

## 单分子纳米孔测序技术及其应用研究进展

曹影<sup>1</sup>, 李伟<sup>2</sup>, 褚鑫<sup>3</sup>, 吴珂<sup>1</sup>, 刘海舟<sup>1</sup>, 刘翟<sup>1,4</sup>

1 中国科学院武汉病毒研究所, 湖北 武汉 430071

2 中国科学院微生物研究所, 北京 100101

3 中国科学院重大科技任务局, 北京 100864

4 中国科学院大学, 北京 100049

曹影, 李伟, 褚鑫, 等. 单分子纳米孔测序技术及其应用研究进展. 生物工程学报, 2020, 36(5): 811–819.

Cao Y, Li W, Chu X, et al. Research progress and application of nanopore sequencing technology. Chin J Biotech, 2020, 36(5): 811–819.

**摘要:** 测序技术在通量和成本方面有了较大的改进, 以单分子纳米孔测序技术为代表的第三代测序技术更是以其超长读长、实时检测和可以直接检测碱基甲基化修饰等优势在医学及生命科学等领域作出了较大贡献。文中就单分子纳米孔测序技术的原理进行了简要描述, 并对其在临床、动物、植物、细菌及病毒等领域的应用和其未来的发展方向进行了讨论。

**关键词:** 纳米孔测序技术, 临床研究, 动物基因组, 植物基因组, 细菌基因组, 病毒基因组

## Research progress and application of nanopore sequencing technology

Ying Cao<sup>1</sup>, Wei Li<sup>2</sup>, Xin Chu<sup>3</sup>, Ke Wu<sup>1</sup>, Haizhou Liu<sup>1</sup>, and Di Liu<sup>1,4</sup>

1 Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan 430071, Hubei, China

2 Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China

3 Bureau of Major R&D Programs, Chinese Academy of Sciences, Beijing 100864, China

4 University of Chinese Academy of Sciences, Beijing 100049, China

**Abstract:** Sequencing technology has been greatly improved in terms of throughput and cost. The single-molecule nanopore DNA sequencing, one of the major branches of the third-generation sequencing technology, has made great contributions in the fields of medicine and life sciences due to its advantages of ultra-long reading length, real-time detection and direct detection of base methylation modification, *etc.* This article briefly describes the principle of nanopore sequencing technology, and discusses its application in clinical, animal, plant, bacterial and virus fields and its future development direction.

**Keywords:** nanopore sequencing technology, clinical research, animal genome, plant genome, bacterial genome, viral genome

**Received:** August 15, 2019; **Accepted:** February 24, 2020

**Supported by:** Beijing Major Science and Technology Projects (No. D171100002217001).

**Corresponding authors:** Di Liu. Tel: +86-27-87198739; E-mail: liud@wh.iov.cn

Haizhou Liu. Tel: +86-27-87198739; E-mail: liuhz@wh.iov.cn

北京市重大科技专项 (No. D171100002217001) 资助。

系统发育和遗传变异分析是我们认识和研究物种的重要方法,而获得高质量的基因组序列数据是开展相关分析的必要前提。自测序技术出现以来,核苷酸测序技术经历了飞速发展,尤其是以高通量著称的第二代测序技术的出现,更使得其越来越成为一种高效、快速、低成本的研究检测手段,被普遍应用于分子生物学、进化生物学、宏基因组学、医学以及法医学等领域,并取得了大量的成果,极大拓展了对物种的认知<sup>[1-7]</sup>。然而,由于测序技术自身的限制,第一代测序技术不适合多样本的大规模测序;第二代测序技术也因为读长较短、高 GC 偏性而在应用上有所局限。

以牛津纳米孔技术公司 (Oxford Nanopore Technologies, ONT) 的单分子纳米孔测序技术 (Single-molecule nanopore DNA sequencing) 和太平洋生物科学公司 (Pacific Biosciences, PacBio) 的单分子实时测序技术 (Single molecule real-time sequencing, SMRT) 为代表的第三代测序技术,因其集高通量、超长读长和可以直接检测碱基甲基化修饰等优势,越来越多地被应用于基因组装配、表观遗传标记、转录组学及宏基因组学等领

域<sup>[8-11]</sup>。与 SMRT 技术相比,单分子纳米孔测序技术不仅在读长的长度方面更具优势 (单读长可达 2 Mb)<sup>[12-13]</sup>,而且其设备体积较小,便于携带,因此在极端条件和实验条件相对落后的疫情地区进行样品的实时检测方面具有明显的优势。文中就单分子纳米孔测序技术的原理及其应用进行概述,并对其存在的不足与未来发展方向进行讨论。

