

第三代测序技术的研究进展与临床应用

谭聃^{1,3}, 欧铜^{1,2,3}

- 1 汕头大学 医学院, 广东 汕头 515000
- 2 深圳市罗湖医院集团 医学检验中心, 广东 深圳 518000
- 3 深圳市罗湖医院集团 众循精准医学研究院, 广东 深圳 518000

谭聃, 欧铜. 第三代测序技术的研究进展与临床应用. 生物工程学报, 2022, 38(9): 3121-3130.

TAN D, OU T. Research progress and clinical application of the third-generation sequencing techniques. Chin J Biotech, 2022, 38(9): 3121-3130.

摘 要: Sanger 测序法, 又称第一代测序技术, 作为测序金标准推动了人类基因组“工作框架图”的绘制, 但通量低、成本高的缺点限制了其进一步大规模应用。第二代测序技术, 又称下一代测序技术, 因其通量高、成本低等优点使得基因测序在基础研究与临床诊疗中得到广泛应用, 但短读长一直是其不可回避的技术短板。第三代测序技术的出现, 因其具有长读长优势, 为基因序列上复杂重复区域解析与高质量基因组组装提供了新的技术手段。近年来, 第三代测序技术进一步发展完善, 同时在肿瘤、免疫、生殖等相关领域逐步体现出临床应用价值。本文将综述第三代测序技术的研究进展与临床应用。

关键词: 第三代测序; 单分子测序; 临床应用

Received: January 23, 2022; **Accepted:** April 11, 2022; **Published online:** April 15, 2022

Supported by: Key Medical Discipline Construction Project of Luohu District, Shenzhen, Guangdong, China

Corresponding author: OU Tong. E-mail: outong@zxbiomed.org

基金项目: 广东省深圳市罗湖区医学重点学科建设项目 (精准医学实验室)

Research progress and clinical application of the third-generation sequencing techniques

TAN Dan^{1,3}, OU Tong^{1,2,3}

1 Shantou University Medical College, Shantou 515000, Guangdong, China

2 Medical Laboratory of Shenzhen Luohu Hospital Group, Shenzhen 518000, Guangdong, China

3 Shenzhen Following Precision Medical Research Institute of Luohu Hospital Group, Shenzhen 518000, Guangdong, China

Abstract: The Sanger sequencing techniques, also known as the first-generation sequencing techniques and the gold standard of sequencing, have promoted the completion of “working draft” of the human genome, but the disadvantages of low throughput and high cost limit its large-scale application. The second-generation sequencing techniques, also known as the next-generation sequencing techniques, have widely used in basic research and clinical application because of its high throughput and low cost, but the short reads has always been an unavoidable shortcoming. Then, the emergence of the third-generation sequencing techniques, with the long reads, provides new technology selection for the analysis of complex repetitive regions on genome sequences and the assembly of high-quality genomes. In recent years, the third-generation sequencing techniques have been further developed, and have gradually demonstrated the clinical application value. This article reviewed the research progress and clinical application of the third-generation sequencing techniques.

Keywords: the third-generation sequencing techniques; single-molecule sequencing; clinical application

经过 40 多年的发展, 解析 DNA 遗传信息的测序技术已发展至第三代, 测序规模也从几千个碱基上升到数百万个基因组; 基因测序对科学技术发展的推动作用及影响力可媲美显微镜技术^[1], 在遗传学、免疫学、肿瘤学、微生物学等多个领域取得突出表现^[2-4]。

20 世纪 70 年代, 基于双脱氧链终止法为原理的第一代测序技术出现, 将 DNA 聚合酶合成片段、同位素标记核苷酸和凝胶电泳识别核苷酸相结合^[5], 至今仍是测序技术的金标准并运用于临床诊断^[6], 其代表性平台有 ABI Prism 3700^[7]。20 世纪 90 年代, 第二代测序技术 (又称下一代测序) 出现, 其代表性平台有 Roche 454、ABI SoliD、IonTorrent、Illumina^[8]、华大基因, 目前二代测序技术在市场上占据主导地位, 但二代测

序需要对测序片段进行扩增且测序读长短, 易引入扩增偏差且无法对复杂重复区域进行解析, 在临床应用上常需与其他方法结合使用^[9]。

2008 年, 第三代测序技术 (又称单分子测序) 出现, 具有读长长及可实现单分子与实时测序等优势^[10]。相比第二代测序的短读长、无法完整解析复杂重复序列, 第三代测序能达到 1 Mb 量级的碱基测序读长, 且当检测到复杂区域时, 如 GC 富集区、大量串联重复区等, 测序结果仍具有重要的参考价值^[11-13]。三代测序用时可压缩至数小时之内, 结合其长读长优势, 已被相关研究者优先选择用于未知基因组从头测序组装^[14]。因尚处在技术发展期, 其碱基识别准确度仍存在较高错误率 (<13%) 且成本相对较高, 限制了现阶段其在临床上的

