

• 综 述 •

## 单分子测序技术在肿瘤诊断中的应用研究进展

盛杰, 王玉欣, 齐江发, 杜庆, 徐瑶

武汉科技大学 生命科学与健康学院 生物医学研究院, 湖北 武汉 430081

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.

**摘 要:** 肿瘤的发生发展通常与多个基因的遗传突变、表达异常有关, 全面深入地分析肿瘤基因组、转录组及表观遗传组学对于快速剖析疾病特定基因簇及修饰位点至关重要。之前, 研究者们主要采用第二代测序技术进行有效信息的挖掘, 然而随着研究的不断深入, 第二代测序技术表现出诸如序列拼接困难、低丰度难以检出等缺点。因此, 单分子测序技术以其独特的优势应运而生, 文中主要对单分子测序技术在几种常见肿瘤中的研究现状进行了综述, 并展望了其在临床诊断中的应用前景。

**关键词:** 单分子测序技术, 单分子实时测序, 肿瘤, 临床诊断

## Progress in the application of single molecular sequencing in tumor diagnosis

Jie Sheng, Yuxin Wang, Jiangfa Qi, Qing Du, and Yao Xu

*Institute of Biology and Medicine, College of Life Science and Health, Wuhan University of Science and Technology, Wuhan 430081, Hubei, China*

**Abstract:** Tumor development is usually related to the genetic mutation and abnormal expression of multiple genes. Comprehensive analysis of tumor genome, transcriptome and epigenetics is very important for the rapid identification of disease-specific gene clusters and modification sites. Previously, the next-generation sequencing technology was mainly used to explore the information of genomes, however, it cannot meet the requirement of mechanical researches due to several problems such as difficult sequence assembly and leak detection of the low abundance factors. Therefore, single molecule sequencing technology gradually emerged with its unique superiority. This paper reviews the research processes of single molecule sequencing technology in several human tumors, and prospects its application in clinical diagnosis.

**Keywords:** single molecular sequencing technology, single molecule real-time, tumor, clinical diagnosis

**Received:** May 21, 2019; **Accepted:** September 5, 2019

**Supported by:** National Natural Science Foundation of China (Nos. 31600617, 81802008), Scientific Research Program of Hubei Provincial Education Department (No. Q20181102).

**Corresponding authors:** Yao Xu. Tel/Fax: +86-27-68893368; E-mail: xuyao0307@wust.edu.cn

国家自然科学基金 (Nos. 31600617, 81802008), 湖北省教育厅科学研究计划 (No. Q20181102) 资助。

网络出版时间: 2019-09-19

网络出版地址: <http://kns.cnki.net/kcms/detail/11.1998.Q.20190918.1016.002.html>

DNA 测序技术能够快速、准确地检测基因遗传突变及疾病相关基因的异常表达,因而在肿瘤临床诊断中具有潜在的应用价值。基因测序技术经历了几个重要的发展阶段,第一代测序技术主要包括 Sanger 测序法(双脱氧链终止法)<sup>[1]</sup>和化学降解法<sup>[2]</sup>,但其测序费用昂贵、耗时长,因此难以普及。随着技术的不断发展与改进,第二代测序技术应运而生,主要包括 Roche 公司的 454 技术、Illumina 公司的 Solexa 技术和 ABI 公司的 SOLID 技术,第二代测序技术显著提高了测序通量,并且能够通过结合 PCR 扩增和荧光标记成像技术获取全部的遗传信息<sup>[3]</sup>。与第一代相比,第二代测序技术大幅度地降低测序成本、节约测序时间,但其读长较短,给基因组重复区域的序列组装、序列匹配等下游分析造成了诸多困难。前期在肿瘤中的测序结果显示,短的读取长度会引起基因组高复杂度的区域组装错误,此外,还存在测序间隙(Gap)区增多、低丰度难以检出等缺点<sup>[4]</sup>。针对以上问题,第三代测序技术,即单分子测序逐渐发展起来,弥补了第二代测序读长较短的劣势,可以对基因组的某些区域直接测序。在肿瘤检测方面,单分子测序技术具有其独特的优势,不仅可以检测基因遗传突变,还可以通过检测表观遗传修饰位点<sup>[5]</sup>、长链非编码 RNA(Long non-coding RNA, lncRNA)丰度<sup>[6]</sup>等对肿瘤进行分级和分型。本文将重点阐述单分子测序技术在乳腺癌、非小细胞肺癌、卵巢癌、前列腺癌中的研究现状,探讨其临床应用前景。