## 1 测序技术的发展

### 1.1 第一代测序技术的原理及其特点

1977年,化学降解法 (Maxam-Gilbert method)<sup>[14]</sup>和双脱氧链终止法 (Chain termination method)<sup>[15]</sup>的相继出现标志着第一代测序技术的诞生 (图 1)。化学降解法涉及到 5 个各自独立的反应,是直接对目的 DNA 分子进行处理,使用不同的化学试剂将脱氧核糖核酸 (Deoxyribonucleic acid, DNA) 分别打碎成不同长度的片段后,通过聚丙烯酰胺凝胶电泳对不同长度的片段进行分离。该操作可以避免合成时造成的错误,但相对双脱氧链终止法较为复杂,逐渐被后者取代。双脱氧链终止法的原理是

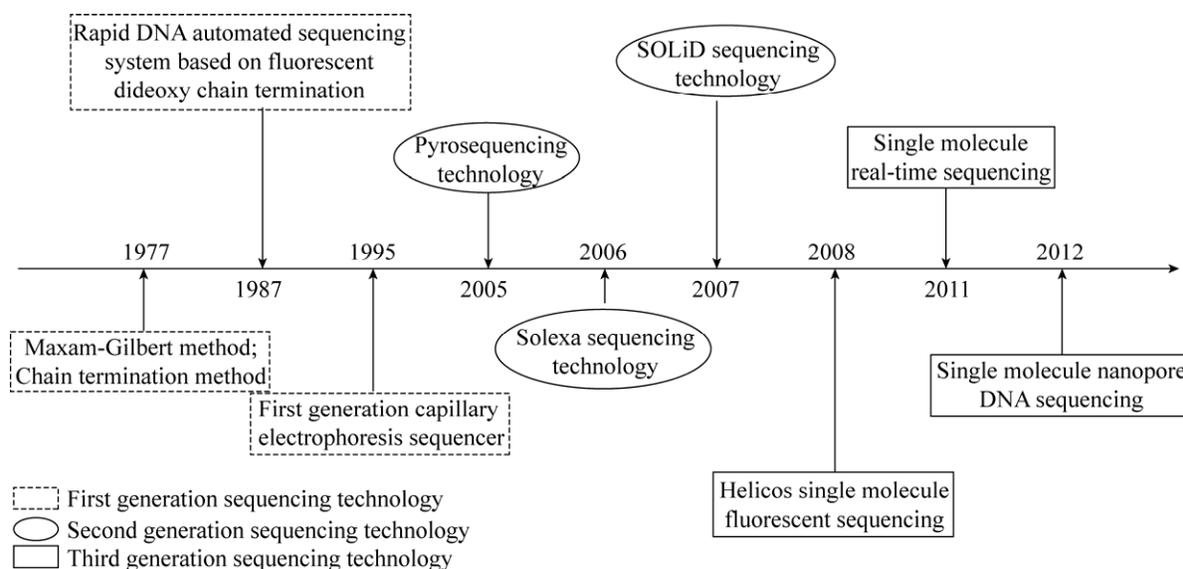


图 1 测序技术的发展

Fig. 1 Development of sequencing technology.

借助双脱氧核苷三磷酸 (Dideoxynucleoside triphosphate, ddNTP) 和脱氧核糖核苷三磷酸 (Deoxy-ribonucleoside triphosphate, dNTP) 具有相似结构这一特性, 其可以在聚合酶的作用下结合到目的模板上, 但因缺少一个羟基而使合成反应终止。在 4 个独立的反应中分别加入一种独特的 ddNTP 和 4 种 dNTP, 后同样经过聚丙烯酰胺凝胶电泳将不同长度的片段分离, 根据末端 ddNTP 的不同得到 DNA 分子的碱基片段。双脱氧链终止法是第一代测序技术的主要代表, 至今仍在普遍使用; 之后出现的商用荧光双脱氧测序技术就是在此基础上的改进。

第一代测序技术产生的读长 (Reads) 长度为 700–900 bp (表 1), 准确率高达 99.999%<sup>[16]</sup>, 设备运行时间短, 适用于通量要求低的快速研究项目; 但一个反应仅能得到一条读长, 且大规模测序成本依然较高, 一般适用于少量样本的小规模测序。

