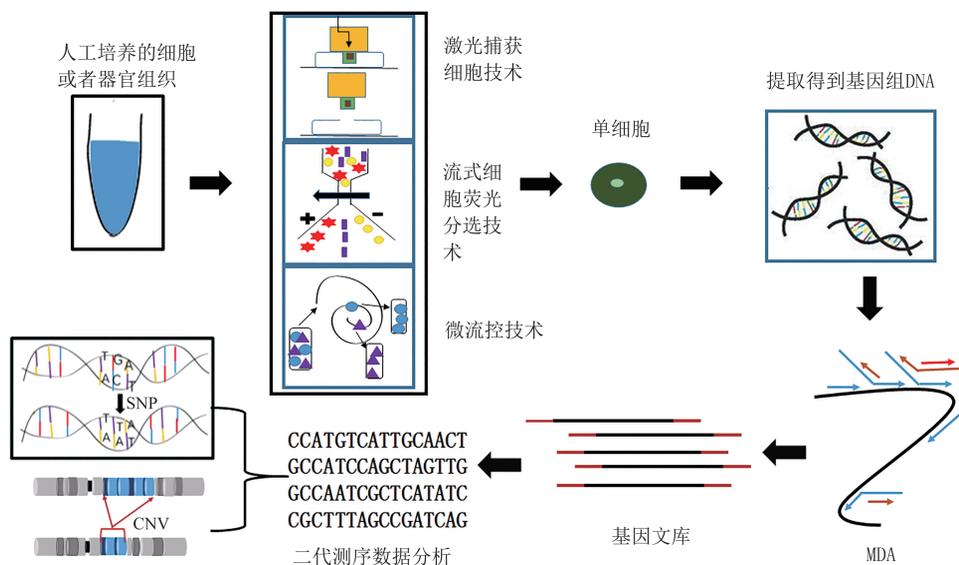


Fig. 1 Single cell genome sequencing process of MDA method

图1 MDA单细胞基因组测序流程

从样品细胞群(培养的细胞或者组织)中通过各种单细胞分离技术,分离得到单细胞,对单细胞进行裂解提取其基因组,用MDA方法扩增得到基因组文库,随后进行测序,对数据分析,进行SNP、CNV检测,细胞类型鉴定等.

## 2.4 多重退火和基于环状循环扩增技术

多重退火和基于环状循环扩增技术 (multiple annealing and looping-based amplification cycles, MALBAC) 是美国国家科学院院士谢晓亮教授所带领的团队在 2012 年发表了一种新的 WGA 方

法<sup>[28]</sup>, 其引入了拟线性预放大, 以减少与非线性放大相关的偏差. MALBAC 结合了 MDA 和基于 PCR 的 WGA 技术, 以人的单细胞的 DNA 片段 (10~100 kb) 作为模板. 具体原理如图 2 所示.