## 1 单分子测序技术简介

### 1.1 单分子实时测序技术

单分子实时测序技术(Single molecule real-time, SMRT)是指在使用 4 种不同荧光标记的脱氧核糖核苷酸(dNTPs)进行连续的模板合成过程中,从 DNA 聚合酶中获取单分子实时测序数据的一项技术<sup>[7]</sup>。实验原理是将待测 DNA 模板结合的 DNA 聚合酶分子固定在一个纳米级检测室(Zero-

mode waveguides, ZMWs)的底部,然后加入 4 种不同的荧光标记的 dNTPs,使用激光照射 ZMW 的底部。由于 ZMW 是一种简单的纳米孔结构,且该孔比单个激光波长还要短以至于激光在孔的附近发生光的衍射,只能激发 ZMW 检测室<sup>[3]</sup>。因此,在 ZMW 区域内的碱基携带的荧光基团能够被直接检测到(图 1)。SMRT 技术可以在高浓度下分离单个分子进行光学分析,从而确定单分子水平上的荧光变化<sup>[8]</sup>。

SMRT 技术采用的荧光标记方法也是其一大特色。虽然之前有其他的测序方法使用了荧光标记,但是其荧光基团被标记在甲基上,在 DNA 链延伸的过程中不会被去掉,造成空间位阻,引起序列错读<sup>[9]</sup>。SMRT 技术是在核苷酸的磷酸基团上进行荧光标记,随着 DNA 链的延伸,磷酸键断开,荧光标记被去掉,不会造成空间位阻<sup>[10]</sup>,因此,SMRT 技术可以读取更长的 DNA 序列。此外,由于检测室底部的 DNA 聚合酶活性会受到激光照射的影响,限制了 DNA 序列的连续合成<sup>[3]</sup>,但该技术已经是 DNA 测序史上的一个里程碑。

### 1.2 单分子测序技术与第二代测序技术的比较

与第二代测序相比,单分子测序技术具有其独特的优势(表 1)。本课题组早期利用第二代测序平台检测的肿瘤转录组数据显示,有许多关键基因序列被漏检,尤其是具有特殊结构的区域。主要是由于第二代测序技术需进行 PCR 扩增,而高 GC 或 AT 含量都不易进行 PCR 扩增,进而影响基因组测序的完整性。此外,高 GC 含量的序列难以被打断成小片段,因而无法读取。高 AT 含量的序列只能通过降低引物延伸温度等方法得到轻微的改善,最终导致 PCR 扩增片段库中缺乏富含 AT 的片段<sup>[11]</sup>。总而言之,第二代测序技术在建库过程中需要扩增和分段环节,二者都可能在读取和序列组装中引入大量偏差和错误<sup>[12-14]</sup>。然而,单分子测序技术无需进行 PCR 扩增,可以完全避免产生上述影响,还能最大限度地节约成本。

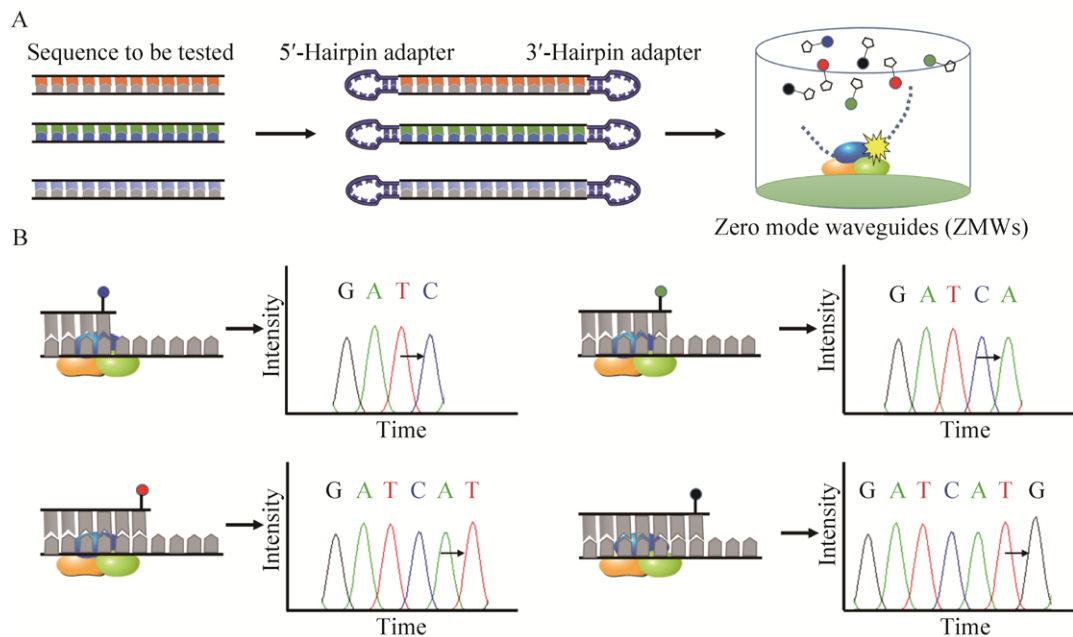


图 1 单分子实时测序的原理示意图

Fig. 1 Schematic diagram of single molecule real-time sequencing.