## 1.2 第二代测序技术原理及其特点

第二代测序技术的原理主要为边合成边测序 (Sequencing by synthesis, SBS) 和连接法测序 (Sequencing by Ligation, SBL)。边合成边测序的代表为罗氏旗下的 454 测序技术和 Illumina 公司的 Solexa 测序技术, 两者检测重点不同。454 测序技术主要通过将合成反应中产生的焦磷酸基团 (PPi) 转化为可见光信号从而判断反应是否发生, 在每轮反应中只会加入一个 dNTP, 若反应发生, 通过光学系统即可获特异的检测峰, 峰值的高低和反应中匹配的碱基数成正比; 其 GS FLX+ 系统

最大的读长可达 1 000 bp<sup>[17]</sup>, 测序准确性与 Sanger 测序相当, 然而, 由于该平台是对产生的焦磷酸进行检测, 当序列中存在同聚物时其准确性会受影响<sup>[18]</sup>。Solexa 测序技术是针对 dNTP 本身进行检测, 在每个 dNTP 上均连接不同颜色的荧光基团和可阻断基团, 随着合成反应的进行, 通过检测加入 dNTP 的荧光基团的颜色推断 DNA 序列的组成, 每次延伸反应中 4 种 dNTP 的浓度均匀, 因此可有效地避免掺入错误<sup>[19]</sup>。

与边合成边测序的思想不同, 连接法测序不通过聚合酶链式反应 (Polymerase chain reaction, PCR) 以待测序列作为模板链合成新的 DNA 链, 而是利用 DNA 连接酶将一段寡核苷酸探针连接到模板上进行测序。连接法测序的代表是美国应用生物系统公司 (Applied Biosystems Inc., ABI) 的 SOLiD 测序平台, 其主要利用 DNA 连接酶可以将两条相互匹配的核苷酸序列连接在一起的原理来识别 DNA 序列中特定位置的核苷酸。通过一段已知的序列与待测序列进行互补, 并且已知的两个碱基对应一个荧光信号, 随着连接反应的进行测序仪记录荧光染料信号, 并断裂连接后的互补序列准备进行第二轮连接, 如此反复, 每一个位点都会被检测两次, 可以显著校正误差<sup>[20]</sup>。

三种主要的第二代测序平台在建库之初都需要对目的片段进行打断 (读长长度较短, Illumina 平台为 100–300 bp, 454 平台平均长度也只在 500 bp 左右), 且均需对信号进行放大才能达到检测要求, 因此会存在一些丰度较低的序列无法被大量扩增而造成信息的丢失, 且 PCR 过程中也

表 1 测序技术特点比较

Table 1 Comparison of sequencing technology characteristics

	First-generation sequencing technology	Next-generation sequencing technology	Third-generation sequencing technology
Reads length	700–900 bp	50–500 bp	<2 Mb
Advantage	Long reads; Accurate	High through; Inexpensive	Ultra-long reads; High through
Disadvantage	Low through; Expensive	Short reads; Equipment expensive	Poor accuracy; Analysis software
Genotype	Major genotype	Population characteristics of variation sites	All genotypes in the population

可能引入错配碱基；与第一代测序技术相比，第二代测序技术通量提高了成千上万倍，且单条序列成本非常低廉，适用于大规模高通量的测序需求，现在仍是科研市场的主力平台。

## 2 单分子纳米孔测序技术原理及特点

牛津纳米孔科技公司的纳米孔测序平台的核心是一个带有 2 048 个纳米孔、由专用集成电路控制的测序芯片 (Flow cell)。单链 DNA 分子穿过纳米孔时，由于不同的碱基的形状大小有差异，与孔内环糊精分子发生特异性反应而引起电阻变化。纳米孔的两侧有一恒定电压，因此可以检测到纳米孔中电流的变化，从而反映出通过纳米孔的 DNA 分子的碱基排列情况。自 2017 年以后，该纳米孔测序平台主要使用 1D 和 1D<sup>2</sup> 两种测序策略，其中 1D 测序原理是：基因组 DNA 或 cDNA 分子经接头帮助到达纳米孔附近，在解旋酶的作用下双链 DNA 分子解开为单链，通过孔道蛋白；传感器检测到不同核苷酸通过所引起的电流变化的差异并将其转换为电信号；最后，根据电信号变化的频谱，应用模式识别算法得到碱基类型<sup>[21]</sup>。与 1D 测序策略不同的是，1D<sup>2</sup> 测序策略在建库时会在两条 DNA 分子上加上一种特殊的接头，使得在读取模板链的同时互补链可以附着到膜上，在第一条链离开纳米孔后不久，互补链就有一定概率接着被测序，两条链的数据相互校正，可以帮助提高测序的准确率。