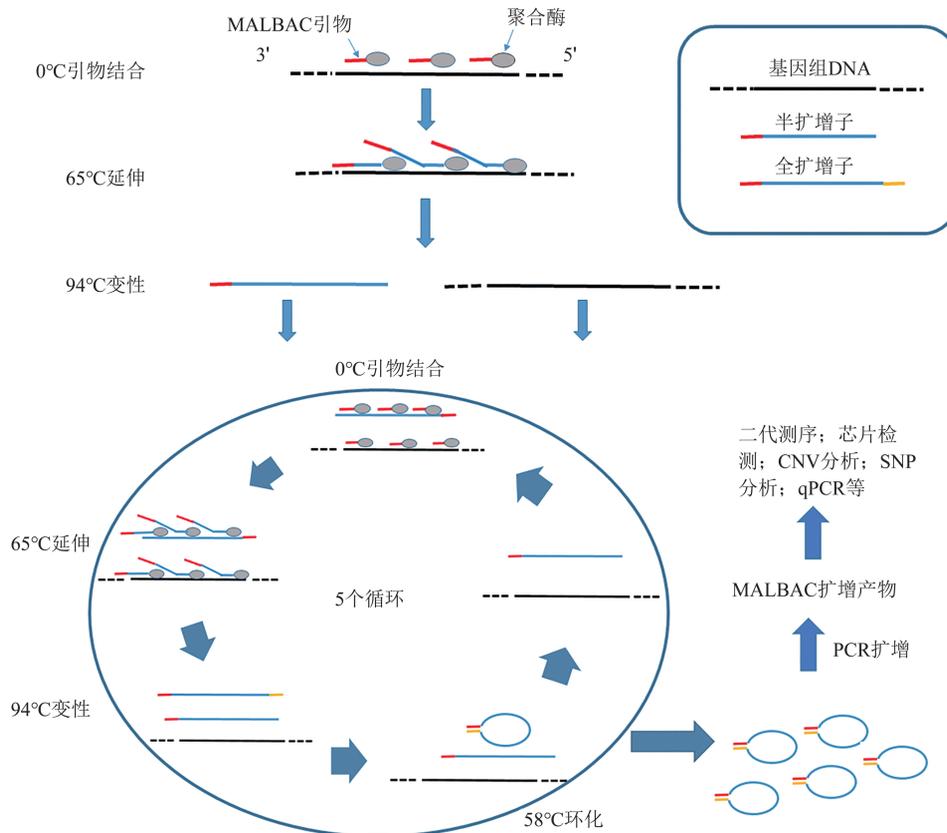


Fig. 2 Schematic diagram of MALBAC<sup>[28]</sup>

图2 MALBAC原理图<sup>[28]</sup>

MALBAC首先在0°C条件下与引物均匀地排列在基因组上, 与基因组结合, 随着温度上升到65°C, 引物开始延伸, 在94°C条件下延伸后的半扩增子与基因组分离, 随后反应又进行0°C-65°C-94°C的循环, 半扩增子形成完整的全扩增子, 温度降低至58°C, 全扩增子环化, 如此扩增反应进行5次, 环化的全扩增子可进行下一步的PCR扩增反应, 最终得到MALBAC扩增产物.

MALBAC由一组随机引物启动扩增, 每个引物具有27 nt通用引物序列和8 nt随机碱基, 随机引物可以在0°C下与模板均匀杂交. 在65°C和具有链置换活性的DNA聚合酶作用下产生具有可变长度(0.5~1.5 kb)的半扩增子, 再经过94°C模板变性、0°C退火和65°C延伸一个循环后半扩增子形成末端互补的全扩增子, 紧接着将温度降低至58°C, 使完整扩增子环化, 以防止DNA进一步扩增和交叉杂交. 把在5个预扩增循环后得到的全扩增子作为模板, 27 nt通用引物作为PCR引物, 通过PCR指数扩增得到下一代测序所需的DNA. 在相同条件

下, MALBAC比MDA方法展现出更高、更均匀的基因组覆盖率, 但同时也存在假阳性偏高的结果, 可能是因为Bst和Taq聚合酶的保真性不高, 以至于扩增过程中准确率低, 所以MALBAC需要多个细胞基因组才能获得更为准确的结果<sup>[32]</sup>.

## 2.5 乳液全基因组扩增

乳液全基因组扩增 (emulsion whole-genome amplification, eWGA) 适用于任何WGA, 其利用油中的少量水性液滴, 均匀扩增单个细胞的基因组. 通过将单细胞基因组DNA片段分配到大量( $10^5$ )皮升液滴中, 允许每个液滴中的一些DNA

片段达到扩增饱和状态. 经过破乳合并液滴后, DNA片段之间的扩增增益的差异显著最小化. 与MDA、MALBAC和DOP-PCR比较<sup>[33]</sup>后的结果表明, eWGA不仅提高了覆盖率, 而且能够以更高的准确度和更高的分辨率同时检测SNV和CNV, 在许多方面优于现有的单细胞扩增方法. 例如, MDA容易受到环境污染, 包括试剂中痕量DNA的污染. 相反, eWGA通过应用小的反应量<sup>[34-35]</sup>, 且其反应缓冲液会被分配到大量分离的液滴中, 而DNA污染物只存在于一小部分液滴中, 不会被过度放大, 因此可以有效减少DNA污染物.

## 2.6 通过转座子插入的线性扩增

2017年, Xie等<sup>[36]</sup>在《科学》杂志(*Science*)上发表了一种新的WGA方法, 即通过转座子插入的线性扩增(linear amplification via transposon

insertion, LIANTI), 其结合了Tn5转座<sup>[37]</sup>和T7体外转录<sup>[38]</sup>用于单细胞基因组分析, 优于现有方法, 能够以千碱基分辨率进行微型CNV检测. PCR的指数扩增会造成扩增偏差和错误, 限制了其在单细胞基因组学中的应用<sup>[39-40]</sup>. 在LIANTI方法中, 来自单个细胞的基因组DNA, 通过特异设计的包含T7启动子LIANTI转座子的Tn5转座而随机片段化. 由T7启动子标记的基因组DNA片段通过体外转录线性扩增成数千个拷贝的RNA, 在逆转录和RNA酶消化后, 合成第二条链, 形成用于DNA文库制备的双链LIANTI扩增子, 其原理如图3所示. LIANTI减少了其他单细胞WGA方法中使用的非特异性扩增和指数扩增, 从而大大降低了扩增偏差和误差.