单分子测序技术的建库和测序步骤比第二代测序更为简便，能够快速地对全基因组序列进行测序拼接，极大地缩短测序时间。针对三代测序产生较高随机测序错误的特点，SMRT 技术可通过置换聚合酶对模板进行多次测序，使模板中每个碱基被多次读取，从而提高测序的准确性<sup>[15]</sup>。

与第二代测序技术不同的是，单分子测序可以更加精确地检测基因组单核苷酸多态性 (Single nucleotide polymorphisms, SNPs)、基因重排和大的结构变异，如拷贝数变异，发现更多新的突变，并对变异位点进行准确定位<sup>[16-17]</sup>。此外，单分子测序还可以直接检测表观遗传信息，当碱基存在修饰时，DNA 聚合酶的合成速度会减慢，不同的修饰类型会引起聚合酶合成过程的“停顿模式”产生微小差异，这些差异将反映在荧光脉冲信号的间隔上，最终通过对应的软件分析即可确定碱基修饰类型<sup>[18]</sup>。目前，单分子测序技术可以检测 5-hC、5-hmU、5-hU、1-mA、6-mA、8-oxoA、BPDE、6-mT、6-mG 等碱基修饰，甚至可以鉴别

传统亚硫酸氢盐测序法无法区分的甲基化修饰和羟甲基化修饰。总之，该技术使得从 DNA 上同时获得基因组序列信息和单核苷酸水平的表观遗传信息成为可能，为深入解析生物体不同代谢过程的表观调控机制提供重要的技术支撑<sup>[19]</sup>。目前，第二代测序技术已经发展成熟且在各个研究领域广泛应用，第三代测序技术还处于快速发展期，不过凭借着其强大的技术优势，未来将在基因测序这个大舞台上大放异彩<sup>[20-23]</sup>。

## 2 单分子测序技术在肿瘤研究中的应用现状

肿瘤的发生与功能基因突变、缺失、异位、重复、修饰等密切相关<sup>[24]</sup>。因此，研究肿瘤基因组学、转录组学、表观遗传组学特征对于探究肿瘤的发生发展、转移、复发的分子机制具有重要的指导作用。单分子测序技术的发展实现了对个体染色体的直接测序，有助于建立每个人完善的基因组信息档案，从分子水平阐明肿瘤的病理生理特征，实现个性化医疗。在肿瘤诊断方面，根

表 1 单分子测序技术与第二代测序技术的各项参数比较

Table 1 The parameters comparison of the single molecular sequencing with the next-generation sequencing technology

Comparison contents	Next generation sequencing	Single molecular sequencing	Reference
Running time	1–6 d	0.5–4 h	[20–21]
Reading length	35–300 bp	3 000 bp on average	[7, 22]
Error rate	1%–10%	3%–15%	[5, 15]
Error type	Mainly produced by GC-rich base sequences	Mainly produced by insert/delete errors, using PacBio Quiver software, the accuracy can reach 99.999%.	[4, 23]
Requirements for samples	High quality sample	No special requirement	[40]
Detection of epigenetic modification sites	No	Detection of methylation, acetylation, etc	[16]
Effects on sequencing result	Yes	No	[11]

据单分子测序的研究结果,可以鉴定出与肿瘤发生发展密切相关的标志分子,并将其应用于肿瘤的早期诊断,对于提高肿瘤患者的生存率具有重要的实际意义<sup>[25]</sup>。

## 2.1 单分子测序技术在乳腺癌中的应用

乳腺癌是一种常见的恶性肿瘤,在女性群体中的发病率远远高于男性,是导致全球范围内女性患癌死亡的主要原因<sup>[26]</sup>。早诊断、早治疗是提高乳腺癌患者生存率的关键。目前,临床上主要通过检测血液中CA153的含量来评价乳腺癌的疾病进程,被认为是乳腺癌患者诊断和监测术后复发、观察疗效较为灵敏的肿瘤标志物之一<sup>[27]</sup>。然而,CA153的单项检查存在一定的局限性,在乳腺癌的早期诊断过程中,其敏感性和特异性表现不佳。为了开发理想的临床诊断模式,提高乳腺癌早期诊断的灵敏度和特异性,科学家们逐步开展了全基因组范围内的位点筛查工作。

基因突变在人类基因组中广泛存在,不仅发生于正常个体的基因组中,与个体发育、组织器官特征密切相关;而且还在疾病个体的基因组中频繁发生,与疾病的发生发展、病理学变化有关<sup>[28–30]</sup>。Garieli等采用单分子测序技术对乳腺癌患者基因组进行测序,结果显示,在乳腺癌相关基因BRCA1上共检测到177个SNPs,其中有40个SNPs位点是之前用第二代测序技术未检测到的,