第三代测序技术在测序过程中不需要通过 PCR 进行信号放大，因此避免了 PCR 反应过程中引入的碱基错配；在整个反应中也不涉及酶的催化反应，理论上只要核酸提取步骤可以得到足够长度的序列，测序步骤就可以对其进行检测；此外，由于甲基化等修饰前后的核苷酸所引起的电阻变化是不同的，所以该测序平台可通过对电信号的识别来判断碱基的甲基化修饰情况。因此，单分子纳米孔测序技术具有高通量、超长读长、

可以直接检测碱基甲基化修饰和体积较小便于携带等优势，在动物、植物、细菌、病毒等的研究中均具有较为广阔的应用空间。

## 3 单分子纳米孔测序技术应用

### 3.1 纳米孔测序技术在临床研究中的应用

近年来，第二代测序技术在遗传病的临床研究及诊断中的应用大大提高了致病基因检测的效率和准确性，使人们对自身基因及其功能有了更深入的认识。但由于第二代测序读长较短且受 GC 含量、基因组重复序列等影响，只能解决部分遗传病基因的检测，在多样性和复杂性疾病致病原因的检出方面作用有限。单分子纳米孔测序技术的读长可达 2 Mb，在基因组结构变异、点突变等各种致病原因的检测方面有较大的优势，很好地弥补了一代和二代测序技术的不足。

目前，越来越多的科研及临床工作者致力于将单分子纳米孔测序技术应用到遗传病基因检测中。例如，Roeck 等<sup>[22]</sup>分别用 Illumina Solexa 和单分子纳米孔测序技术检测了大量阿尔茨海默病患者 ABCA7 基因的提前终止密码子 (Premature termination codon, PTC) 突变情况，发现单分子纳米孔技术比第二代测序技术多检测出 7 个 PTC 突变，解释了不同程度无义介导的 mRNA 降解和转录调控影响 ABCA7 基因的表达，进而影响疾病的严重程度，这一发现可能为阿尔茨海默病的干预性治疗创造机会。此外，Stancu 等<sup>[23]</sup>利用纳米孔测序技术检测病人染色体碎裂重排病，结果显示纳米孔测序技术与 Illumina 测序技术相比，凭借其超长读长的优势在结构变异检测应用中更胜一筹。

### 3.2 纳米孔测序技术在动物基因组研究中的应用

近年来越来越丰富的动物基因组研究极大地推动了人们对于物种演化、病虫害防治及濒危动物的保护等方面的认知及研究。Tyson 等<sup>[24]</sup>用

ONT 科技公司的 MinION 平台对秀丽隐杆线虫的基因组进行了测序, 其组装后的 contig N50 达 1.22 Mb, 含有 145 个重叠群的秀丽隐杆线虫基因组, 覆盖了参考基因组的 99%, 从装配的基因组中鉴定出了乙醛诱导的染色体重排和质粒的插入, 并且首次证明 MinION 测序可用于大于 100 Mb 基因组的装配以及基因组重排的结构分析。Jansen 等<sup>[25]</sup>使用 MinION 平台对欧洲鳗鲡的基因组进行了测序, 组装后基因组中的片段比之前的草图低两个数量级, 极大地提高了组装的准确性。

Bickhart 等使用以纳米孔测序为主的方法, 对家山羊 *Capra hircus* 进行测序, 与之前的基因组相比连续性提高约 400 倍, 并且更好地展现了反刍动物中以重复家族和免疫基因复合体为代表的长度超过 1 kb 的重复结构, 利于鉴定后代品种的适应性变异, 同时为其他反刍动物的遗传学研究提供参考<sup>[26]</sup>。Jain 等<sup>[27]</sup>使用 MinION 报告了人 GM12878 Utah/Ceph 细胞系基因组的测序和组装, 最终产生的数据测序深度为 30 倍, 组装的基因组大小为 28.67 亿碱基, 占参考基因组的 85.8%; 在使用短读长数据进行校正后, 装配精度超过 99.8%。可以发现, 利用纳米孔测序技术能够完整地组装和定位较大片段的主要组织相容性复合物 (MHC) 基因座, 进一步证实了纳米孔测序技术在基因组重测序方面的应用潜力, 为结构变异和部分表观遗传修饰的检测提供了有效的技术手段。

### 3.3 纳米孔测序技术在植物基因组研究中的应用

自 2000 年第一个植物基因组拟南芥被破译以来<sup>[28]</sup>, 近 20 年里有 300 多种植物基因组被相继测序并发布, 覆盖了各种粮食、油料、蔬菜、药用及果类作物。植物基因组的大小从几十 Mb 到一百多 Gb 不等, 其多倍性、高杂合以及多重复区域等问题一直是阻碍人们获得高质量植物基因组序列的主要原因。大部分 Illumina 测序得到