大规模应用^[12]。代表性第三代测序技术有 HeliScope 单分子测序技术 (HeliScope single molecule sequencing)、实时单分子测序技术 (single-molecule real-time sequencing, SMRT)、纳米孔测序技术 (Oxford nanopore sequencing) 以及 Geno Care 单分子测序技术。检索 Pubmed 数据库近 21 年基于第二代、第三代测序技术的研究, 相关研究数量都呈逐年增加趋势; 虽然目前的研究与应用仍以二代测序技术为主, 但基于三代测序的研究数量已经与 5 年前基于二代测序的研究数量持平 (图 1)。

1 HeliScope 单分子测序技术

1.1 测序原理

2003 年, Helicos 单分子测序技术被报道, 相比于二代测序的短读长, 它可以将长序列片段化后再进行测序^[15-16]。该技术的测序原理与

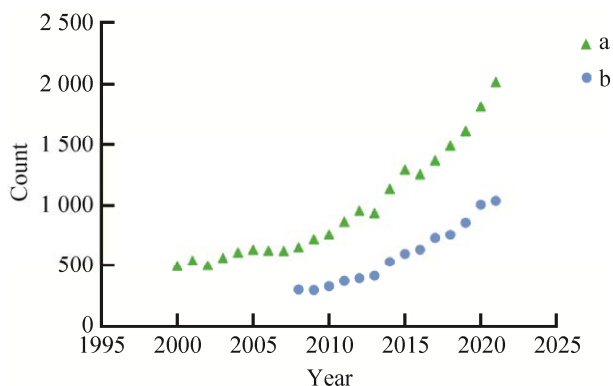


图 1 近 21 年基于第二代与三代测序技术的研究数量 数据来源于 Pubmed 数据库

Figure 1 Quantitative studies based on the second-generation and third-generation sequencing techniques over the last 21 years (data from Pubmed <http://pubmed.ncbi.nlm.nih.gov>). a: search the Pubmed website for the number of articles related to “the second-generation sequencing techniques”; b: search the Pubmed website for the number of articles related to “the third-generation sequencing techniques”.

主要流程如下 (图 2): (1) 将 DNA 大片段剪切成小片段后加上 poly(A) 尾; (2) 剪切后的片段通过 poly(A) 尾与测序芯片中测序单元 (flow cell) 上带有荧光标记的 poly(T) 寡核苷酸进行原位杂交完成装载; (3) 逐一加入荧光标记的单色核苷酸和聚合酶孵育, 使得 DNA 模板链延伸, 发生延伸反应的序列片段发出荧光; (4) 荧光被共聚焦显微镜识别并拍摄图像转化为相应的核苷酸信息; (5) 清洗荧光染料; (6) 洗涤后, 已延伸的 DNA 片段上的荧光物质被移除, 再重复以上步骤识别下一轮单个碱基直至片段测序完成; (7) 数据系统整理成序列图像^[8,17-18]。整个测序过程无需进行基因扩增即可直接读取基因序列, 避免了扩增步骤带来的测序误差^[19]。

对比后续研发的三代测序平台, HeliScope 技术与二代测序技术具有一定的相似性, 如带有荧光标记的反应体系反复添加、结合、识别、清洗的过程耗时长, 且单次测序读长偏短^[10]。

HeliScope 测序作为第三代测序技术的开端, 实现了“一次一个碱基”的单分子测序, 但其并未解决第二代测序短读长的技术短板^[20]。

1.2 临床应用

HeliScope 平台在病毒基因组测序、人类基因组拷贝数变异检测、人类转录组谱整合等方面显示了良好的应用前景^[21-23]。如 Yu 等将 HeliScope 测序与二代测序结合, 分析 22 个临床组织样本的总体基因表达情况, 结果显示两种测序方法数据集高度相关, 且当两种测序方法检测单个组织时, 该组织 75% 基因高度一致, 17% 基因数据集显示高表达基因异质性。该研究揭示 HeliScope 测序在改进基因表达模型注释与补充组织转录组生物信息方面可与二代测序互补, 对更全面认识人体不同组织转录组谱的正常状态及识别临床疾病相关组织标志物具有重要应用价值^[23]。

