



专论与综述

单分子实时测序技术在环境微生物研究中的应用

韩迎亚 杨乔乔 王倩楠 安贤惠 刘有华 李联泰*

江苏海洋大学 江苏 连云港 222005

摘要: 1975 年, 第一个 cDNA 的基因组——噬菌体 ϕ X174 基因组的测序完成, 标志着测序时代的开始。1996 年以来人类基因组计划的发起极大地推动了测序技术的应用和发展, 第二代测序技术也被称为高通量测序技术。而单分子 DNA 测序技术是最近 10 年内发展起来的新一代的测序技术, 称为第三代测序技术, 其中包括单分子实时测序(Single molecule real time sequencing, SMRT)、真正单分子测序和单分子纳米孔测序等技术。SMRT 测序技术有超长读长、测序周期短、不需要模板扩增和直接检测表观修饰位点等特点, 为研究人员提供了新的选择。本文综述了 SMRT 测序技术的基本原理、性能以及它在微生物 16S rRNA 基因、微生物全基因组以及微生物宏基因组测序等方面的应用, 并分析了 SMRT 测序技术在环境微生物应用中的优势以及存在的问题。

关键词: Sanger 测序法, SMRT 测序技术, 基因组学, 环境微生物

Application of single molecule real time sequencing in environmental microorganisms research

HAN Ying-Ya YANG Qiao-Qiao WANG Qian-Nan AN Xian-Hui LIU You-Hua
LI Lian-Tai*

Jiangsu Ocean University, Lianyungang, Jiangsu 222005, China

Abstract: In 1975, the sequencing of the first cDNA genome, the phage ϕ X174 genome, marked the beginning of the sequencing era. Since 1996, the launch of the Human Genome Project has greatly promoted the application and development of sequencing technology. The second generation sequencing technology is also known as high-throughput sequencing technology. Single-molecule DNA sequencing technology is a new generation of sequencing technology called third-generation sequencing technology, which developed in the last 10 years. It includes single-molecule real-time sequencing, true single-molecule sequencing and single-molecule nanopore sequencing. Because SMRT sequencing technology has the characteristics of long reading length, short sequencing period, without requirement for template amplification and direct detection of apparent modification sites, it provides researchers with new

Foundation items: Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD); Postgraduate Research & Innovation Programs of Jiangsu (KYCX18_2569); Postgraduate Research & Practice Innovation Programs of Jiangsu (SJCX19_0989, SJCX19_0990)

*Corresponding author: Tel: 86-518-85895427; E-mail: lilt@hhit.edu.cn

Received: 10-10-2018; **Accepted:** 31-07-2019; **Published online:** 17-09-2019

基金项目: 江苏高校优势学科建设资助(PAPD); 江苏省研究生科研创新计划(KYCX18_2569); 江苏省研究生科研与实践创新计划(SJCX19_0989, SJCX19_0990)

*通信作者: Tel: 0518-85895427; E-mail: lilt@hhit.edu.cn

收稿日期: 2018-10-10; **接受日期:** 2019-07-31; **网络首发日期:** 2019-09-17

options. This paper reviews the basic principle and performance of SMRT sequencing technology and its application in microbial 16S rRNA gene, microbial whole genome sequencing and microbial macrogenome sequencing, and analyzes the advantages and problems of SMRT sequencing technology in environmental microorganisms.

Keywords: Sanger sequencing, SMRT sequencing technology, Genomics, Environmental microorganisms

DNA 序列是生命遗传和进化的基础。为了对 DNA 进行测序, 1977 年 Maxam 和 Gilbert 发明了化学降解法^[1]。同年, Sanger 的双脱氧终止法问世^[2], 引发了生物学革命, 这些测序方法也被称为第一代测序技术。应用这些方法最终完成了迄今为止世界上最大的生物协作项目: 人类基因组计划^[3]。

Sanger 测序法对现代分子生物学研究起到了极大的促进作用。随着大规模基因组学的兴起, 各种高通量、低成本的测序技术应运而生。21 世纪, 第二代测序技术^[4]和第三代测序技术^[5]相继问世。近 10 年, 第二代测序技术出现并且发展成熟, 包括焦磷酸测序技术(454)^[6]、Solexa 测序技术^[7]以及 Solid 测序技术^[8]。因为这些平台测序的高通量及低价格, 引起生命科学研究方法的大变革^[9], 极大地推进了基因组学的发展(表 1)。然而, 其不足之处也日益凸显。第二代测序技术读长相对较短, 这给组装基因组带来了巨大困难; 同时, 由于该技术是建立在 PCR 基础上的, 扩增后的 DNA 分子片段数目与扩增前有所偏差, 对基因表达分析也有很大影响^[16]。