有24(60%)个是新检测到的Reads。此外,短的插入/缺失突变(1–2 bp)也较先前的二代测序结果更为丰富<sup>[31]</sup>。由此可见,BRCA1基因的遗传变异可能是乳腺癌的易感因素,将其与现有的检测手段相结合,为乳腺癌的风险分析提供遗传学基础。针对基因组中的罕见变异,单分子测序技术还设计了SMRTbell<sup>TM</sup>模块,将其放入单分子测序反应中,能通过对单个分子的多次传递进行重复读取,从而提高检出率<sup>[32]</sup>。在前期乳腺癌单分子测序结果的基础上,本课题组在乳腺癌患者和健康个体中验证了VEGF基因的变异情况,结果发现与对照组相比,大部分乳腺癌患者VEGF基因的启动子区存在18 bp序列的缺失,该突变能够直接影响VEGF的启动子活性及其表达,促进了乳腺癌细胞的迁移和血管形成,提示该突变可能作为乳腺癌的诊断标志(Guo等未发表数据)。

非编码RNA参与细胞的生长、增殖、分裂和分化过程,在重要功能基因表达调控方面发挥着不可或缺的作用,其表达变化可导致肿瘤的发展和转移<sup>[33–34]</sup>。越来越多的证据表明,lncRNA在乳腺癌的发生发展中具有重要的调控作用<sup>[35–36]</sup>。Jonsson等利用单分子测序技术对雌激素诱导的管腔A型乳腺癌细胞系(MCF7和T47D)进行转录组测序,共检测到2 000个功能基因及1 000个

lncRNA 受到 ER $\alpha$  的调控, 进一步研究证实, 长链非编码 RNA 分子 LINC01016 和 LINC00160 是 ER $\alpha$  的直接靶基因, 敲低 LINC00160 能够显著抑制乳腺癌细胞增殖<sup>[37]</sup>。由此可见, 这些长链非编码 RNA 可以作为乳腺癌的诊断标志, 在乳腺癌的临床检测中具有潜在的应用价值。研究者们之所以选择单分子测序技术, 是因为单分子测序会以一种无偏倚的方式精确检测到低丰度的转录本, 对所有 lncRNA 进行全面分析。对这些 lncRNA 基因的特性及其作用模式的解析, 有助于改进现有的癌症诊断、监测和靶向治疗方法。综上, 通过单分子测序技术已经筛选了多个与乳腺癌相关的靶标, 但要将其应用于临床诊断, 还需针对灵敏度、特异性问题等展开深入的研究和评价。

## 2.2 单分子测序技术在非小细胞肺癌中的应用

肺癌在全球范围内的发病率、死亡率逐年上升, 严重危害人类健康。有 80% 的肺癌属于非小细胞肺癌 (Non-small cell lung cancer, NSCLC), 组织类型以腺癌和鳞状细胞癌最为常见<sup>[38]</sup>。小细胞癌通常对化疗和放射治疗反应敏感, 而 NSCLC 往往需要采用手术治疗。目前, 在临床个体化诊疗过程中, 由于大部分肿瘤患者无法进行组织活检, 进行初诊的非小细胞肺癌病人往往已到了患病晚期, 错失了手术治疗的时机<sup>[39]</sup>。因此, 寻找一种全新的、便捷的、准确诊断非小细胞肺癌的方法迫在眉睫, 尤其是患病早期的诊断对于提高患者生存率至关重要。

早在 1948 年, 研究人员便在外周血中发现了游离 DNA, 且肿瘤患者外周血中 DNA 含量显著高于正常人, 随后在这些 DNA 中检测到了突变的原癌基因, 并且与原发肿瘤一致, 证实了循环肿瘤 DNA (Circulating tumor DNA, ctDNA) 的存在。ctDNA 是一种液体活检分子, 具有分离简单、易于监测等优点, 并且 ctDNA 的水平随肿瘤负荷程度的变化而变化。对 ctDNA 进行测序分析能够直接反映肿瘤的进展情况, 有望成为跟踪肿瘤进

化的有效工具<sup>[40]</sup>。由于肿瘤患者外周血中大多数的游离 DNA 来自于正常细胞, ctDNA 只占外周血游离 DNA 的一小部分。因此, 对 ctDNA 含量进行检测的技术, 必须能够对微量的 ctDNA 进行准确检测与定量。第二代测序技术在检测 ctDNA 过程中存在速度慢、检出率低等问题<sup>[41-42]</sup>, 而单分子测序不仅提高了通量, 能够对微量 ctDNA 进行检测, 还大大提高了检测速率, 降低测序成本, 因而具有良好的前景。研究者们通过对非小细胞肺癌患者外周血中的 ctDNA 进行测序分析, 发现 EGFR 基因中 62.5% 的突变在 ctDNA 中被检测到 (包含耐药突变 p.T790M), 进一步证实了 ctDNA 可以作为潜在的肿瘤标志物进行临床检测<sup>[43]</sup>。此外, 非小细胞肺癌的发生被发现与功能基因的序列突变有关, DiBardino 等采用第二代测序技术对 22 份非小细胞肺癌的标本中 467 个基因进行测序分析, 共检测到 204 个基因变异位点, 并且大部分突变都是位于 TP53、EGFR、EPHB1、MLL3、APC 等非小细胞肺癌相关基因上<sup>[44]</sup>。鉴于与 NSCLC 肿瘤相关临床致癌突变的多样性, 急需一种更全面的方法, 能够提高致癌位点的检出率, 以评估患者的肿瘤突变情况。Shi 等通过单分子测序技术评估了 154 例非小细胞肺癌标本中上皮生长因子受体 EGFR、KRAS、BRAF 和 ALK 突变的类型、频率和丰度, 结果发现有 44.2% 的突变位于 EGFR 基因上, 且突变多为杂合型, 说明 EGFR 突变是非小细胞肺癌的易感因素<sup>[17]</sup>。目前, 在非小细胞肺癌基因组测序的文献中大多采用了第二代测序技术, 相信随着单分子测序的不断发展和应用, 会有越来越多的疾病敏感位点被检测出来。