的读长较短, 组装的基因组连续性欠佳, 454 测序技术虽然在读长上较其他第二代测序技术有明显的优势, 但是还是远短于单分子纳米孔测序技术产生的读长。发表在 *Nature Plant* 上的一篇文章<sup>[29]</sup>巧妙运用 MinION/PromethION 结合 Bionano 光学图谱和 Illumina 测序平台的测序策略, 将白菜型油菜、甘蓝型油菜、裂果蕪基因组组装到 contig N50>5 Mb, 并组装出包含代表全部染色体或染色体臂的 scaffolds。同时, 研究者还与已经发布的组装数据进行了比较, 发现该研究组装得到的 contig N50 是以前报道的 100–450 倍; 最重要的是, 相较于已发布数据组装出的 446.8 Mb 基因组, 该研究组装得到了 528.8 Mb 大小的基因组, 填补了之前研究没有覆盖到的 82 Mb 的区域。可以发现, 在纳米孔准确度还不是特别高的今天, 多平台测序结合的策略能在保证基因组测序结果可靠性的同时, 为我们带来单一测序手段所无法发现的信息, 不仅为基因组测序研究带来重大突破, 更为人们进一步改良经济作物提供了可能的生物途径, 也必将促进人类健康的研究。

### 3.4 纳米孔测序技术在细菌基因组研究中的应用

细菌表型的差异不像动物与植物那样明显, 因此, 通过全基因组序列比较对细菌进行分类至关重要。Quick 等<sup>[30]</sup>于 2014 年首先使用纳米孔测序技术对大肠杆菌 K-12 的标准菌株 MG1655 进行测序; 随后 Loman 等<sup>[31]</sup>于 2015 年对 Quick 测序数据的组装过程进行了优化, 证明单分子纳米孔测序技术产生的数据质量足以完整地组装细菌基因组并准确地重建基因顺序和方向。Tanaka 等<sup>[32]</sup>测试了纳米孔测序数据对弧菌科 (Vibrionaceae) 细菌基因组分类的能力, 并比较了弧菌科下属 5 个种的菌株仅使用纳米孔测序结果进行组装和基于纳米孔测序加 Illumina Solexa 测序混合组装的结果, 通过总体基因组相关性指数和多位点序列分析的比较发现, 仅使用纳米孔测序结果进行

组装的错误不影响平均核苷酸同一性。Li 等则通过提高纳米孔测序读长的长度和准确度来进行基于 16S 核糖体 RNA (16S ribosomal ribonucleic acid, 16S rRNA) 的细菌分析, 得到物种的准确分类以及在 0.1% 丰度下对物种的鉴定和相对丰度的定量分析<sup>[33]</sup>。

纳米孔测序的长读长在检测基因组中的重复序列与结构变异等方面也拥有明显优势。Kranz 等<sup>[34]</sup>有机结合了纳米孔测序技术的长读长和 Illumina 测序高准确度的特点, 发现了在野生型菌株中存在而在基因工程改造菌株中尚未发现的 1 420 bp 长转座子侧翼和含 ORF 的序列, 并成功检测到包括 3 种已知的基因工程改造在内的 7 种结构变异。Giordano 等<sup>[35]</sup>使用包括 MinION 在内的 3 种测序平台, 对酿酒酵母 S288C 菌株进行了测序并比较了不同软件对基因组的组装效果。在测序深度为 31X 时, 只使用纳米孔测序数据进行组装准确率为 98%, 而结合 Illumina 测序数据进行校正时, 可使准确率提高到 99.98%, 该研究证明了纳米孔测序技术在微生物中应用的可能性, 为其他微生物基因组相关研究提供了很好的示范作用。

### 3.5 纳米孔测序技术在病毒基因组研究中的应用

对病毒研究来说, 纳米孔测序技术超长的读长使其可以覆盖绝大多数病毒基因组, 在病毒相关疫情监测、病毒全基因组测序、病毒的进化与遗传变异等方面的研究拥有显著优势。在几内亚埃博拉疫情爆发期间, Joshua 等<sup>[36]</sup>基于 MinION 平台轻便便携的特点设计了一整套仅一张普通实验台即可容纳的测序系统, 对疫情进行实时基因组监测。即使在当地资源有限的条件下, 也可以在接收阳性埃博拉样品后不到 24 h 产生结果, 整个测序过程仅需 15–60 min, 为疫情的监测提供了一个良好的示范。2019 年 1 月, 尼日利亚 20 个州和联