第二代测序技术出现不久, 最新一代的测序方法应运而生, 即单分子测序技术(Single molecule sequencing, SMS), 也被称为第三代测序技术(Third generation sequencing, TGS), 因为其采用单分子读取技术, 有着更快的读取速度, 同时不需要 PCR 扩增, 进一步降低了测序的成本。目前主流的第三代测序技术主要包括单分子纳米孔测序技术、真正单分子测序技术和单分子实时测序技术(SMRT)^[17]。其中单分子纳米孔测序技术以超长读长和轻便见长, 但由于其测序错误率过高无法在研究中推广; 而真正单分子测序由于测序技术费用偏高, 项目基本处于停滞状态^[14]。大部分第三代测序技术仍处于研发的阶段, 只有 SMRT 测序相对成熟。目前微生物多样性研究中以第二代测序为主, 而基因组测序以及环境总基因测序以第三代测序为主。但随着 SMRT 测序技术的进一步发展, 其必将成为微生物学研究中不可缺少的一环。因此, 系统地了解 SMRT 测序技术对微生物研究有着重要意义。本文介绍了 SMRT 测序技术的原理性能及其在环境微生物中的应用。

表 1 测序技术发展比较

Table 1 Development and comparison of sequencing technology

测序技术		优点	缺点	应用	主要平台	参考文献
Sequencing technology		Advantages	Disadvantages	Application	Main platform	References
Sanger	Maxam-Gilbert	Longer reads;	Low output;	Single gene sequencing	ABI 3730 XL	[10]
	Sanger	High accuracy	High cost			
NGS	Roche 454	High output;	Shorter reads;	Genome sequencing	454 Sequencing	[11]
	Solexa	High accuracy	Assembly is difficult		HiSeq-2000	[12]
	Solid				MiSeq	[13]
	Ion Torrent				HiSeq-2500 Ion Torrenttm	[6]
TGS	SMRT	Longer reads;	High cost;	Genome sequencing	SMRT	[14]
	SMS	No GC-skew;	High random error rate;		Heliscope/Helicos	[15]
	NSMS	Read time in real-time	Lack of analysis software		NSM	

1 SMRT 测序技术原理

2009 年, Pacific Biosciences 公司推出了 SMRT 测序技术。和其他的两个单分子测序技术原理相同, SMRT 测序技术也是采用边合成边测序的策略。它的原理是:当 DNA 模板被聚合酶捕获后, 4 种不同荧光标记的 dNTP 通过布朗运动随机进入检测区域并与聚合酶结合, 与模板匹配的碱基生成化学键的时间远远长于其他碱基停留的时间。因此统计荧光信号存在时间的长短, 可区分匹配的碱基与游离碱基。通过统计 4 种荧光信号与时间的关系图, 即可测定 DNA 模板序列^[17]。

SMRT 测序技术的核心是零膜波导孔(Zero mode waveguide, ZMW), 它是一种直径为 20 nm–50 nm 的纳米孔, 底部固定有 DNA 聚合酶^[18]。它的直径远小于检测激光的波长, 因此当激光打在 ZMW 底部时, 激光无法穿过, 在 ZMW 的底部发生衍射, 只可以照亮很小的区域。DNA 聚合酶就被固定在这里。只有在此波长处, 碱基携带的荧光基团才能被激活从而被检测到。每个 ZMW 都只固

定一个 DNA 聚合酶, 当一个 ZMW 结合少于或超过一个 DNA 模板时, 它所产生的测序结果在后续数据分析时被过滤掉, 由此来保证每个可用的 ZMW 都是一个单独的 DNA 合成体系(图 1)。

2 SMRT 的测序特点

2.1 超长的读长

2011 年初, PacBio 发布了 PacBio RS 测序仪。虽然最初的平均读长相对较短, 平均错误率高^[20], 但近年来该技术有了很大的改进。2013 年 Ferrarini 等^[21]使用 PacBio RS 平台对 *Potentilla micrantha* 叶绿体基因组进行测序, 在纠错后, 共回收了 28 638 个 PacBio RS 读数, 平均读长为 1 902 bp, 由于更长的读长和更低的 GC 偏差, 较第二代测序有着更高的准确度。目前为止, PacBio 公司基于 SMRT 测序技术共推出了 3 款测序仪, 第一款产品 PacBio RS 在 2011 年正式发布并商用, 平均读长 3 kb 左右; 2013 年 4 月发布了升级版 PacBio RS II, 平均读长为 10–15 kb; 2015 年 10 月推出全新升级的三代测序仪 PacBio Sequel 测序系统, 平均读长为 8–12 kb (表 2)^[22]。