## 2.3 单分子测序技术在卵巢癌中的应用

卵巢癌是最致命的妇科恶性肿瘤之一, 其明显特征是发病隐匿, 病程进展迅速, 在患病早期即可发生转移、扩散<sup>[45]</sup>。据统计, 70%–80% 的卵巢癌患者在进行入院初诊时已处于疾病晚期, 5 年

生存率仅 20%–30%，而早期的卵巢癌患者生存率可达 90%。因此，提高卵巢癌的早期临床诊断水平，对于争取最佳的治疗时机和改善预后都具有重要意义。

研究发现，在卵巢癌患者的基因组中存在大规模的拷贝数变异，导致转录元件的异常结合和转录错误。Ayhan 等研究发现，CCNE1 基因拷贝数的增加与卵巢癌相关基因 ARID1A、PIK3CA 和 ZNF217 等的异常转录和表达有关，并且携带该结构变异的患者预后较差<sup>[46]</sup>。因此，对卵巢癌患者基因组和转录本的全面解析有助于开发新的治疗策略和靶向药物。近年来，第二代高通量测序技术的发展加速了人们对遗传基础的认识，在一定程度上解释了卵巢癌的遗传发病机理<sup>[47-48]</sup>。然而，第二代测序平台有多种限制因素，在序列装配过程中容易引入错误<sup>[12]</sup>。通常情况下，肿瘤组织中的转录本具有丰富的多样性，目前的工作流程无法准确地解析转录起始位点、剪切连接、聚腺苷酸化的位置和基因融合事件<sup>[14]</sup>。第三代测序技术，如 SMRT，可以产生较长的读取序列 (>10 kb)，有时可以跨越整个转录单元，避免了对异构体进行后续组装<sup>[49]</sup>。单分子测序技术还使用了一系列的方法来纠正错误，包括多次读取循环模板，以及结合第二代高保真度短片段的混合测序方法，最终提高测序的准确性<sup>[32]</sup>。Jing 等整合了第二代测序平台和单分子测序平台，对卵巢癌患者的完整基因组和转录组进行测序分析，获得了受转移性上皮性卵巢癌影响的患者基因组多维数据集，特别是通过单分子测序鉴定了大量未注释的新的转录本、新的长链非编码 RNA 和基因嵌合体，并对基因转录起始、剪接、多聚腺苷酸和融合位点进行准确定位，该研究发现了多个与卵巢癌相关的功能基因转录本和非编码 RNA<sup>[50]</sup>。Winterhoff 等通过测序比较了卵巢癌上皮细胞和基质细胞中的转录组数据，结果显示，调节细胞外基质和上皮间质转化的基因在基质细胞中高表达<sup>[51]</sup>。然而，这些基因能否用于卵巢癌的临床诊

断及监测，还需在体内外对其分子调控机制展开深入研究。

#### 2.4 单分子测序技术在前列腺癌中的应用

前列腺癌是一种男性常见的恶性肿瘤，在欧美发达国家发病率很高，是男性癌症并引起死亡的第二大原因，仅在美国每年就有约 3 万人因此而死亡<sup>[52]</sup>。我国前列腺癌的发病率也呈逐年增高的趋势，在男性泌尿系统肿瘤中占居第三位。研究表明，前列腺癌的肿瘤异质性特征使其在进展速度上产生显著差异，目前，临床上尚无有效的方法能够在治疗前准确区分惰性肿瘤和侵袭性肿瘤。因此，开发精准的诊断和预后工具，给患者提供个性化的治疗和监测平台，对于识别易发生前列腺癌的种系突变、评估特定的体细胞遗传事件与临床病理信息的相关性至关重要。