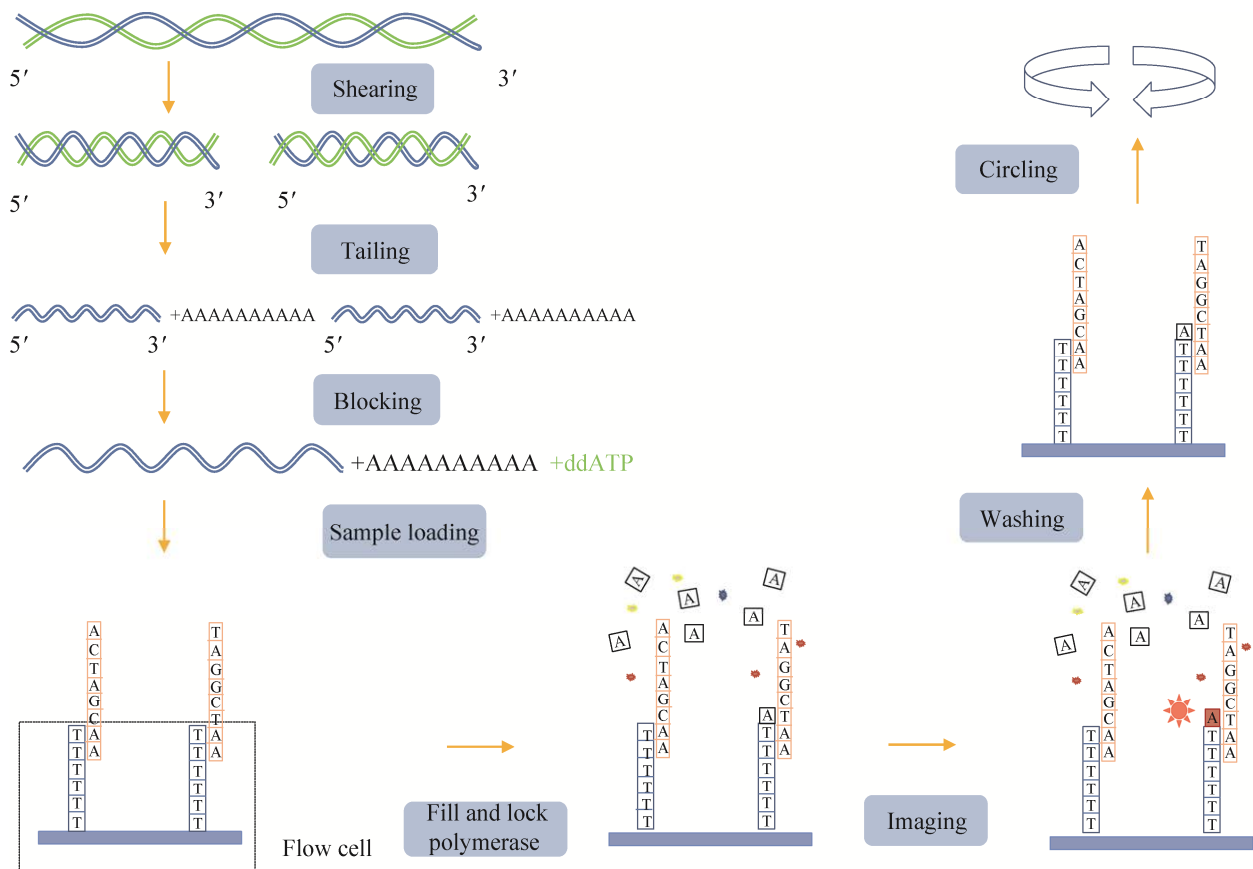


图2 HeliScope 测序平台工作流程示意图

Figure 2 Overview of the HeliScope sequencing process.

2 实时单分子测序技术

2.1 测序原理

实时单分子测序技术 (SMRT) 为目前三代测序技术中使用率较高的测序技术, 其代表性平台有 Pacific Biosciences (PacBio) 科技公司研发的各类产品^[24]。该平台测序原理与主要的测序流程如下 (图 3): (1) 将 DNA 模板与发夹接头序列连接形成闭合环状单链 DNA 模板作为测序单位用以建立 SMRT 模板库; (2) 组装好的测序单位会加载到具有零模波导 (zero-mode waveguide, ZMW) 的 SMRT 芯片单元上并与 ZMW 底部的 DNA 聚合酶结合进行测序; (3) 测序反应时, 聚合酶围绕 SMRT 模板工作, 利用带有荧光的 dNTP 延长 DNA 片段; (4) 在

延长过程中, 与模板碱基配对后的 dNTPs 荧光基团会被 ZMW 底部发射的激光激发捕获, 同时配对碱基与游离碱基停留时间有所差异, 以此来确定碱基信息^[25-26]。PacBio 测序产品主要有 RS II b、Sequel、Sequel II, 其中 Sequel II 平台提供高精度 (circular consensus sequencing, CCS) 和长读长 (continuous long reads, CLR) 读两种测序模式以供选择, 不同产品及其不同模式的读长、精确率、数据读取量等参数存在一定的差异, 可适用不同应用场景^[12]。SMRT 的测序平台能对 GC 重复区、碱基修饰区识别保持高精确率, 还能完整检测基因亚型、新基因等, 但其精准率受聚合酶有效工作时间影响较大, 而通过增加测序次数能有效降低测序过程中的随机误差^[27-29]。