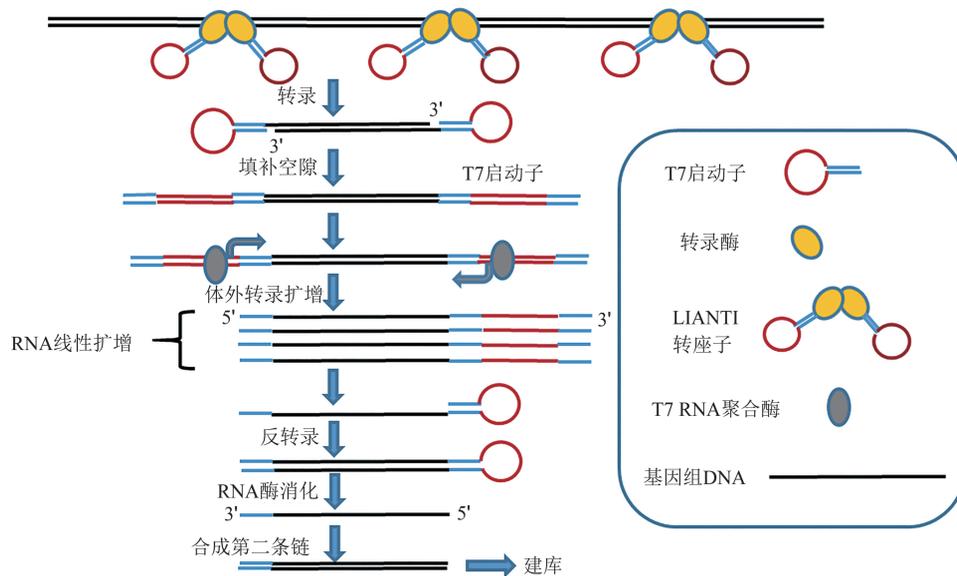


Fig. 3 Schematic diagram of LIANTI<sup>[36]</sup>

图3 LIANTI原理图<sup>[36]</sup>

LIANTI对来自单个细胞的基因组DNA, 随机片段化并用LIANTI转座子标记, 随后DNA聚合酶缺口延伸转换单链T7启动子在每个片段的两端环成双链T7启动子. 进行体外转录过夜以将基因组DNA片段线性扩增成基因组RNA, 其能够在3'末端自我启动. 在逆转录、RNase消化和第二链合成后, 形成标记有独特分子条形码的双链LIANTI扩增子, 代表来自单个细胞的原始基因组DNA的扩增产物, 并准备用于DNA文库制备和下一代测序.

## 2.7 利用微流体反应器进行单链测序

SISSOR (single-stranded sequencing using microfluidic reactors) 是利用微流体反应器进行单链测序的一种方法. 该方法用于精确的单细胞基因组测序和单元型分析<sup>[41]</sup>. 微流体装置由四个模块组成: 单细胞捕获模块、细胞裂解与链分离模块、

分区模块和扩增模块. 首先从细胞悬浮液中捕获单个细胞; 随后捕获的细胞被裂解, 双链染色体DNA分子被碱性溶液分离; 之后分离的DNA片段使用旋转泵随机分布, 并分割成24个相同的区室; 最后每个分区被中和并下推到扩增模块中, 由MDA进行扩增. 每个腔内的放大产物被检索, 最

终收集每个区室的扩增基因组DNA合并转换成条形码测序文库 (barcoded sequencing libraries), 并使用Illumina的合成短读测序进行测序. 利用冗余序列信息和基于单倍体型减少序列错误, 平均500 kb长的DNA片段可以组装成单倍体重叠群, 测序结果错误率低至 $10^{-8}$ , 而且在更均匀的扩增和更准确的序列比对后可以进一步提高性能. 从单个细胞获得准确的基因组序列和单倍体信息将使得基因组测序的应用能够满足不同的临床需求. 10x Genomics作为单细胞领域的后起之秀, 巧妙使用了Barcoding (条形码) 和Microfluidics (微流体) 技术, 在单细胞分离、扩增原理上具有明显的优势. 与二代测序数据相结合, 在Scaffold的组装上能够达到媲美三代测序的组装效果.

## 2.8 主流方法的比较

De Bourcy等<sup>[34]</sup>对MDA、MALBAC和PEP三种单细胞扩增方法作了定量比较, 发现MDA的扩增均匀性与扩增产率成反比, 而MALBAC和PEP则不明显. 当扩增产率高于 $10^6$ 时, MDA的覆盖率最低; 当扩增产率大于 $2.5 \times 10^3$ 时, MDA检测CNV效果较差; 当扩增产率小于 $5 \times 10^6$ 时, 3种方法的从头测序 (*de novo*) 拼接效果无差异. 聂等<sup>[42]</sup>发现, MDA技术不适合用于短串联重复序列 (short tandem repeat, STR), 等位基因缺失和伪等位基因严重, 相比之下PEP技术分型准确率较高, 但其对小片段的扩增偏倚性可能会导致大片段扩增失败. 在单细胞水平上对成纤维细胞样品进行低深度高通量测序, 比较MALBAC和MDA这两种方法的染色体拷贝数变异 (CNV) 检测率和等位基因丢失率 (allele dropout, ADO) 的系数<sup>[43]</sup>. 结果