邦首都直辖区拉沙热 (Lassa fever) 疫情进入紧急状态。英格兰公共卫生署的 Kafetzopoulou 等<sup>[37]</sup>利用纳米孔测序技术, 通过对 36 个基因组及 120 份临床样本进行实时宏基因组测序分析, 揭示了拉沙热病毒 (LASV) 的多样化及其与早期发现的毒株的系统发育相关性, 该研究在较短时间内缓解了人们对拉沙热在人际间广泛传播的恐惧, 使公共卫生资源得到了合理分配, 还进一步指出 LASV 防治重点是加强社区鼠类控制、环境卫生和食品储存安全, 是纳米孔测序技术应用的在疫情诊断中应用的经典案例。

### 3.6 纳米孔测序技术在 RNA 直接测序中的应用

不经反转录、无需扩增的 RNA 直接测序能获得全长的链特异性 RNA, 无测序偏好性, 并同时记录碱基修饰, 使我们可以研究原始状态下的 RNA 病毒基因组的本质。RNA 直接测序方法首次被描述<sup>[38]</sup>以来, 已被相继用于流感病毒基因组的直接测序和冠状病毒的研究<sup>[11,39]</sup>。Adrian 等在冠状病毒的研究中指出, 直接 RNA 测序具有重建大量全长 RNA 病毒基因组克隆型的潜力。RNA 直接测序还可以对在病毒生命周期中起重要作用的病毒信使核糖核酸 (messenger RNA, mRNA) 等进行分析, 可以识别和量化剪接类型并进行碱基修饰检测, 而这些在以往的方法中是无法做到的。

RNA 甲基化修饰是表观转录组学的重要组成部分, 其受到甲基转移酶、去甲基化酶等的动态调控, 与基因表达调控密切相关。在过去的几年中, 转录组学领域经历了巨大的发展, 主要的限制是缺乏通用的方法在转录组范围内定位 RNA 修饰。Huanle 等<sup>[40]</sup>借助纳米孔测序平台对 RNA 进行直接测序, 高精度地检测出 N<sup>6</sup>-甲基腺苷 (m<sup>6</sup>A) RNA 修饰, 准确率达 90%。该结果通过直接 RNA 测序鉴定 RNA 修饰的概念验证, 并为将来探索其他 RNA 修饰提供借鉴, 为研究 RNA 修饰在其天然 RNA 环境中的生物学作用开辟了途径。

## 4 总结与展望

单分子纳米孔测序技术以其特有的优势在临床、动物、植物、微生物及病毒基因组研究中发挥着不可或缺的作用。与第二代测序平台相比,纳米孔测序技术不仅可以实时获取基因组序列,产生的超长读长也使其在全基因组范围内结构变异的检测具有更可靠的结果;而且,它利用电信号的独特的测序原理使其可以检测到4种胞嘧啶的碱基修饰,为表观遗传学研究奠定了基础。与太平洋生物科学公司的RS II平台相比,单分子纳米孔测序技术无需DNA聚合酶的链式反应,所以其测序读长不受DNA聚合酶活性的影响。这一特点使其在重测序读长方面更具优势,也为我们获得更大的全长基因组并对其进行研究提供了可能。

然而,较高的错误率仍然是阻碍其应用的关键。目前MinION测序仪的单碱基准确率约85%<sup>[41]</sup>,修正后的一致性序列的准确率约97%<sup>[27]</sup>。对于动植物这样基因组较大且变异率相对较低的生物来说,增加测序深度可达到一定的纠错效果,对研究结果的影响不是特别大。但是临床检测通常对数据的准确率要求特别严苛,再者,病毒基因组较小且变异率较高,因此错误率造成的影响是不可忽视的。已发表的研究中,多是将病毒“准种”作为一个整体来看待,得到其群体水平的一致性(Consensus)序列,从而对其进行进一步的系统进化方面的分析。对于病毒来说,我们得到的往往是一个高度相关但又不完全相同的动态种群,利用一致性序列来代表这个动态种群的全部信息显然不够准确,因此需要我们借助更高精度的测序方法对群体中非主要基因组进行研究。

从纳米孔测序技术在动物、植物、微生物等多个研究对象的应用中我们可以发现,利用多测序平台对对象进行综合分析是现阶段所用的主要测序手段,并且已经取得了很好的成果<sup>[23,29,42]</sup>。然而,提高纳米孔测序平台的准确度,使其可以

更好为科研工作者服务仍然是我们努力的方向。单分子纳米孔测序技术需要向着更高通量、更高的准确率以及更高的自动化程度方向发展。随着纳米孔测序技术准确率的不断提高,其将在科学研究中发挥越来越重要的作用。