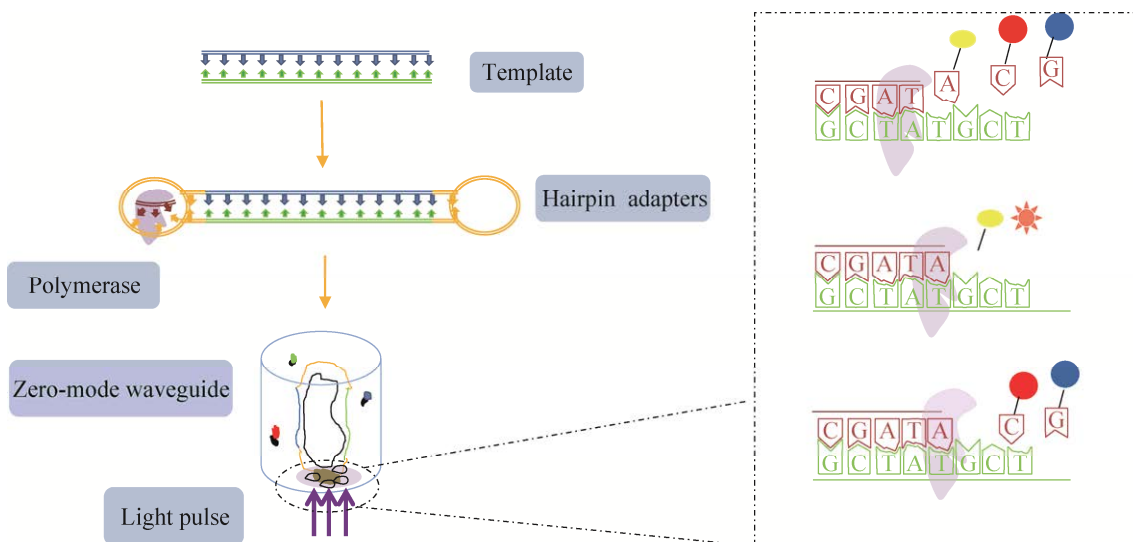


图3 实时单分子测序平台工作流程示意图

Figure 3 Overview of the single-molecule real-time sequencing process.

2.2 临床应用

由于具备长读长的优势，实时单分子测序在肿瘤学、免疫学、遗传学、生殖医学等领域展开了新的应用前景，也使得临床中对患者进行基因测序、发现等位致病基因并开发新的治疗策略等有了更多的技术方法^[30]。

在肿瘤学方面，Vasan等^[31]通过PacBio平台对多名乳腺癌患者进行基因测序发现多发*pik3ca*基因突变，其中95%的突变是同一等位基因上顺式排列的双突变；多发双突变导致了乳腺癌相关磷酸肌醇-3激酶活性增加，加速了肿瘤细胞的增殖和肿瘤的生长，且与单突变相比，双突变对激酶抑制剂的敏感性增加，有助于完善临床肿瘤治疗用药方案；而仅有SMRT测序技术的长读长模式才能对所有位于不同蛋白质结构域的*pik3ca*的双突变进行分析，对明确肿瘤患者的分子病因并精准治疗具有重要意义。

在免疫学方面，由于自然杀伤细胞的免疫球蛋白受体(killer immunoglobulin-like receptors, KIR)区域内的重组、变异和重复元件使得通过二代测序技术无法完整解析单倍型DNA，Roe

等^[32]提出了一种将PacBio长读测序与探针相结合有效捕获并组装二倍体人类KIR单倍型的方法；该方法仅使用18个探针就覆盖了KIR标准参考序列的97%，精准率达99.97%，首次对人类二倍体KIR单倍型进行完全测序与分析，可快速应用于临床诊疗。

在遗传学方面，Mizuguchi等^[33]通过PacBio长读长测序模式对患有智力障碍综合征的双胞胎进行了外显子基因组测序，精确识别了来自母系的12 kb染色体倒转；倒转直接破坏了*cpne9*和*brpf1*两个基因的正常表达，而*brpf1*异常是常染色体显性畸形面相及遗传性智力发育障碍的重要致病机理，与患者临床表征一致；该病例充分说明长读长测序对探索罕见遗传疾病中的亚显微性结构突变的优越性。

在生殖医学方面，SMRT长读长模式有助于直接检测致病基因的亲本来源，为临床预测其在后代的发生频率提供依据。Wilbe等^[34]使用SMRT测序技术对一对夫妇多次因超声诊断结构异常而反复终止妊娠的流产组织进行基因分析，鉴定出了与小儿先天性侏儒痴呆综合征

相关的 *ptpn11* 基因突变, 并确定了 *ptpn11* 突变等位基因的亲本来源及发生频率, 对该夫妇新一轮胚胎植入具有指导意义。Xu 等^[35]通过 SMRT 测序技术对地中海贫血相关基因 *hba1/2* 和 *hbb* 进行全长测序, 揭示了第二代测序技术难以解析的基因完整变异信息, 有望弥补传统测序技术漏检与分型困难的缺憾。