CNV检测的成功率在MALBAC组为100%, 在MDA组为91.67%, MALBAC组CNV检测的变异系数明显优于MDA组. 在MALBAC组中等位基因检测的总ADO率为4.55%, 显著低于MDA组中观察到的22.5%. 当5个或更多个细胞作为初始模板时, MDA的ADO率显著降低, 以致于这两种方法之间的差异变得不明显. 利用MALBAC和MDA两种WGA方法对 $\beta$ -地中海贫血基因分型和单核苷酸多态性进行CNV检测<sup>[44]</sup>, 并统计分析了两种方法之间的扩增效率、阳性预测值、灵敏度和ADO值. 发现MDA的扩增率和ADO率均优于MALBAC, MDA的敏感性和阳性预测值均高于MALBAC, 证实MDA比MALBAC更适合用于SNP检测. 然而, 在单细胞水平上使用低深度NGS进行CNV检测时, MALBAC比MDA更稳定和均一. 因此MALBAC更适合于CNV检测, 而用MDA对SNV检测会好得多. 谢晓亮教授团队<sup>[36]</sup>对比了LIANTI、MDA、MALBAC和DOP-PCR方法, 结果表明, LIANTI扩增效率及均匀性是最优的, 并且基因组覆盖率可达到97%, 是目前CNV检测最准确的方法. SISSOR技术利用双链共有序列和长亚单倍体片段组装, 同时提供比使用类似单细胞扩增和测序平台的其他技术更高的每碱基测序准确度和更长的单倍型. 长片段读取方法<sup>[45]</sup>通过比较来自许多细胞的多个单链文库的共有序列来减少错误, 而SISSOR技术利用仅来自单个细胞的两个互补链的共识来进行错误校正, 是迄今为止最精确的单细胞基因组测序<sup>[41]</sup>. 关于各种方法之间的简易比较我们列于表2中.

Table 2 Comparison of mainstream WGA methods

表2 主要的WGA方法比较

WGA方法	扩增原理	引物组成	酶	产物长度	覆盖率
PEP	基于PCR的完全随机引物扩增法	含15个碱基的随机引物	DNA聚合酶	<2 kb	~50%
DOP-PCR	基于PCR的部分随机引物扩增法	引物中含6个随机引物的简并引物	DNA聚合酶	<2 kb	~40%
MDA	多重置换恒温扩增法	引物中含有6个随机引物	phi29 DNA聚合酶	<100 kb	~70%
MALBAC	多重退火环状循环恒温扩增和部分随机引物PCR法结合	27个通用引物和8个随机引物	Bst酶; Taq DNA聚合酶	<2 kb	~90%
LIANTI	线性扩增	包含T7启动子	转录酶; T7 RNA聚合酶	~400 bp	~97%
SISSOR	多重置换恒温扩增法	引物中含有6个随机引物	phi29 DNA聚合酶	<100 kb	~70%

### 3 应用

#### 3.1 在肿瘤研究和临床诊断中的应用

基因组畸变在相同组织学类型的癌症之间不同,没有哪两种肿瘤具有相同的体细胞遗传畸变特征,且不同的肿瘤细胞间存在不同的致病力差异,恶性肿瘤细胞为适应宿主体内复杂的微环境,肿瘤单细胞间也会表现出不同的变异和形态差异<sup>[46]</sup>.单细胞测序技术最早应用于癌症细胞研究领域,通过对肾肿瘤细胞进行单细胞扩增测序发现细胞间的变异频率和病变位置存在很大差异,且没有明显的细胞亚群.经过对单细胞基因组的分析有助于开发更为有效的细胞靶向治疗,也有利于研究肿瘤的发展趋势和转移机制<sup>[47]</sup>.单细胞测序技术还将应用于无创监测,即对血液中的单个循环肿瘤细胞(circulating tumor cell, CTC)进行测序,以跟踪原发性和转移性肿瘤中的突变.多项研究已经表明,肺癌<sup>[9]</sup>、前列腺癌<sup>[48]</sup>和结肠癌<sup>[49]</sup>的CTCs中,原发性和转移性肿瘤中50%以上的突变可以被检测到.通过在治疗过程中多个时间点对CTC进行测序,肿瘤学家可以追踪突变进化,并在耐药性出现之前迅速改变治疗策略.

#### 3.2 在发育生物学研究中的应用

单细胞扩增技术的不断发展使得通过单细胞测序技术研究个体发育成为可能.通过MALBAC方法对亚洲男性的99个精子进行全基因组扩增测序,研究结果显示常染色体非整倍性的增加,交叉频率的降低,并绘制了重组图谱<sup>[50]</sup>.罗伯逊易位(Robertsonian translocation, RT)是男性不育、反复出现妊娠丢失和出生缺陷的常见原因.通过使用组合的单细胞全基因组扩增和测序方案,综合分析了来自两个RT载体患者的88个单精子的染色体拷贝数<sup>[51]</sup>.在每个精子的全基因组基础上鉴定染色体畸变,发现先前报道的染色体间效应可能不存在于RT携带者,为研究携带染色体疾病的患者提供强有力的工具.染色体平衡重排(chromosomal rearrangements, CR)是人类中最常见的遗传异常.Gui等<sup>[52]</sup>研究了CR携带者囊胚内细胞团(inner cell mass, ICM)和滋养外胚层(trophectoderm, TE)的染色体组成之间的一致性程度,结果CR携带者ICM和TE的染色体组成高度一致,而基于卵裂期荧光原位杂交(fluorescence in situ hybridization, FISH)的胚胎植入前遗传学诊断(preimplantation genetic

diagnosis, PGD)不准确.因此,建议使用囊胚TE活组织用于基于二代测序技术(next generation sequencing, NGS)的PGD的检查,以鉴定来自CR携带者胚胎中的染色体非平衡.应用单细胞全基因组扩增技术结合二代测序技术进行胚胎植入前遗传学筛查<sup>[53]</sup>,对废弃胚胎继续培养后得到23个优质囊胚,从中取148个TE细胞,结果发现39个细胞染色体存在异常,来源于14个囊胚.因此,单细胞扩增与二代测序技术结合可用于筛选染色体结构异常的胚胎,改善试管婴儿的质量,减少出生缺陷.