此外,改进现有技术并与其结合从而产生新的技术策略为生物学问题的研究提供新的解决方案是纳米孔测序技术的研究人员以及其他科研人员应该重点关注的领域。目前广受关注的10×Genomics公司的单细胞测序技术的基本思想是改进标签技术,给每个细胞及其遗传物质加上一个标签,从而批量对其进行处理;空间转录组测序技术的基础则是结合冷冻切片和标签技术确定细胞的位置信息,然后借助第二代测序技术对固定位置信息的细胞的转录组信息进行分析,从而为各种生物学问题的研究提供一个新的视角。作为研究人员,我们应考虑自身科研目的和需求,充分利用现有平台间的优势,取长补短,构建适宜的测序方案,从而实现新的科学发现。

## REFERENCES

- [1] Ratner L, Haseltine W, Patarca R, et al. Complete nucleotide sequence of the AIDS virus, HTLV-III. *Nature*, 1985, 313(6000): 277–284.
- [2] Barrell BG, Bankier AT, Drouin J. A different genetic code in human mitochondria. *Nature*, 1979, 282(5735): 189–194.
- [3] Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature*, 2001, 409(6822): 860–921.
- [4] Qin JJ, Li RQ, Raes J, et al. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature*, 2010, 464(7285): 59–65.
- [5] Jones MA, Bhide S, Chin E, et al. Targeted polymerase chain reaction-based enrichment and next generation sequencing for diagnostic testing of congenital disorders of glycosylation. *Genet Med*, 2011, 13(11): 921–932.
- [6] Rivas MA, Beaudoin M, Gardet A, et al. Deep resequencing of GWAS loci identifies independent

- rare variants associated with inflammatory bowel disease. *Nat Genet*, 2011, 43(11): 1066–1073.
- [7] Kim EH, Lee HY, Yang IS, et al. Massively parallel sequencing of 17 commonly used forensic autosomal STRs and amelogenin with small amplicons. *Forens Sci Int: Genet*, 2016, 22: 1–7.
- [8] Fuselli S, Baptista RP, Panziera A, et al. A new hybrid approach for MHC genotyping: high-throughput NGS and long read MinION nanopore sequencing, with application to the non-model vertebrate Alpine chamois (*Rupicapra rupicapra*). *Heredity*, 2018, 121(4): 293–303.
- [9] Mondal TK, Rawal HC, Chowrasia S, et al. Draft genome sequence of first monocot-halophytic species *Oryza coarctata* reveals stress-specific genes. *Sci Rep*, 2018, 8: 13698.
- [10] Shin H, Lee E, Shin J, et al. Elucidation of the bacterial communities associated with the harmful microalgae *Alexandrium tamarense* and *Cochlodinium polykrikoides* using nanopore sequencing. *Sci Rep*, 2018, 8: 5323.
- [11] Keller MW, Rambo-Martin BL, Wilson MM, et al. Direct RNA sequencing of the coding complete influenza A virus genome. *Sci Rep*, 2018, 8: 14408.
- [12] High-quality, Long Reads Produce Exceptional Results. 2018 Available from: <https://www.pacb.com/smrt-science/smrt-sequencing/smrt-sequencing-modes/>.
- [13] Longer and longer: DNA sequence of more than two million bases now achieved with nanopore sequencing [EB/OL]. [2019-09-24]. <https://nanoporetech.com/about-us/news/longer-and-longer-dna-sequence-more-two-million-bases-now-achieved-nanopore>.
- [14] Maxam AM, Gilbert W. A new method for sequencing DNA. *Proc Natl Acad Sci USA*, 1977, 74(2): 560–564.
- [15] Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA*, 1977, 74(12): 5463–5467.
- [16] Liu L, Li YH, Li SL, et al. Comparison of next-generation sequencing systems. *J BioMed Res Int*, 2012, 2012: 251364.
- [17] Miller W, Hayes VM, Ratan A, et al. Genetic diversity and population structure of the endangered marsupial *Sarcophilus harrisii* (Tasmanian devil). *Proc Natl Acad Sci USA*, 2011, 108(30): 12348–12353.
- [18] Margulies M, Egholm M, Altman WE, et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, 2005, 437(7057): 376–380.
- [19] Børsting C, Morling N. Next generation sequencing and its applications in forensic genetics. *Forensic Sci Int Genet*, 2015, 18: 78–89.
- [20] Chen C, Wan HS, Zhou QH. The next generation sequencing technology and its application in cancer research. *CJLC*, 2010, 13(002): 154–159 (in Chinese). 陈琛, 万海粟, 周清华. 新一代基因测序技术及其在肿瘤研究中的应用. *中国肺癌杂志*, 2010, 13(002): 154–159.
- [21] Jain M, Olsen HE, Paten B, et al. The oxford nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome Biol*, 2016, 17: 239.
- [22] de Roeck A, van den Bossche T, van der Zee J, et al. Deleterious ABCA7 mutations and transcript rescue mechanisms in early onset Alzheimer's disease. *Acta Neuropathol*, 2017, 134(3): 475–487.
- [23] Stancu MC, van Roosmalen MJ, Renkens I, et al. Mapping and phasing of structural variation in patient genomes using nanopore sequencing. *Nat Commun*, 2017, 8: 1326.
- [24] Tyson JR, O'Neil NJ, Jain M, et al. MinION-based long-read sequencing and assembly extends the *Caenorhabditis elegans* reference genome. *Genome Res*, 2018, 28(2): 266–274.
- [25] Jansen HJ, Liem M, Jong-Raadsen SA, et al. Rapid *de novo* assembly of the European eel genome from nanopore sequencing reads. *Sci Rep*, 2017, 7: 7213.
- [26] Bickhart DM, Rosen BD, Koren S, et al. Single-molecule sequencing and chromatin conformation capture enable *de novo* reference assembly of the domestic goat genome. *Nat Genet*, 2017, 49(4): 643–650.
- [27] Jain M, Koren S, Miga KH, et al. Nanopore sequencing and assembly of a human genome with ultra-long reads. *Nat Biotechnol*, 2018, 36(4): 338–345.
- [28] Arabidopsis Genome Initiative. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, 2000, 408(6814): 796–815.
- [29] Belser C, Istace B, Denis E, et al. Chromosome-scale assemblies of plant genomes