3 纳米孔测序技术

3.1 测序原理

1995 年, Kasianowicz 等^[36]提出了纳米孔可作为传感器用于生物学研究, 2014 年, 作为重点技术攻关对象, 纳米孔测序商业化便携式设备研发成功^[37]。纳米孔测序技术无需聚合酶参与的合成环节, 可直接对解链后的 DNA 单链进行测序。基于纳米孔测序在技术领域的重大突破, 部分学者也将其称之为第四代测序技术^[38]。纳米孔测序技术的主要原理与测序流程如下(图 4): (1) 将 DNA 分子连接到带有运动蛋白的测序接头上; (2) 带有接头的 DNA 及混合物加载至带有纳米孔的流动池中; (3) 测序开始时, 动力蛋白解开 DNA 双链, 并与电场力一起驱动带有负电荷的单链 DNA 以受控速率穿过纳米孔; (4) 不同碱基经过纳米孔产生不同电

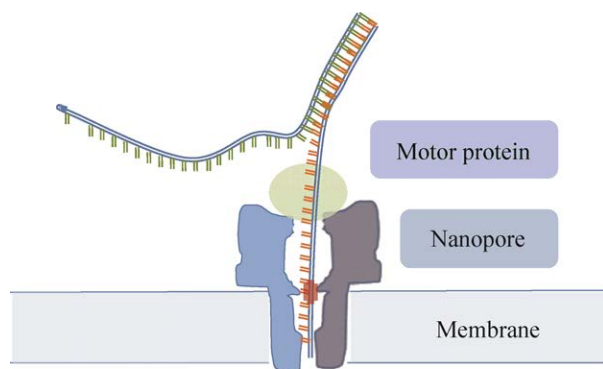


图 4 纳米孔测序平台工作流程示意图

Figure 4 Overview of the nanopore sequencing process.

信号被捕捉从而确定碱基信息^[39]。纳米孔测序无需成像系统与相关设备, 极具便携性, 因此可以把其测序设备带至任何需要测序的地点^[40]。测序结果可通过 USB 接口输出及手机小程序分析使得实时测序可为临床服务。同时, 纳米孔测序前无需 DNA 合成及扩增, 这样既对读长无明显限制 (>4 kb) 又简化了测序步骤, 还能对特殊结构(如碱基甲基化修饰等)进行直接检测^[41]。但在测序过程中, 碱基前进速度等带有随机性, 间隔太短时碱基识别易错失, 导致存在较高错误率, 限制了纳米孔测序结果价值。近年来, 不少学者在优化测序体系与开发分析算法等方面着手, 对纳米孔测序结果进行校正, 如利用最新 Bonito CRF 分析算法能显著提高单序列读长准确度^[42-44]。国内自主研发纳米孔测序仪 QNOME 系列同样在读长、通量、准确读等方面具备一定优势, 助力国内单分子测序临床端与科研端的应用 (www.qitantech.com)。

3.2 临床应用

由于便捷性与高效性并存, 纳米孔测序应用广泛^[45]。在传染病学方面, Zhu 等^[46]收集新型冠状病毒肺炎患者支气管肺泡灌洗液样本, 通过纳米孔测序技术对其进行了病原微生物测序, 分析了新型冠状病毒在内的 23 种病原体基因序列。埃博拉疫情防控中, Quick 等^[47]通过纳米孔技术对 142 个埃博拉病毒样本进行基因组测序, 仅在 1 h 内通过 20-25×测序深度就明确了样本的病毒基因型, 为在复杂的疫情防控环境中实时监测病毒基因组变异提供了参考方法。

在肿瘤学方面, 基于 DNA 甲基化是神经系统肿瘤分类的重要依据, Euskirchen 等^[48]探索纳米孔全基因组测序在神经系统肿瘤的分子诊断潜力, 高效建立了全基因组甲基胞嘧啶图谱, 通过该图谱的参考数据精准识别原发脑肿瘤, 且在临床背景下依然具备分子诊断的稳健性与

特异性。在慢性粒细胞白血病患者及部分急性淋巴细胞性白血病患者对一线药物酪氨酸激酶抑制剂产生耐药的情况下, Minervini 等^[49]通过纳米孔测序技术对耐药患者的肿瘤基因组进行测序, 揭示其中约有 25%–50% 患者 *bcr-abl1* 激酶结构域突变, 且在疾病的加速进展期及复发期突变频率增加, 可见纳米孔测序有望用于评估疾病进展并为耐药肿瘤治疗提供新思路。

在遗传学方面, 微单倍型作为一种新型遗传标记, 极具应用潜力。Wang 等^[50]利用国内自主研发的纳米孔测序平台 QNome-9604 进行了不同批次不同个体的微单倍型测序分析, 总体分型准确率达到 99.83%, 并与 MinION 测序平台平行对比, 结果充分显示了 QNome-9604 在微型单倍体分型研究领域的优势与潜力, 且开发的生信分析软件提高了自动识别分型的效率, 为未来大规模应用于遗传学、临床医学领域奠定基础。