前列腺特异抗原 (Prostate specific antigen, PSA) 是目前被广泛应用于临床诊断前列腺癌的关键指标，但是 PSA 存在灵敏度和特异性较低的缺点，不能准确鉴别前列腺良性增生和前列腺癌，在早期诊断方面表现不佳<sup>[53]</sup>。随着分子检测技术的不断发展，越来越多的与前列腺癌发生发展相关的基因、变异位点、非编码 RNA 等特异性标志物被鉴定出来。基因表达差异数据显示，前列腺癌基因 3 (PCA3) 在癌组织中的表达水平是正常组织的 6–34 倍，进一步临床检测结果证明 PCA3 的敏感性和特异性均高于 PSA<sup>[54]</sup>。Zhang 等采用第二代测序和全基因组关联分析方法筛选了与前列腺癌相关的 SNP 位点，共鉴定出位于 11 个数量性状基因的 56 个 SNPs 位点，其基因型频率在前列腺癌患者与正常人之间存在显著差异，其中，RGS17 和 ASCL2 两个位点可能是前列腺癌的主要风险因素<sup>[55]</sup>。针对前列腺癌的空间异质性、瘤内异质性和瘤间异质性，单分子测序技术能够检测来自前列腺不同侧肿瘤细胞的完整的基因组信息，并对差异位点进行准确定位，有助于了解前列腺癌的进展和克隆进化。Su 等对不同部位的前

列腺癌组织进行单分子测序, 结果发现 TP53 突变存在显著差异, 这对于制定前列腺癌的诊断和治疗方案具有重要的指导意义<sup>[56]</sup>。此外, Tevz 等采用单分子测序技术 PacBio SMRT, 对前列腺癌患者 RLN1 和 RLN2 两个转录本进行检测, 发现融合的 RLN1-RLN2 在前列腺癌细胞中高表达, 预示其可能参与前列腺癌的发病过程<sup>[57]</sup>。然而, 这些因子参与前列腺癌发生发展的分子调控机制还有待进一步阐明。总之, 通过对高通量测序数据的有机整合, 筛选出具有高灵敏度和高特异性的联合检测体系, 将成为未来前列腺癌临床诊断的新标准。

### 3 总结与展望

单分子测序技术不仅能够从基因遗传突变方面和表观遗传修饰方面解析人类基因组多样性, 还可以对人类重大疾病如肿瘤进行分级、分型。在遗传多态性检测方面, 单分子测序技术能够对每一条 DNA 链进行单独测序, 利用了聚合酶自身的反应速度, 大大减少了测序时间, 不需要进行聚合酶链式反应, 读长较长, 能够对复杂基因组进行测序。在表观遗传学方面, 单分子测序技术能够直接检测 DNA 的碱基修饰位点, 而这些修饰又常常与癌症相关, 弥补了之前测序技术不能涉及的领域。虽然单分子测序具有明显的技术优势, 但该技术的测序错误率较高, 极大限制了其应用和普及。因此, 研发更高效的纠错算法或提高测序准确率才能推动单分子测序技术的发展, 有助于解析个体化遗传图谱及疾病相关的敏感位点。

由于单分子测序摒弃了二代测序中大量使用的 PCR 技术, 能够实现最自然条件下 DNA 原始序列的测序, 因此, 在肿瘤诊断方面具有广阔的应用前景。目前, 已经有多个潜在的肿瘤标志物被鉴定出来。然而, 这些标志因子所参与的信号通路、分子调节机制, 及其能否组合使用进行临

床诊断等问题还有待进一步研究。相信在不久的将来, 随着技术的不断改进和完善, 单分子测序必然会进一步推动基因组学、转录组学和表观遗传组学深度覆盖鉴定和定量技术的进步, 为开发更加有效的肿瘤临床诊断策略提供坚实的技术支撑。

### REFERENCES

- [1] Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA*, 1977, 74(12): 5463–5467.
- [2] Berenbaum MR. *Proceedings of the national academy of sciences-its evolution and adaptation*. *Proc Natl Acad Sci USA*, 2019, 116(3): 704–706.
- [3] Ortiz-Cuaran S, Scheffler M, Plenker D, et al. Heterogeneous mechanisms of primary and acquired resistance to third-generation EGFR inhibitors. *Clin Cancer Res*, 2016, 22(19): 4837–4847.
- [4] Buermans HPJ, den Dunnen JT. Next generation sequencing technology: advances and applications. *Biochim Biophys Acta*, 2014, 1842(10): 1932–1941.
- [5] Ameer A, Kloosterman WP, Hestand MS. Single-molecule sequencing: towards clinical applications. *Trends Biotechnol*, 2019, 37(1): 72–85.
- [6] Jenjaroenpun P, Wongsurawat T, Pereira R, et al. Complete genomic and transcriptional landscape analysis using third-generation sequencing: a case study of *Saccharomyces cerevisiae* CEN.PK113-7D. *Nucleic Acids Res*, 2018, 46(7): e38.
- [7] Ardui S, Ameer A, Vermeesch JR, et al. Single molecule real-time (SMRT) sequencing comes of age: applications and utilities for medical diagnostics. *Nucleic Acids Res*, 2018, 46(5): 2159–2168.
- [8] Zhu P, Craighead HG. Zero-mode waveguides for single-molecule analysis. *Annu Rev Biophys*, 2012, 41: 269–293.
- [9] Braslavsky I, Hebert B, Kartalov E, et al. Sequence information can be obtained from single DNA molecules. *Proc Natl Acad Sci USA*, 2003, 100(7): 3960–3964.
- [10] Korlach J, Bjornson KP, Chaudhuri BP, et al. Real-time DNA sequencing from single polymerase molecules. *Methods Enzymol*, 2010, 472: 431–455.
- [11] Aird D, Ross MG, Chen WS, et al. Analyzing and