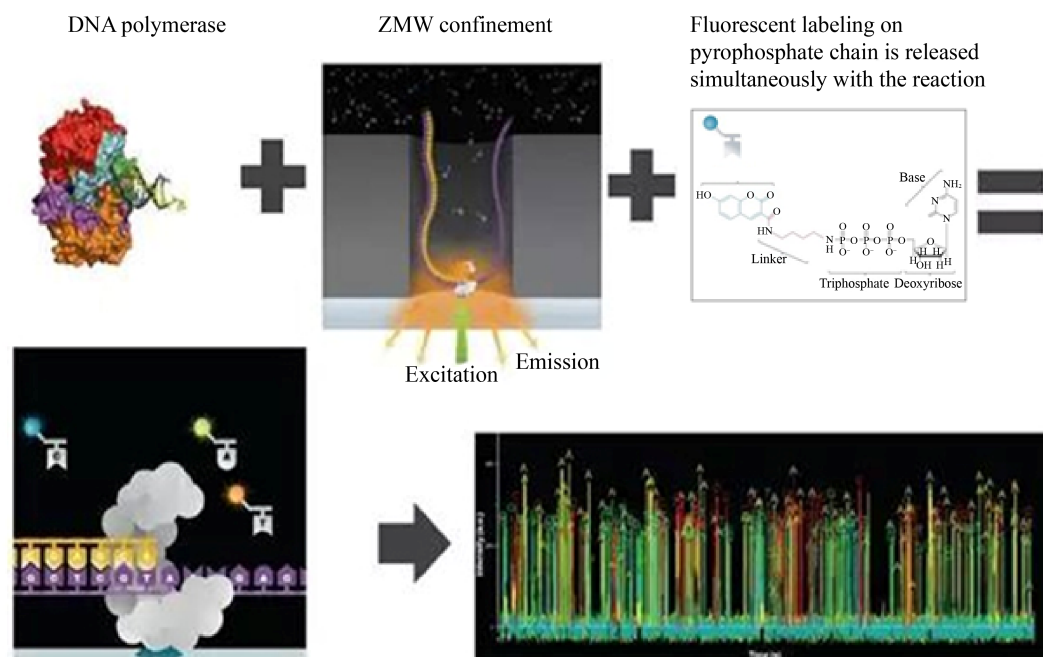


图 1 SMRT 测序技术原理^[19]

Figure 1 Principle of SMRT sequencing technology^[19]

表 2 PacBio RS II 与 PacBio Sequel 的比较^[11,22]Table 2 Comparison between PacBio RS II and PacBio Sequel^[11,22]

Type	PacBio RS II	PacBio Sequel
Average reads length	10–15 kb	8–12 kb
Run time	240 min	240 min
Error rate	10%–15%	10%–15%
Total output	500 Mb–1 Gb	4–7 Gb
Accuracy	99.99%	99.99%
Of reads	About 50 k	About 500 k
Run price	\$400	\$850
Application	Sequencing of small genomes such as microorganisms and target regions	Genome denovo assembly; Animal and plant genome re-sequencing and full-field transcriptome sequencing

2.2 测序的速度较快

测序速度快是 SMRT 测序技术的另一个特点。与第二代测序技术相比,它简化了建库和测序步骤。相比动辄运行数天的第二代测序技术,SMRT 测序每个 Run 运行时间最短近 30 min^[19]。在对时间有要求的情况下,SMRT 有很大的优势。较短的运行时间对于应对传染病暴发尤为重要,在短时间内得到变异微生物的基因组,可为快速和准确地研究暴发起因以及治疗策略提供基础^[23]。

2.3 无需模板扩增

基因组的 GC 含量直接影响到 DNA 序列的测定。第二代测序技术中文库构建和测序过程中都有 PCR 扩增步骤,高 GC 或低 GC 含量的基因组区域不易被 PCR 扩增,导致在测序过程中测序覆盖度不足^[24]。另外,在文库构建时需将 DNA 打断成适当大小的片段,由于高 GC 含量区域不易被打断,使得这些片段过大而在长度筛选时被舍弃^[25]。由于 SMRT 测序是单分子测序技术,没有 PCR 扩增步骤^[26],结合 SMRT 测序超长读长的特点,可以完成长片段的高 GC 含量区域测序,从而帮助高 GC 含量基因组完成组装。SMRT 测序这一优势很好地应用到极端微生物的基因组研究中^[27]。同时,无需模板扩增步骤还避免了 PCR 引入的错误,并且只需要使用极少的荧光基团,为今后大幅降低测序试剂成本提供了空间。