#### 3.3 在微生物研究中的应用

单细胞基因组扩增和测序技术在微生物探索领域也发挥了不可替代的作用,相比其他技术,在难以培养、尚不能培养的微生物以及在环境中物种丰度小的微生物中更具优势.蓝藻在早期地球的历史中发挥了至关重要的作用,但蓝藻多样性并未得到广泛开发.Davison等<sup>[54]</sup>扩增了不同生长温度的蓝藻单细胞基因组,从单细胞获得CRISPR间隔区数据,结合宏基因组宿主和病毒序列,研究天然存在的群体中的宿主——病毒共同进化原因.2015年,张琦等<sup>[55]</sup>利用MDA技术扩增甲藻单细胞基因组,通过对其线粒体细胞色素氧化酶1(Cox 1)基因测序结果与NCBI数据库比对,可快速鉴定,也可用于单细胞的遗传学分析和种群遗传学分析.同年,Martijn等<sup>[56]</sup>发现了一株罕见的立克次氏体菌株,通过单细胞基因组扩增测序,其部分基因组的系统发育和比较分析揭示了存在趋化基因和垂直遗传的鞭毛基因.尽管立克次体科和其他立克次氏体家族已经进行了数十年的广泛研究,但其关于宿主关联的起源和演变的许多细节仍然难以捉摸.这一发现表明,立克次体科的祖先可能具有兼性的细胞内生活方式.这一研究结果强调了单细胞基因组学在研究微生物,特别是对于罕见的微生物细胞多样性和进化中的作用.

#### 3.4 在神经科学研究中的应用

利用单细胞基因组方法对正常人脑中的额叶皮质神经细胞进行分析,发现神经型人诱导的多能干细胞衍生的神经元具有比成纤维细胞更大的CNV,内源性人额叶皮质神经元的单细胞测序显示,13%至41%的神经元具有至少一个兆碱基量级的从头CNV,并且一部分神经元具有高度异常的基因组标记.结果表明,嵌合体CNV在人类神经元中是丰富的<sup>[57]</sup>.Gole等<sup>[58]</sup>利用一种大规模并行聚合酶克

隆方法——微孔位移扩增系统 (MIDAS), 这种方法可以 1~2 Mb 的分辨率检测原代人神经细胞中的单拷贝数变化, 发现人神经细胞至少有一个区域含有百万碱基水平的单拷贝数变异. MIDAS 可以进一步表征许多异质细胞群中的基因组多样性, 促进了从单个大肠杆菌细胞中近乎完整的微生物基因组的从头组装. 近年来, 科学家们通过单细胞测序方法研究神经元细胞, 可以用来研究细胞异质性功能的影响及基因组多样性等.

当前单细胞技术的发展在肿瘤、发育生物学、微生物学和神经科学研究中的应用取得了历史性的进展, 更加丰富了人类对自然界生物进化和多样性的认识. 近年来, 单细胞测序技术在医学领域得到广泛的应用, 在血液系统疾病具有高度异质性, 单细胞测序技术在该方面的应用为治疗和预防此类疾病奠定基础. 第 59 届美国血液学会 (ASH) 报道了多项单细胞测序技术与血液疾病的研究结果, 有利于分析正常细胞与恶性细胞间的差异和癌变机制, 并发现潜在的治疗靶标<sup>[59]</sup>. 在循环肿瘤细胞治疗过程中进行非侵入性诊断和监测, 跟踪 CTC 的反应和进化, 可以不断更新治疗方法, 最后通过更准确地模拟对当前药物抗性的发展, 可以防止治疗抗性<sup>[60]</sup>.

## 4 展望和前瞻

在单个人类细胞的基因组中准确检测变异和远程单倍体类型仍然是非常具有挑战性的研究工作. 常见的方法需要使用 DNA 聚合酶和高通量的短尾 DNA 测序, 对单个细胞的基因组进行广泛的体外扩增. 这些方法有两个显著的缺点, 一是聚合酶复制错误可能会产生成千上万的假阳性, 二是相对较短的序列读取几乎不包含任何单倍体的信息. 当然单细胞技术在各个领域的应用成功与否与单细胞全基因组扩增效果相关, 如表 3 所示不同的 WGA 方法适用于不同的应用研究, 对现有的 WGA 方法进行改进或研发新技术是时下的研究热点. 目前单细胞研究技术存在一些问题, 如扩增偏倚性、非特异性扩增、灵敏度不高、重复性差、操作失误率高、存在外来污染等, 针对这些问题, 如何提高扩增准确性、覆盖率和操作便捷是未来科研工作者的研究方向, 如何在微量样品中达到高灵敏性, 是 WGA 方法在法医学中发挥作用、在发育生物学中朝着无创产检的方向发展的关键. 随着 WGA 方法的不断更新发展, 相信可以在更多的领域发挥作用, 为科学研究提供坚实的基础.