- minimizing PCR amplification bias in Illumina sequencing libraries. *Genome Biol*, 2011, 12(2): R18.
- [12] Alkan C, Sajjadian S, Eichler EE. Limitations of next-generation genome sequence assembly. *Nat Methods*, 2011, 8(1): 61–65.
- [13] Steijger T, Abril JF, Engström PG, et al. Assessment of transcript reconstruction methods for RNA-seq. *Nat Methods*, 2013, 10(12): 1177–1184.
- [14] Niu BF, Fu LM, Sun SL, et al. Artificial and natural duplicates in pyrosequencing reads of metagenomic data. *BMC Bioinformatics*, 2010, 11: 187.
- [15] Lou DI, Hussmann JA, McBee RM, et al. High-throughput DNA sequencing errors are reduced by orders of magnitude using circle sequencing. *Proc Natl Acad Sci USA*, 2013, 110(49): 19872–19877.
- [16] Gao Y, Deng LW, Yan Q, et al. Single molecule targeted sequencing for cancer gene mutation detection. *Sci Rep*, 2016, 6: 26110.
- [17] Shi J, Yuan M, Wang ZD, et al. Comprehensive profiling and quantitation of oncogenic mutations in non-small cell lung carcinoma using single-molecule amplification and re-sequencing technology. *Tumour Biol*, 2017, 39(2): 1–6.
- [18] Flusberg BA, Webster DR, Lee JH, et al. Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nat Methods*, 2010, 7(6): 461–465.
- [19] Tommas S, Danza K, Pilato B, et al. Innovative technology for cancer risk analysis. *Ann Oncol*, 2011, 22 Suppl 1: i37–i43.
- [20] Rhoads A, Au KF. PacBio sequencing and its applications. *Genomics Proteomics Bioinformatics*, 2015, 13(5): 278–289.
- [21] Quail MA, Smith M, Coupland P, et al. A tale of three next generation sequencing platforms: comparison of ion torrent, pacific biosciences and illumina MiSeq sequencers. *BMC Genomics*, 2012, 13: 341.
- [22] Levy SE, Myers RM. Advancements in next-generation sequencing. *Annu Rev Genomics Hum Genet*, 2016, 17: 95–115.
- [23] Nakamura K, Oshima T, Morimoto T, et al. Sequence-specific error profile of Illumina sequencers. *Nucleic Acids Res*, 2011, 39(13): e90.
- [24] De S, Ganesan S. Looking beyond drivers and passengers in cancer genome sequencing data. *Ann Oncol*, 2017, 28(5): 938–945.
- [25] Milos PM. Emergence of single-molecule sequencing and potential for molecular diagnostic applications. *Expert Rev Mol Diagn*, 2009, 9(7): 659–666.
- [26] Harbeck N, Gnant M. Breast cancer. *Lancet*, 2017, 389(10074): 1134–1150.
- [27] Li X, Dai DN, Chen B, et al. Clinicopathological and prognostic significance of cancer antigen 15–3 and carcinoembryonic antigen in breast cancer: a meta-analysis including 12, 993 patients. *Dis Markers*, 2018, 2018: 9863092.
- [28] Hiremath PJ, Kumar A, Penmetsa RV, et al. Large-scale development of cost-effective SNP marker assays for diversity assessment and genetic mapping in chickpea and comparative mapping in legumes. *Plant Biotechnol J*, 2012, 10(6): 716–732.
- [29] Barbosa C, Peixeiro I, Romão L. Gene expression regulation by upstream open reading frames and human disease. *PLoS Genet*, 2013, 9(8): e1003529.
- [30] Fuxman Bass JI, Sahni N, Shrestha S, et al. Human gene-centered transcription factor networks for enhancers and disease variants. *Cell*, 2015, 161(3): 661–673.
- [31] Gabrieli T, Sharim H, Fridman D, et al. Selective nanopore sequencing of human BRCA1 by Cas9-assisted targeting of chromosome segments (CATCH). *Nucleic Acids Res*, 2018, 46(14): e87.
- [32] 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.
- [33] Romano G, Veneziano D, Acunzo M, et al. Small non-coding RNA and cancer. *Carcinogenesis*, 2017, 38(5): 485–491.
- [34] Slaby O, Laga R, Sedlacek O. Therapeutic targeting of non-coding RNAs in cancer. *Biochem J*, 2017, 474(24): 4219–4251.
- [35] Soudyab M, Iranpour M, Ghafouri-Fard S. The role of long non-coding RNAs in breast cancer. *Arch Iran Med*, 2016, 19(7): 508–517.
- [36] Klinge CM. Non-coding RNAs in breast cancer: intracellular and intercellular communication. *Non-Coding RNA*, 2018, 4(4): 40.
- [37] Jonsson P, Coarfa C, Mesmar F, et al.