4 Geno Care 单分子测序技术

4.1 测序原理

Geno Care 单分子测序技术是以全内反射光学为原理, 与 HeliScope 的测序原理仅有部分不同: (1) 序列 3' 末端添加的是 polyA/T 双链接头, 提高效率; (2) 同时每个测序循环添加两种 dNTP (www.genemind.com)。其代表性单分子基因测序仪 Geno Care 1600 无需扩增, 且样本制备只需 3 步就可直接读取测序芯片上的 DNA 分子, 简化了建库流程, 降低了前期样本投入量的要求, 保证了样本的高利用率。该测序技术自带内嵌式生物信息分析系统可自动化生成测序结果, 如: 无创产前、拷贝数变异、胚胎植入前以及其他定制化生物信息分析系统, 检测周期缩至 24 h 内, 为临床使用提供了高效便捷的解决方案^[51]。

4.2 临床应用

Li 等利用 Geno Care 单分子测序技术对患有 Phelan-McDermid 综合症的异卵双胞胎及其父母进行基因检测, 结果显示两名患儿有一段完全相同的 6 Mb 杂合缺失, 与父母基因检测结果相比这是一个新发突变, 且在这 6 Mb 的缺失区域中, 包含了 45 个蛋白编码基因但却不包括 *shank3* 基因, 打破了传统观点认为 *shank3* 是该病最密切相关的致病基因, 两名患儿的染色体异常均源于胚胎, 通过传统的方法无法很好地进行产前诊断, 而更高精度的检测技术能在产前实现胚胎的染色体疾病检测, 有效防控出生缺陷^[52]。

5 总结与展望

随着分子生物学的发展与“精准医疗”概念的提出, 第三代测序技术已历经多代次更迭, 性能也趋于优化与稳定。面对二代测序短读长、操作复杂、耗时长等特点, 三代测序利用现代光学、纳米技术等手段捕获序列碱基信息力求弥补缺陷, 但提升准确度、降低测序成本还是主要挑战, 近年来, 已有公布的应用软件为三代测序的比对、纠错、数据模拟等提供方案。第三代测序技术各平台部分技术参数总结见表 1。相较于早期研发的 Helicos 单分子测序, 实时单分子测序、纳米孔测序、Geno Care 单分子测序在有效测序读长与单位时间内测序通量等方面进一步提高。

以三代测序技术平台为基础的临床应用是分子诊断领域的重要方向, 也逐步展示其优势与应用前景, 便携的检测设备与缩短的测序时长使其有望用于床旁或特定环境快速检测; 可直接覆盖复杂区域的长读长与简化的实验操作有望突破现有的测序应用瓶颈, 直接转化成对肿瘤、罕见遗传疾病等的诊疗潜能, 提高精准诊疗水平, 推动人类个体化医疗发展^[53]。

表 1 第三代平台间的部分参数比较

Table 1 Partial parameter comparison between the third-generation sequencing techniques

Sequencing technology	Platform	Read length	Time spent (h)	Throughput per flow cell	Read accuracy	References
Helicos	HeliScope	25–55 bp	>48	28.0 Gb	>99%	[8]
Pacific Biosciences	RSII	>60 bp	>6	2.0 Gb	87%–92%	[12]
	Sequel	>100 bp	>20	20.0 Gb	87%–92%	
	Sequel II	>200 bp	>30	160.0 Gb	87%–92%	
Oxford Nanopore	Flongle	>2 Mb	>16	2.0 Gb	87.0%–98.3%	[12]
	MinION	>1 500 kb	>48	30.0 Gb	87.0%–98.3%	(https://nanoporetech.com)
	GridION	>1 500 kb	>48	150.0 Gb	87.0%–98.3%	
	PromethION	>1 000 kb	>72	8.6 Tb	98.3%	
QNome	QNome9604	150 kb	> 8	1.5 Gb	99.8%	(https://qitantech.com)
Geno Care	Geno Care1600	<50 kb	<24	16.0 Gb	>99.0%	(http://www.genemind.com)