Table 3 Comparison of application of mainstream WGA methods

表3 主流的WGA方法应用比较

WGA方法	优点	缺点	应用
PEP	操作简单, 易改进, 对模板要求较低, 起始模板量低	产量低, 完全随机引物易产生扩增偏差, 保真性差	LOH分析, STR分析等
DOP-PCR	操作简单, 对模板要求低, 起始模板量低	扩增偏差大	FISH, SNP分析, SSCP分析等
MDA	模板要求低, 起始模板量低, 产量高, 忠实性好	对模板质量要求高, 产生非特异性产物	SNV检测, NGS, STR分析, 单细胞测序, Southern印迹, 限制性片段长度多态性分析等
MALBAC	操作简单, 模板要求低, 起始模板量低, 产量高, 结果可靠, 可重复, 扩增均匀度高	模板拷贝数低时易出现扩增偏差和非特异性扩增	染色体分析, CNV检测, SNV检测, CGH, 单细胞测序等
LIANTI	操作简单, 覆盖率高可达97%, 扩增均匀	扩增过程中仍存在偏差, 无法进行微CNVs的精确检测	染色体分析, CNV检测, SNV检测, CGH, 单细胞测序等
SISSOR	更精确的单细胞基因组测序精度, 更长的单倍型	缺乏集成的测序文库制备, 不能并行处理和多个单细胞测序的可扩展性	单细胞测序, 单倍体分析等

LOH分析: 杂合性缺失分析; STR分析: 短片段重复序列分析; FISH: 荧光原位杂交; SNP分析: 单核苷酸多态性分析; SSCP分析: 单链构象多态性分析; SNV检测: 单核苷酸变异检测; CNV检测: 拷贝数变异检测; NGS: 第二代测序技术; CGH: 比较基因组杂交.

## 参 考 文 献

- [1] de Souza N. Single-cell methods. *Nat Methods*, 2010, 7(1): 35
- [2] Navin N, Kendall J, Troge J, *et al.* Tumour evolution inferred by single-cell sequencing. *Nature*, 2011, 472(7341): 90-94
- [3] Kalisky T, Quake S R. Single-cell genomics. *Nat Methods*, 2011, 8(4): 311-314