2.4 测序存在随机性的错误

SMRT 技术的测序速度很快,每秒约 10 个 dNTP,但快速的测序也带来了明显的缺点——错误偏高,这几乎是目前单分子测序技术的通病。SMRT 测序的错误率大概为 15%,碱基错测率大约为 1%,其他的错误主要是单碱基的插入和缺失。通过校准错误来提高序列准确性需要较高覆盖度和占用大量的计算机网络资源,这相对来说比较难以实施,其最根本的解决方法是通过革新技术来提高测序的准确性。SMRT 的测序错误都是随机的,并非系统错误,系统错误是无法通过提高测序覆盖度校准的^[28]。

3 SMRT 测序在环境微生物研究中的应用

目前 SMRT 测序技术应用于微生物 16S rRNA 基因测序、微生物基因组组装、转录组测序、甲基化分析和基因组重测序等方面。

3.1 微生物 16S rRNA 基因测序

SMRT 在微生物 16S rRNA 基因中应用广泛。2006 年, Sogin 等^[29]首次成功地将高通量测序用于深海环境微生物多样性调查,一直以来 16S rRNA 基因高通量测序片段的选择都存在争议^[30],全长 DNA 测序彻底终止了这一争论。SMRT 测序技术在复杂环境微生物的研究中所具备的优势已经多次被证实, Mosher 等^[31]通过不断完善 PacBio RS 系统和 SMRT 平台提高了 16S rRNA 基因扩增的准确性,并有望将环境微生物鉴定到种水平。2013 年 Hu 等^[32]采集了多位由不同方式分娩的新生儿的粪便,采用 PacBio RS 平台对粪便中微生物 16S rRNA 基因进行测序,结果显示新生儿胎粪的细菌群落丰度与成人粪便的样品丰度不同,非糖尿病母亲所生新生儿胎粪中变形菌门(*Proteobacteria*)比例明显较高(胎粪为 71%,成年人粪便为 3.1%),拟杆菌门(*Bacteroidetes*)的比例较低(胎粪为 2.5%,成年人粪便为 42.8%)。与非糖尿病组相似,来自糖尿病与妊娠期糖尿病母亲所生的新生儿胎粪较成人粪便也显示出变形菌门的富集以及拟杆菌门的减少。由

此可见,不同分娩方式对新生儿的粪便微生物没有影响,而母亲患病状态对新生儿肠道微生物组成有显著的影响。

2016年,本研究组 Li 等^[33]从海水虾蟹贝混养池内采集 6 个海水样品,与同期采集的非养殖池中 6 个样品进行对比;通过 Illumina MiSeq 测序分析了海水虾蟹贝混养池中细菌群落的多样性;研究发现,与非养殖池相比,海水养殖明显改变了养殖池中细菌群落的结构和丰度,观察到新增的门有 7 个,分别为迷踪菌门(*Elusimicrobia*)、纤维杆菌门(*Fibrobacteres*)、梭杆菌门(*Fusobacteria*)、柔膜菌门(*Tenericutes*)、MVP-21、SM2F11 和 WCHB1-60;减少或消失的门有 5 个,分别为 Candidate_division_OD1、Candidate_division_OP11、Candidate_division_BRC1、黏胶球形菌纲(*Lentisphaerae*)、脱铁杆菌门(*Deferribacteres*)。研究结果有助于更好地了解复杂水产养殖生态系统中细菌群落的结构和丰度,有望在改善海水养殖水质和疾病早期预警等方面提供有用的信息。然而,由于二代测序读长短的特性,无法获取物种全部相关代谢功能信息,从而限制了我们对菌群代谢功能的深入理解,随着第三代 SMRT 测序技术的出现,我们有望从根本上解决这一瓶颈,通过对 16S rRNA 基因全长序列的高通量精准测序,更深入地了解海水养殖环境细菌群落结构组成及代谢的多样性。

2016 年 Singer 等^[34]采用 PacBio RS II 和 Illumina MiSeq 两种测序平台对 23 种细菌和 3 种古菌组成的人工菌群以及取自湖水的天然微生物群落样本进行了群落多样性分析;结果显示,门阶元的菌群组成较为相似,但在更小的阶元如属和种水平存在差异;SMRT 测序的结果显著降低了物种分类信息的不确定性;计算机模拟的结果也表明,短独长的单 V 区测序可能严重低估某些特定类群的微生物,比如水体样本中参与氮循环和甲烷代谢的物种。因此,运用 SMRT 测序技术进行微生物研究将显著提升菌群多样性组成谱解析的精准性