- Single-molecule sequencing reveals estrogen-regulated clinically relevant lncRNAs in breast cancer. *Mol Endocrinol*, 2015, 29(11): 1634–1645.
- [38] Shi C, Zheng Y, Li Y, et al. Association between clinical characteristics and the diagnostic accuracy of circulating single-molecule amplification and resequencing technology on detection epidermal growth factor receptor mutation status in plasma of lung adenocarcinoma. *J Clin Lab Anal*, 2018, 32(2): e22271.
- [39] Calabrese F, Lunardi F, Pezzuto F, et al. Are there new biomarkers in tissue and liquid biopsies for the early detection of non-small cell lung cancer? *J Clin Med*, 2019, 8(3): 414.
- [40] Yi X, Ma JH, Guan YF, et al. The feasibility of using mutation detection in ctDNA to assess tumor dynamics. *Int J Cancer*, 2017, 140(12): 2642–2647.
- [41] Lin Y, Wu Z, Guo W, et al. Gene mutations in gastric cancer: a review of recent next-generation sequencing studies. *Tumour Biol*, 2015, 36(10): 7385–7394.
- [42] Ma QC, Ennis CA, Aparicio S. Opening Pandora's box--the new biology of driver mutations and clonal evolution in cancer as revealed by next generation sequencing. *Curr Opin Genet Dev*, 2012, 22(1): 3–9.
- [43] Bartels S, Persing S, Hasemeier B, et al. Molecular analysis of circulating cell-free DNA from lung cancer patients in routine laboratory practice: a cross-platform comparison of three different molecular methods for mutation detection. *J Mol Diagn*, 2017, 19(5): 722–732.
- [44] DiBardino DM, Rawson DW, Saqi A, et al. Next-generation sequencing of non-small cell lung cancer using a customized, targeted sequencing panel: emphasis on small biopsy and cytology. *Cytojournal*, 2017, 14: 7.
- [45] Mills K, Fuh K. Recent advances in understanding, diagnosing, and treating ovarian cancer. *F1000Res*, 2017, 6: 84.
- [46] Ayhan A, Kuhn E, Wu RC, et al. *CCNE1* copy-number gain and overexpression identify ovarian clear cell carcinoma with a poor prognosis. *Mod Pathol*, 2017, 30(2): 297–303.
- [47] The Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. *Nature*, 2011, 474(7353): 609–615.
- [48] Patch AM, Christie EL, Etemadmoghadam D, et al. Whole-genome characterization of chemoresistant ovarian cancer. *Nature*, 2015, 521(7553): 489–494.
- [49] Roberts RJ, Carneiro MO, Schatz MC. The advantages of SMRT sequencing. *Genome Biol*, 2013, 14(7): 405.
- [50] Jing Y, Zhang Y, Zhu H, et al. Hybrid sequencing-based personal full-length transcriptomic analysis implicates proteostatic stress in metastatic ovarian cancer. *Oncogene*, 2019, 38(16): 3047–3060.
- [51] Winterhoff BJ, Maile M, Mitra AK, et al. Single cell sequencing reveals heterogeneity within ovarian cancer epithelium and cancer associated stromal cells. *Gynecol Oncol*, 2017, 144(3): 598–606.
- [52] Miyahira AK, Soule HR. The 23rd annual prostate cancer foundation scientific retreat report. *Prostate*, 2017, 77(10): 1093–1106.
- [53] Caram MEV, Skolarus TA, Cooney KA. Limitations of prostate-specific antigen testing after a prostate cancer diagnosis. *Eur Urol*, 2016, 70(2): 209–210.
- [54] de Kok JB, Verhaegh GW, Roelofs RW, et al. *DD3<sup>PCA3</sup>*, a very sensitive and specific marker to detect prostate tumors. *Cancer Res*, 2002, 62(9): 2695–2698.
- [55] Zhang P, Xia JH, Zhu J, et al. High-throughput screening of prostate cancer risk loci by single nucleotide polymorphisms sequencing. *Nat Commun*, 2018, 9(1): 2022.
- [56] Su F, Zhang W, Zhang DL, et al. Spatial intratumor genomic heterogeneity within localized prostate cancer revealed by single-nucleus sequencing. *Eur Urol*, 2018, 74(5): 551–559.
- [57] Tevz G, McGrath S, Demeter R, et al. Identification of a novel fusion transcript between human relaxin-1 (*RLN1*) and human relaxin-2 (*RLN2*) in prostate cancer. *Mol Cell Endocrinol*, 2016, 420: 159–168.

(本文责编 郝丽芳)