- [4] Huang L, Ma F, Chapman A, *et al.* Single-cell whole-genome amplification and sequencing: methodology and applications. *Annu Rev Genomics Hum Genet*, 2015, **16**(1): 79-102
- [5] Wang Y, Navin N E. Advances and applications of single-cell sequencing technologies. *Mol Cell*, 2015, **58**(4): 598-609
- [6] Lasken R S. Single-cell genomic sequencing using multiple displacement amplification. *Curr Opin Microbiol*, 2007, **10**(5): 510-516
- [7] Blainey P C. The future is now: single-cell genomics of bacteria and archaea. *Fems Microbiol Rev*, 2013, **37**(3): 407-427
- [8] Yan L., Yang M, Guo H, *et al.* Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat Struct Mol Biol*, 2013, **20**(9): 1131-1139
- [9] Ni X, Zhuo M, Su Z, *et al.* Reproducible copy number variation patterns among single circulating tumor cells of lung cancer patients. *Proc Natl Acad Sci U S A*, 2013, **110**(52): 21083-21088
- [10] Vogelstein B, Papadopoulos N, Velculescu V E, *et al.* Cancer genome landscapes. *Science*, 2013, **339**(6127): 1546-1558
- [11] Hou Y, Fan W, Yan L, *et al.* Genome analyses of single human oocytes. *Cell*, 2013, **155**(7): 1492-1506
- [12] Gawad C, Koh W, Quake S R. Single-cell genome sequencing: current state of the science. *Nat Rev Genet*, 2016, **17**(3): 175-188
- [13] Yan L, Huang L, Xu L, *et al.* Live births after simultaneous avoidance of monogenic diseases and chromosome abnormality by next-generation sequencing with linkage analyses. *Proc Natl Acad Sci U S A*, 2015, **112**(52): 15964-15969
- [14] Zhang K, Martiny A C, Reppas N B, *et al.* Sequencing genomes from single cells by polymerase cloning. *Nat Biotechnol*, 2006, **24**(6): 680-686
- [15] Marcy Y, Ishoev T, Lasken R S, *et al.* Nanoliter reactors improve multiple displacement amplification of genomes from single cells. *PLoS Genet*, 2007, **3**(9): 1702-1708
- [16] Streets A M, Zhang X, Cao C, *et al.* Microfluidic single-cell whole-transcriptome sequencing. *Proc Natl Acad Sci U S A*, 2014, **111**(19): 7048-7053
- [17] Stepanauskas R, Sieracki M E. Matching phylogeny and metabolism in the uncultured marine bacteria, one cell at a time. *Proc Natl Acad Sci U S A*, 2007, **104**(21): 9052-9057
- [18] Ma Z, Zhou Y, Collins D J, *et al.* Fluorescence activated cell sorting *via* a focused traveling surface acoustic beam. *Lab Chip*, 2017, **17**(18): 3176-3185
- [19] Shapiro E, Biezuner T, Linnarsson S. Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat Rev Genet*, 2013, **14**(9): 618-630
- [20] Bonner R F, Emmert-Buck M, Cole K, *et al.* Laser capture microdissection: molecular analysis of tissue. *Science*, 1997, **278**(5342): 1481-1483
- [21] Frumkin D, Wasserstrom A, Itzkovitz S, *et al.* Amplification of multiple genomic loci from single cells isolated by laser microdissection of tissues. *BMC Biotechnol*, 2008, **8**(1): 1-16
- [22] Zhang L, Gui X, Schmitt K, *et al.* Whole genome amplification from a single cell: Implications for genetic analysis. *Proc Natl Acad Sci USA*, 1992, **89**(13): 5847-5851
- [23] Telenius H, Carter N P, Bebb C E, *et al.* Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single degenerate primer. *Genomics*, 1992, **13**(3): 718-725
- [24] Van K, Kang Y J, Shim S R, *et al.* Genome-wide scan of the soybean genome using degenerate oligonucleotide primed PCR: an example for studying large complex genome structure. *Genes & Genomics*, 2012, **34**(5): 467-474
- [25] Blagodatskikh K A, Kramarov V M, Baronva E V, *et al.* Improved DOP-PCR (iDOP-PCR): A robust and simple WGA method for efficient amplification of low copy number genomic DNA. *PLoS One*, 2017, **12**(9): e0184507
- [26] Viguera E, Canceill D, Ehrlich S D. Replication slippage involves DNA polymerase pausing and dissociation. *The Embo*, 2001, **20**(10): 2587-2595
- [27] Dean F B, Hosono S, Fang L, *et al.* Comprehensive human genome amplification using multiple displacement amplification. *Proc Natl Acad Sci USA*, 2002, **99**(8): 5261-5266
- [28] Zong C, Lu S, Chapman A R, *et al.* Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. *Science*, 2012, **338**(6114): 1622-1626
- [29] Hosono S, Farqi A F, Dean F B, *et al.* Unbiased whole-genome amplification directly from clinical samples. *Genome Res*, 2003, **13**(5): 954-964
- [30] Yilmaz S, Allgaier M, Hugenholtz P. Multiple displacement amplification compromises quantitative analysis of metagenomes. *Nat Methods*, 2010, **7**(12): 943-944
- [31] Leung K, Klaus A, Lin B K, *et al.* Robust high-performance nanoliter-volume single-cell multiple displacement amplification on planar substrates. *Proc Natl Acad Sci USA*, 2016, **113**(30): 8484-8489
- [32] Lasken R S. Single-cell sequencing in its prime. *Nat Biotechnol*, 2013, **31**(3): 211-212
- [33] Fu Y, Li C, Lu S, *et al.* Uniform and accurate single-cell sequencing based on emulsion whole-genome amplification. *Proc Natl Acad Sci U S A*, 2015, **112**(38): 11923-11928
- [34] Bourcy C F A, Vlamincik I D, Kanbar J N, *et al.* A quantitative comparison of single-cell whole genome amplification methods. *PLoS One*, 2014, **9**(8): e105585
- [35] Blainey P C, Quake S R. Digital MDA for enumeration of total nucleic acid contamination. *Nucleic Acids Research*, 2010, **39**(4): e19
- [36] Chen C, Xing D, Tan L, *et al.* Single-cell whole-genome analyses by linear amplification *via* transposon insertion (LIANTI). *Science*, 2017, **356**(6334): 189-194
- [37] Goryshin I Y, Reznikoff W S. Tn5 *in vitro* transposition. *Biological Chemistry*, 1998, **273**(13): 7367-7374
- [38] Gelder R N V, Zastrow M E V, Yool A, *et al.* Amplified RNA synthesized from limited quantities of heterogeneous cDNA. *Proc Natl Acad Sci U S A*, 1990, **87**(5): 1663-1667
- [39] Reuter J A, Spacek D V, Pai R K, *et al.* Simul-seq: combined DNA and RNA sequencing for whole-genome and transcriptome