和全面性。Gesudu 等^[35]利用 SMRT 测序技术研究马奶酒中细菌多样性和发酵过程中细菌种群的演替动态,结果表明细菌群落的结构差异可归因于地理位置。Mosher 等^[31]使用 PacBio RS II 及 P4/C2 试剂测得沉淀物样本微生物 16S rRNA 基因片段平均长度为 1 419–1 431 bp。在 Benítez-Páez 等^[36]和 Schloss 等^[37]的研究中,使用全长 16S rRNA 基因序列对物种多样性、微生物组成和微生物进化开展研究,对实验的准确性和分辨率能够带来显著的提升。同时,将研究深入到种水平,而不是局限在属水平上^[38]。2017 年, Cao 等^[39]在重庆 7 个地区获取了 38 份 10 年以上的泡菜盐水样品,并运用 SMRT 测序技术分组对其 16S rRNA 基因全长进行测序,结果显示,酸度越低,物种多样性越高,乳酸菌属物种也越多,且较高酸度而言更易使人致病。2017 年 Jia 等^[40]利用 SMRT 测序技术检测了中药九味强效丸的多种成分,成功地恢复了属于 ITS2 和 psbA-trnH 区域的 5 416 个和 4 342 个环状共有序列读数,结果表明,SMRT 可以检测传统草药产品中的物质,SMRT 测序中的误差不会影响识别多种处方物种和几种掺杂物的能力。可能成为生物检测有价值的工具。

3.2 微生物全基因组测序

在无参考序列情况下,凭借生物信息学分析方法直接对物种序列进行拼接、组装,最终获得该物种的基因组图谱,称为全基因组测序^[41]。全基因组测序有利于人们深入了解物种的基因组组成以及分子进化,目前由 SMRT 测序技术完成的完整基因组测序有很多。2017 年 Lhee 等^[42]通过形态和分子相结合的方法从 KR01 菌株中发现了具有光合作用的新物种(*Paulinella micropora* sp. nov.), 菌株编号为: KX897545;通过构建 SMRTbell 库,并使用 PacBio RS II 测序平台测序获得 16 Gb 的原始数据,然后通过分级基因组组装过程(Hierarchical genome-assembly process, HGAP)^[43]获取基因组全长,并经由 Illumina HiSeq 测序平台进一步确定,最终通过比较证实为新的种^[42]。2017 年, van Kan

等^[44]采用 SMRT 测序技术和第二代测序技术相结合的方式对葡萄孢菌全基因组测序, 得到由 18 条染色体组装的新基因组, 测序深度和完整性得到了大幅度的提高。2017 年 Hui 等^[45]使用 SMRT 测序技术分析了 3 种清酒曲中的细菌微生物群, 结果产生总共 39 121 个高质量序列, 包括 14 354 个细菌和 24 767 个真菌序列读数, 属于 5 个细菌门和 2 个真菌门, 分别由变形菌门和子囊菌门(*Ascomycota*)控制; 在属的水平上, 苍白杆菌属(*Ochrobactrum*)和海洋嗜杀酵母属(*Wickerhamomyces*)分别为最丰富的细菌和真菌属; 主要的细菌和真菌物种分别是苍白杆菌(*Ochrobactrum lupini*)和海洋嗜杀酵母(*Wickerhamomyces anomalus*); 本研究将 3 种清酒曲样品的微生物群组成分析为物种水平的精确度。该结果可用于进一步开发传统的发酵产品, 尤其是制曲的优化。同时, 这项研究表明, SMRT 是分析食品样品中微生物成分的有力工具。Chan 等^[46]通过第三代测序技术报告了一种来自鲭鱼生鱼片中分离的奈氏西地西菌 SSMD04 的完整基因组序列, 这是第一份关于此菌种完整的基因组文献。同样, Wibberg 等^[47]利用第三代测序技术完成了对假单胞菌属产碱假单胞菌(*Pseudomonas pseudoalcaligenes*) CECT5344 的完整基因组测序; 由于该菌可将氰化物消化并转变为无毒化合物, 所以可以将其应用到含氰化物废物的处理中; 此次产碱假单胞菌的完整测序使我们对其基因组有了充分认识, 预示着它在未来污染物的生物降解中可发挥重要作用。