在未来的一段时间里,三代测序将与一代、二代测序共存互补,而单分子测序商业产品的不断创新研发,也预示未来低成本、高效率的三代测序工具将应用更加广泛。

REFERENCES

- [1] Shendure J, Balasubramanian S, Church GM, et al. DNA sequencing at 40: past, present and future. *Nature*, 2017, 550(7676): 345-353.
- [2] Kozińska A, Seweryn P, Sitkiewicz I. A crash course in sequencing for a microbiologist. *J Appl Genet*, 2019, 60(1): 103-111.
- [3] Nakagawa H, Fujita M. Whole genome sequencing analysis for cancer genomics and precision medicine. *Cancer Sci*, 2018, 109(3): 513-522.
- [4] Xuan JK, Yu Y, Qing T, et al. Next-generation sequencing in the clinic: promises and challenges. *Cancer Lett*, 2013, 340(2): 284-295.
- [5] Sanger F, Coulson AR. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J Mol Biol*, 1975, 94(3): 441-448.
- [6] Dubucs C, Chassaing N, Sergi C, et al. re-focusing on agnathia-otocephaly complex. *Clin Oral Investig*, 2021, 25(3): 1353-1362.
- [7] Ansorge WJ. Next-generation DNA sequencing techniques. *N Biotechnol*, 2009, 25(4): 195-203.
- [8] Pareek CS, Smoczynski R, Tretyn A. Sequencing technologies and genome sequencing. *J Appl Genet*, 2011, 52(4): 413-435.
- [9] Yohe S, Thyagarajan B. Review of clinical next-generation sequencing. *Arch Pathol Lab Med*, 2017, 141(11): 1544-1557.
- [10] Schadt EE, Turner S, Kasarskis A. A window into third-generation sequencing. *Hum Mol Genet*, 2010, 19(R2): R227-R240.
- [11] Hodgkinson A, Chen Y, Eyre-Walker A. The large-scale distribution of somatic mutations in cancer genomes. *Hum Mutat*, 2012, 33(1): 136-143.
- [12] Logsdon GA, Vollger MR, Eichler EE. Long-read human genome sequencing and its applications. *Nat Rev Genet*, 2020, 21(10): 597-614.
- [13] Van Dijk EL, Jaszczyszyn Y, Naquin D, et al. The third revolution in sequencing technology. *Trends Genet*, 2018, 34(9): 666-681.
- [14] McCarthy A. Third generation DNA sequencing: pacific biosciences' single molecule real time technology. *Chem Biol*, 2010, 17(7): 675-676.
- [15] Milos P. Helicos BioSciences. *Pharmacogenomics*, 2008, 9(4): 477-480.
- [16] Thompson JF, Steinmann KE. Single molecule sequencing with a HeliScope genetic analysis system. *Curr Protoc Mol Biol*, 2010, Chapter 7: Unit7.10.
- [17] 俞晓玲, 姜文倩, 郑玲, 等. 单分子测序技术及应用研究进展. *生物化学与生物物理进展*, 2020, 47(1): 5-16.
Yu XL, Jiang WQ, Zheng L, et al. The application and research progress of single molecule sequencing technology. *Prog Biochem Biophys*, 2020, 47(1): 5-16 (in Chinese).
- [18] 刘岩, 吴秉铨. 第三代测序技术: 单分子即时测序.

- 中华病理学杂志, 2011, 40(10): 718-720.
- Liu Y, Wu BQ. Third generation sequencing technology: single-molecule real-time sequencing. *Chin J Pathol*, 2011, 40(10): 718-720 (in Chinese).
- [19] Hart C, Lipson D, Ozsolak F, et al. Single-molecule sequencing: sequence methods to enable accurate quantitation. *Methods Enzymol*, 2010, 472: 407-430.
- [20] Ozsolak F, Milos PM. Transcriptome profiling using single-molecule direct RNA sequencing. *Methods Mol Biol*, 2011, 733: 51-61.
- [21] Harris TD, Buzby PR, Babcock H, et al. Single-molecule DNA sequencing of a viral genome. *Science*, 2008, 320(5872): 106-109.
- [22] Pushkarev D, Neff NF, Quake SR. Single-molecule sequencing of an individual human genome. *Nat Biotechnol*, 2009, 27(9): 847-850.
- [23] Yu NYL, Hallström BM, Fagerberg L, et al. Complementing tissue characterization by integrating transcriptome profiling from the Human Protein Atlas and from the FANTOM5 consortium. *Nucleic Acids Res*, 2015, 43(14): 6787-6798.
- [24] Heather JM, Chain B. The sequence of sequencers: the history of sequencing DNA. *Genomics*, 2016, 107(1): 1-8.
- [25] Travers KJ, Chin CS, Rank DR, et al. A flexible and efficient template format for circular consensus sequencing and SNP detection. *Nucleic Acids Res*, 2010, 38(15): e159.
- [26] 杨春, 吴文雪, 于小淇, 等. 单分子测序技术法医学应用研究进展. *中国法医学杂志*, 2021, 36(3): 302-305, 309.
- Yang C, Wu WX, Yu XQ, et al. Research progress of single molecule sequencing in forensic medicine. *Chin J Forensic Med*, 2021, 36(3): 302-305, 309 (in Chinese).
- [27] Loomis EW, Eid JS, Peluso P, et al. Sequencing the unsequenceable: expanded CGG-repeat alleles of the fragile X gene. *Genome Res*, 2013, 23(1): 121-128.
- [28] Rhoads A, Au KF. PacBio sequencing and its applications. *Genomics Proteomics Bioinformatics*, 2015, 13(5): 278-289.
- [29] Shin SC, Ahn DH, Kim SJ, et al. Advantages of single-molecule real-time sequencing in high-GC content genomes. *PLoS One*, 2013, 8(7): e68824.
- [30] Eichler EE. Genetic variation, comparative genomics, and the diagnosis of disease. *N Engl J Med*, 2019, 381(1): 64-74.
- [31] Vasan N, Razavi P, Johnson JL, et al. Double PIK3CA mutations in *Cis* increase oncogenicity and sensitivity to PI3K α inhibitors. *Science*, 2019, 366(6466): 714-723.
- [32] Roe D, Williams J, Ivery K, et al. Efficient sequencing, assembly, and annotation of human KIR haplotypes. *Front Immunol*, 2020, 11: 582927.
- [33] Mizuguchi T, Okamoto N, Yanagihara K, et al. Pathogenic 12-kb copy-neutral inversion in syndromic intellectual disability identified by high-fidelity long-read sequencing. *Genomics*, 2021, 113(1 Pt 2): 1044-1053.
- [34] Wilbe M, Gudmundsson S, Johansson J, et al. A novel approach using long-read sequencing and ddPCR to investigate gonadal mosaicism and estimate recurrence risk in two families with developmental disorders. *Prenat Diagn*, 2017, 37(11): 1146-1154.
- [35] Xu LP, Mao AP, Liu H, et al. Long-molecule sequencing: a new approach for identification of clinically significant DNA variants in α -thalassemia and β -thalassemia carriers. *J Mol Diagn*, 2020, 22(8): 1087-1095.
- [36] Kasianowicz JJ, Brandin E, Branton D, et al. Characterization of individual polynucleotide molecules using a membrane channel. *PNAS*, 1996, 93(24): 13770-13773.
- [37] Deamer D, Akeson M, Branton D. Three decades of nanopore sequencing. *Nat Biotechnol*, 2016, 34(5): 518-524.
- [38] Feng YX, Zhang YC, Ying CF, et al. Nanopore-based fourth-generation DNA sequencing technology. *Genomics Proteomics Bioinformatics*, 2015, 13(1): 4-16.
- [39] Lin B, Hui JN, Mao HJ. Nanopore technology and its applications in gene sequencing. *Biosensors*, 2021, 11(7): 214.
- [40] Petersen LM, Martin IW, Moschetti WE, et al. Third-generation sequencing in the clinical laboratory: exploring the advantages and challenges of nanopore sequencing. *J Clin Microbiol*, 2019, 58(1): e01315-e01319.
- [41] Kono N, Arakawa K. Nanopore sequencing: review of potential applications in functional genomics. *Dev Growth Differ*, 2019, 61(5): 316-326.
- [42] Karst SM, Ziels RM, Kirkegaard RH, et al. High-accuracy long-read amplicon sequences using unique molecular identifiers with Nanopore or PacBio sequencing. *Nat Methods*, 2021, 18(2): 165-169.
- [43] Rauluseviciute I, Drabløs F, Rye MB. DNA methylation data by sequencing: experimental