- profiling. *Nat Methods*, 2016, **13**(11): 953-958
- [40] Zahn H, Steif A, Laks E, *et al.* Scalable whole-genome single-cell library preparation without preamplification. *Nat Methods*, 2017, **14**(2): 167-173
- [41] Chu W K, Edge P, Lee H S, *et al.* Ultraaccurate genome sequencing and haplotyping of single human cells. *Proc Natl Acad Sci U S A*, 2017, **114**(47): 12512-12517
- [42] 聂同钢, 马妍, 徐晓宁, 等. 全基因组扩增在微量检材 DNA 分型中的应用. *中国法医学杂志*, 2017, **32**(2): 175-181  
Nie T G, Ma Y, Xu X N, *et al.* *Chinese Journal of Forensic Medicine*, 2017, **32**(2): 175-181
- [43] Liu W Q, Zhang H M, Hu D, *et al.* The performance of MALBAC and MDA methods in the identification of concurrent mutations and aneuploidy screening to diagnose beta-thalassaemia disorders at the single- and multiple-cell levels. *J Clin Lab Anal*, 2018, **32**(2): 1-8
- [44] He F, Zhou W, Cai R, *et al.* Systematic assessment of the performance of whole-genome amplification for SNP / CNV detection and beta-thalassaemia genotyping. *J Hum Genet*, 2018, **63**(4): 407-416
- [45] Peters B A, Kermani B G, Sparks A B, *et al.* Accurate whole-genome sequencing and haplotyping from 10 to 20 human cells. *Nature*, 2012, **487**(7406): 190-195
- [46] Lipinski K A, Barber L J, Davies M N, *et al.* Cancer evolution and the limits of predictability in precision cancer medicine. *Trends Cancer*, 2016, **2**(1): 49-63
- [47] Xu X, Hou Y, Yin X, *et al.* Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. *Cell*, 2012, **148**(5): 886-895
- [48] Lohr J G, Adalsteinsson V A, Cibulskis K, *et al.* Whole-exome sequencing of circulating tumor cells provides a window into metastatic prostate cancer. *Nat Biotechnol*, 2014, **32**(5): 479-484
- [49] Heitzer E, Auer M, Gasch C, *et al.* Complex tumor genomes inferred from single circulating tumor cells by array-CGH and next-generation sequencing. *Cancer Res*, 2013, **73**(10): 2965-2975
- [50] Lu S, Zong C, Fan W, *et al.* Probing meiotic recombination and aneuploidy of single sperm cells by whole-genome sequencing. *Science*, 2012, **338**(6114): 1627-1630
- [51] Sha Y W, Sha Y K, Ji Z Y, *et al.* Comprehensive genome profiling of single sperm cells by multiple annealing and looping-based amplification cycles and next-generation sequencing from carriers of robertsonian translocation. *Ann Hum Genet*, 2017, **81**(2): 91-97
- [52] Gui B, Yao Z, Li Y, *et al.* Chromosomal analysis of blastocysts from balanced chromosomal rearrangement carriers. *Reproduction*, 2016, **151**(4): 455-464
- [53] 朱玉蓉, 刘春玲, 杨胜, 等. 应用全基因组扩增结合二代测序技术对废弃胚胎进行植入前遗传学筛查. *中华医学遗传学杂志*, 2018, **35**(3): 337-341  
Zhu Y R, Liu C L, Yang S, *et al.* *Chinese Journal of Medical Genetics*, 2018, **35**(3): 337-341
- [54] Davison M, Hall E, Zare R, *et al.* Challenges of metagenomics and single-cell genomics approaches for exploring cyanobacterial diversity. *Photosynthesis Research*, 2015, **126**(1): 135-146
- [55] 张琦, 刘永健, 柳圭泽. 海洋甲藻的单细胞 PCR 技术初步研究. *海洋环境科学*, 2015, **34**(04): 611-615  
Zhang Q, Liu Y J, Liu G Z. *Marine Environmental Science*, 2015, **34**(04): 611-615
- [56] Martijn J, Schulz F, Zaremba-Niedzwiedzka K, *et al.* Single-cell genomics of a rare environmental alphaproteobacterium provides unique insights into Rickettsiaceae evolution. *Isme J*, 2015, **9**(11): 2373-2385
- [57] McConnell M J, Lindberg M R, Brennand K J, *et al.* Mosaic copy number variation in human neurons. *Science*, 2013, **342**(6158): 632-637
- [58] Gole J, Gore A, Richards A, *et al.* Massively parallel polymerase cloning and genome sequencing of single cells using nanoliter microwells. *Nat Biotechnol*, 2013, **31**(12): 1126-1132
- [59] 宋鸽, 蔺亚妮, 汝昆. 单细胞测序在血液系统疾病中的应用进展. *白血病·淋巴瘤*, 2018, **27**(2): 73-75  
Song G, Lin Y N, Ru K. *Leukemia·Lymphoma*, 2018, **27**(2): 73-75
- [60] Tsoucas D, Yuan G C. Recent progress in single-cell cancer genomics. *Curr Opin Genet Dev*, 2017, **42**(1): 22-32

## Single Cell Whole Genome Amplification Technology and Application\*

XU Xiao-Li, WU Ling-Juan, YAN Ren-Xiang\*\*

(College of Biological Sciences and Engineering, Fuzhou University, Fuzhou 350100, China)

**Abstract** Cells in the same tissue often have similar structures and functions. However, it is found that each cell is heterogeneous by sequencing cells. Single-cell whole-genome amplification technology is the premise of single-cell sequencing. This technology can be used to reveal the differences of single-cell genomic structures and it also plays an important role in tumor research, developmental biology, and microbial research, and it has already become one of hotspots in life science research technology. The difficulty of single-cell whole genome amplification techniques lies in the isolation of single cells and the amplification of whole genomes. This paper introduces the popular single-cell separation technology and single-cell whole genome amplification technology and also compares the advantages and disadvantages of each technology. The application of the technology in tumor research, developmental biology and microbiology research was also emphatically discussed.

**Key words** single cell separation technique, single cell whole genome amplification technique, comparison, application.

**DOI:** 10.16476/j.pibb.2018.0259

---

\* This work was supported by a grant from The National Natural Science Foundation of China (31500673).

\*\* Corresponding author.

Tel: 86-591-22866376, E-mail: yanrenxiang@fzu.edu.cn

Received: December 28, 2018 Accepted: March 5, 2019