3.3 微生物环境总基因组测序

微生物环境总基因组学是指某一特定环境中所有微生物基因组的总和^[48]。传统的培养方法并不能代表自然界中微生物多样性的范围。为了更好地研究微生物的多样性, Rondon 等^[49]通过构建宏基因组文库研究了土壤微生物的多样性, 揭开了环境微生物研究新的篇章。随着高通量测序价格大幅下跌, 获得大批原始宏基因组测序数据已经不再是难题, 而真正的研究瓶颈在于数据分析环节。其中,

微生物参考基因组缺乏是宏基因组数据分析主要障碍。目前, 已有参考基因组的微生物数量与自然界存在的微生物数量相去甚远。因此, 从复杂的宏基因组数据中完整而准确地构建微生物基因组草图成为分析流程的首要任务^[50]。第二代测序技术由于测序片段短的问题导致组装困难, 第三代测序技术有望彻底解决这一问题。

2016 年, Frank 等^[51]使用代表复杂微生物群落的宏基因组样本, 将 PacBio CCS 和 Illumina HiSeq 数据的应用性能和分类学算法进行了比较; 采用了两种平台相结合的方法对 8 个沼气反应器内的微生物宏基因组进行研究, 分别单独组装与两种方式混合组装; 结果表明, 混合组装的方式得到的组装序列长度高于单独组装; 如在长度大于 25 kb 重叠群的累积核苷酸方面, PacBio 和 HiSeq 的总和为 21.01 Mb 而混合组装的为 26.8 Mb。该试验结果表明 SMRT 测序技术对微生物宏基因组研究有提高作用。2016 年, Singer 等^[34]应用 SMRT 进行微生物群落分析; 主要通过 PacBio 三代全长 16S rRNA 基因和二代 16S rRNA 基因 V4 区来研究加拿大 Sakinaw 湖水样本的微生物多样性, 结果显示, 三代全长在种属鉴定上多样性更丰富, 提高了种属鉴定的分辨率和准确度; 同时, 三代全长 16S rRNA 基因鉴定出了更多的在湖水不同深度下关于氮循环和产甲烷相关的菌种, 而在二代测序中这种微生物的丰度被严重低估。2017 年, Driscoll 等^[52]从美国俄勒冈州克拉马斯湖中采集细菌并共培养, 然后采用 PacBio 测序平台进行宏基因组测序, 成功组装出 3 个微生物基因组草图, 研究表明 SMRT 测序技术在低多样性微生物群落宏基因组组装中具有可行性。2018 年 Wang 等^[53]通过 SMRT 测序的方法分析了奶粉在两种不同生产环境中的细菌群落; 即小规模生产的室内环境和大规模的工厂生产, 相比大规模的工厂生产, 室内环境生产过程中的细菌群落变化相对较小, 而大规模的工厂生产导致微生物含量明显变化。这种技术的实施可以增强目前在乳制品加工中食品质量的保证。

4 小结

近 10 年来, DNA 的测序技术进入了突飞猛进的时代, 各种测序仪相继问世。从第一代到第三代都各有优势。Sanger 测序通量低、读长较长、准确率高, 对于一些小量测序仍是最佳选择。而高通量、成本低的第二代测序技术已经发展成熟, 这在大型基因组测序和重测序中得到了广泛的使用。SMRT 测序则以众多优势在微生物的研究中得到广泛应用, 并且将在更多的领域发挥重要作用。在将来的一段时间内, 这些测序技术都会在不同的领域各自发挥优势, 并且相互补足^[16]。由于前两代测序技术的一些不足, 人们越来越关注单分子测序。以 SMRT 测序技术为代表的第三代测序技术终将会取代第二代测序技术, 成为环境微生物研究的主流手段。但它还存在测序费用较高、错误率偏高等一些问题, 另外 SMRT 测序技术需要用到酶, 所以如何保持酶的活性与稳定性仍然是一个重要的问题; 而同时, SMRT 技术需要荧光标记, 怎样提高信号灵敏度同时又不能把信号变成噪音增加荧光背景也是一个需要解决的问题。人们还没有真正将 SMRT 测序技术用于复杂环境微生物的研究, 因此, 这方面还有很大的空间可以进行探索。