- using nanopore long reads and optical maps. *Nat Plants*, 2018, 4(11): 879–887.
- [30] Quick J, Quinlan AR, Loman NJ. A reference bacterial genome dataset generated on the MinION™ portable single-molecule nanopore sequencer. *GigaScience*, 2014, 3(1): 22.
- [31] Loman NJ, Quick J, Simpson JT. A complete bacterial genome assembled *de novo* using only nanopore sequencing data. *Nat Methods*, 2015, 12(8): 733–735.
- [32] Tanaka M, Mino S, Ogura Y, et al. Availability of Nanopore sequences in the genome taxonomy for *Vibrionaceae* systematics: rumoiensis clade species as a test case. *PeerJ*, 2018, 6: e5018.
- [33] Li CH, Chng KR, Boey EJ, et al. INC-Seq: accurate single molecule reads using nanopore sequencing. *Gigascience*, 2016, 5(1): 34.
- [34] Kranz A, Vogel A, Degner U, et al. High precision genome sequencing of engineered *Gluconobacter oxydans* 621H by combining long nanopore and short accurate Illumina reads. *J Biotechnol*, 2017, 258: 197–205.
- [35] Giordano F, Aigrain L, Quail MA, et al. *De novo* yeast genome assemblies from MinION, PacBio and MiSeq platforms. *Sci Rep*, 2017, 7: 3935.
- [36] Quick J, Loman NJ, Duraffour S, et al. Real-time, portable genome sequencing for Ebola surveillance. *Nature*, 2016, 530(7589): 228–232.
- [37] Kafetzopoulou LE, Pullan ST, Lemey P, et al. Metagenomic sequencing at the epicenter of the Nigeria 2018 Lassa fever outbreak. *Science*, 2019, 363(6422): 74–77.
- [38] Garalde DR, Snell EA, Jachimowicz D, et al. Highly parallel direct RNA sequencing on an array of nanopores. *Nat Methods*, 2018, 15(3): 201–206.
- [39] Viehweger A, Krautwurst S, Lamkiewicz K, et al. Direct RNA nanopore sequencing of full-length coronavirus genomes provides novel insights into structural variants and enables modification analysis. *Genome Res*, 2019, 29(9): 1545–1554.
- [40] Liu HL, Begik O, Lucas MC, et al. Accurate detection of m<sup>6</sup>A RNA modifications in native RNA sequences. *Nat Commun*, 2019, 10: 4079.
- [41] Stancu MC, van Roosmalen MJ, Renkens I, et al. Mapping and phasing of structural variation in patient genomes using nanopore sequencing. *Nat Commun*, 2017, 8: 1326.
- [42] Guo L, Winzer T, Yang XF, et al. The opium poppy genome and morphinan production. *Science*, 2018, 362(6412): 343–347.

(本文责编 郝丽芳)