- approaches and recommendations for tools and pipelines for data analysis. *Clin Epigenetics*, 2019, 11(1): 193.
- [44] Sahlin K, Medvedev P. Error correction enables use of Oxford Nanopore technology for reference-free transcriptome analysis. *Nat Commun*, 2021, 12(1): 2.
- [45] Loit K, Adamson K, Bahram M, et al. Relative performance of MinION (Oxford nanopore technologies) versus sequel (Pacific biosciences) third-generation sequencing instruments in identification of agricultural and forest fungal pathogens. *Appl Environ Microbiol*, 2019, 85(21): e01368-e01319.
- [46] Zhu N, Zhang DY, Wang WL, et al. A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med*, 2020, 382(8): 727-733.
- [47] Quick J, Loman NJ, Duraffour S, et al. Real-time, portable genome sequencing for Ebola surveillance. *Nature*, 2016, 530(7589): 228-232.
- [48] Euskirchen P, Bielle F, Labreche K, et al. Same-day genomic and epigenomic diagnosis of brain tumors using real-time nanopore sequencing. *Acta Neuropathol*, 2017, 134(5): 691-703.
- [49] Minervini CF, Cumbo C, Orsini P, et al. Nanopore sequencing in blood diseases: a wide range of opportunities. *Front Genet*, 2020, 11: 76.
- [50] Wang Z, Qin L, Liu J, et al. Forensic nanopore sequencing of microhaplotype markers using QitanTech's QNome. *Forensic Sci Int Genet*, 2022, 57: 102657.
- [51] Hu L, Lin QT, Xie PY, et al. Accurate CNV identification from only a few cells with low GC bias in a single-molecule sequencing platform. *bioRxiv*, 2020, DOI:10.1101/2020.01.21.908897.
- [52] Li S, Xi KW, Liu T, et al. Fraternal twins with Phelan-McDermid syndrome not involving the *SHANK3* gene: case report and literature review. *BMC Med Genomics*, 2020, 13(1): 146.
- [53] 盛杰, 王玉欣, 齐江发, 等. 单分子测序技术在肿瘤诊断中的应用研究进展. *生物工程学报*, 2020, 36(2): 180-188.
- Sheng J, Wang YX, Qi JF, et al. Progress in the application of single molecular sequencing in tumor diagnosis. *Chin J Biotech*, 2020, 36(2): 180-188 (in Chinese).

(本文责编 陈宏宇)