REFERENCES

- [1] Maxam AM, Gilbert W. A new method for sequencing DNA[J]. Proceedings of the National Academy of Sciences of the United States of America, 1977, 74(2): 560-564
- [2] Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors[J]. Proceedings of the National Academy of Sciences of the United States of America, 1977, 74(12): 5463-5467
- [3] Chen Z, Huang W, Fu G, et al. Human genome project progress and prospect[J]. Journal of Nature, 2000, 22(3): 125-133, 148 (in Chinese)
陈竺, 黄薇, 傅刚, 等. 人类基因组计划现状与展望[J]. 自然杂志, 2000, 22(3): 125-133, 148
- [4] Metzker ML. Sequencing technologies—the next generation[J]. Nature Reviews Genetics, 2010, 11(1): 31-46
- [5] Schadt EE, Turner S, Kasarskis A. A window into third-generation sequencing[J]. Human Molecular Genetics, 2011, 19(R2): R227-R240
- [6] Margulies M, Egholm M, Altman WE, et al. Genome sequencing in microfabricated high-density picolitre reactors[J]. Nature, 2005, 437(7057): 376-380
- [7] Bentley DR, Balasubramanian S, Swerdlow HP, et al. Accurate whole human genome sequencing using reversible terminator chemistry[J]. Nature, 2008, 456(7218): 53-59
- [8] Valouev A, Ichikawa J, Tonthat T, et al. A high-resolution, nucleosome position map of *C. elegans* reveals a lack of universal sequence-dictated positioning[J]. Genome Research, 2008, 18(7): 1051-1063
- [9] Reis-Filho JS. Next-generation sequencing[J]. Breast Cancer Research, 2009, 11(3): S12
- [10] Ge YJ, Chen S. Single-molecule real-time sequencing and its applications in microbial epigenetics—a review[J]. Microbiology China, 2017, 44(1): 186-199 (in Chinese)
葛元洁, 陈实. 单分子实时测序及其在微生物表观遗传学中的应用[J]. 微生物学通报, 2017, 44(1): 186-199
- [11] Liu YH, Wang L, Yu L. The principle and application of the single-molecule real-time sequencing technology[J]. Hereditas (Beijing), 2015, 37(3): 259-268 (in Chinese)
柳延虎, 王璐, 于黎. 单分子实时测序技术的原理与应用[J]. 遗传, 2015, 37(3): 259-268
- [12] Tang Y, Liu X. SMRT sequencing and its application in microorganism studies[J]. Biotechnology Bulletin, 2018, 34(6): 48-53 (in Chinese)
唐勇, 刘旭. SMRT 测序技术及其在微生物研究中的应用[J]. 生物技术通报, 2018, 34(6): 48-53
- [13] Lin L, Li YH, Li SL, et al. Comparison of next-generation sequencing systems[J]. Journal of Biomedicine and Biotechnology, 2012, 2012: 251364
- [14] Porreca GJ, Zhang K, Li JB, et al. Multiplex amplification of large sets of human exons[J]. Nature Methods, 2007, 4(11): 931-936
- [15] Ondov BD, Varadarajan A, Passalacqua KD, et al. Efficient mapping of applied biosystems SOLiD sequence data to a reference genome for functional genomic applications[J]. Bioinformatics, 2008, 24(23): 2776-2777
- [16] Wang XC, Yang ZR, Wang M, et al. High-throughput sequencing technology and its application[J]. China Biotechnology, 2012, 32(1): 109-114 (in Chinese)
王兴春, 杨致荣, 王敏, 等. 高通量测序技术及其应用[J]. 中国生物工程杂志, 2012, 32(1): 109-114
- [17] Zhang DF, Ma QY, Yin TM, et al. The third generation sequencing technology and its application[J]. China Biotechnology, 2013, 33(5): 125-131 (in Chinese)
张得芳, 马秋月, 尹佟明, 等. 第三代测序技术及其应用[J]. 中国生物工程杂志, 2013, 33(5): 125-131
- [18] Levene MJ, Korlach J, Turner SW, et al. Zero-mode waveguides for single-molecule analysis at high concentrations[J]. Science, 2003, 299(5607): 682-686
- [19] Rhoads A, Au KF. PacBio sequencing and its applications[J]. Genomics, Proteomics & Bioinformatics, 2015, 13(5): 278-289
- [20] 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[J]. BMC Genomics, 2012, 13: 341
- [21] Ferrarini M, Moretto M, Ward JA, et al. An evaluation of the PacBio RS platform for sequencing and *de novo* assembly of a chloroplast genome[J]. BMC Genomics, 2013, 14: 670
- [22] PacBio. Sequence with confidence[EB/OL]. <http://www.pacificbiosciences.com>

- [23] Liu Y, Wu BQ. Third-generation sequencing technology: single-molecule real-time sequencing[J]. Chinese Journal of Pathology, 2011, 40(10): 718-720 (in Chinese)
刘岩, 吴秉铨. 第三代测序技术: 单分子即时测序[J]. 中华病理学杂志, 2011, 40(10): 718-720
- [24] Aird D, Ross MG, Chen WS, et al. Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries[J]. Genome Biology, 2011, 12(2): R18
- [25] Niu BF, Fu LM, Sun SL, et al. Artificial and natural duplicates in pyrosequencing reads of metagenomic data[J]. BMC Bioinformatics, 2010, 11: 187
- [26] Schadt EE, Turner S, Kasarskis A. A window into third-generation sequencing[J]. Human Molecular Genetics, 2010, 19(R2): R227-R240
- [27] Shin SC, Ahn DH, Kim SJ, et al. Advantages of single-molecule real-time sequencing in high-GC content genomes[J]. PLoS One, 2013, 8(7): e68824
- [28] Tang Y, Liu X. Full-length sequencing of 16S rRNA gene and its analysis based on the SMRT sequencing technology[J]. Biotechnology Bulletin, 2017, 33(8): 34-39 (in Chinese)
唐勇, 刘旭. 基于 SMRT 测序技术的 16S rRNA 基因全长测序及其分析方法[J]. 生物技术通报, 2017, 33(8): 34-39
- [29] Sogin ML, Morrison HG, Huber JA, et al. Microbial diversity in the deep sea and the underexplored "rare biosphere"[J]. Proceedings of the National Academy of Sciences of the United States of America, 2006, 103(32): 12115-12120
- [30] Chakravorty S, Helb D, Burday M, et al. A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria[J]. Journal of Microbiological Methods, 2007, 69(2): 330-339
- [31] Mosher JJ, Bowman B, Bernberg EL, et al. Improved performance of the PacBio SMRT technology for 16S rDNA sequencing[J]. Journal of Microbiological Methods, 2014, 104: 59-60
- [32] Hu JZ, Nomura Y, Bashir A, et al. Diversified microbiota of meconium is affected by maternal diabetes status[J]. PLoS One, 2013, 8(11): e78257
- [33] Li LT, Yan BL, Li SH, et al. A comparison of bacterial community structure in seawater pond with shrimp, crab, and shellfish cultures and in non-cultured pond in Ganyu, Eastern China[J]. Annals of Microbiology, 2016, 66(1): 317-328
- [34] Singer E, Bushnell B, Coleman-Derr D, et al. High-resolution phylogenetic microbial community profiling[J]. The ISME Journal, 2016, 10(8): 2020-2032
- [35] Gesudu Q, Zheng Y, Xi XX, et al. Investigating bacterial population structure and dynamics in traditional koumiss from Inner Mongolia using single molecule real-time sequencing[J]. Journal of Dairy Science, 2016, 99(10): 7852-7863
- [36] Benítez-Páez A, Portune KJ, Sanz Y. Species-level resolution of 16S rRNA gene amplicons sequenced through the MinION™ portable nanopore sequencer[J]. GigaScience, 2016, 5(1): 4
- [37] Schloss PD, Jenior ML, Koumpouras CC, et al. Sequencing 16S rRNA gene fragments using the PacBio SMRT DNA sequencing system[J]. PeerJ, 2016, 4: e1869
- [38] Lee CH, Bowman B, Heiner C, et al. Developments in PacBio® metagenome sequencing: shotgun whole genomes and full-length 16S[C]. International Plant and Animal Genome Conference Asia, 2014
- [39] Cao JL, Yang JX, Hou QC, et al. Assessment of bacterial profiles in aged, home-made Sichuan paocai brine with varying titratable acidity by PacBio SMRT sequencing technology[J]. Food Control, 2017, 78: 14-23
- [40] Jia J, Xu ZC, Xin TY, et al. Quality control of the traditional patent medicine Yimu Wan based on SMRT sequencing and DNA barcoding[J]. Frontiers in Plant Science, 2017, 8: 926
- [41] Cao CX, Han W, Zhang HP. Application of third generation sequencing technology to microbial research[J]. Microbiology China, 2016, 43(10): 2269-2276 (in Chinese)
曹晨霞, 韩琬, 张和平. 第三代测序技术在微生物研究中的应用[J]. 微生物学通报, 2016, 43(10): 2269-2276
- [42] Lhee D, Yang EC, Kim JJ, et al. Diversity of the photosynthetic *Paulinella* species, with the description of *Paulinella micropora* sp. nov. and the chromatophore genome sequence for strain KR01[J]. Protist, 2017, 168(2): 155-170
- [43] Chin CS, Alexander DH, Marks P, et al. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data[J]. Nature Methods, 2013, 10(6): 563-569
- [44] van Kan JA, Stassen JHM, Mosbach A, et al. A gapless genome sequence of the fungus *Botrytis cinerea*[J]. Molecular Plant Pathology, 2017, 18(1): 75-89
- [45] Hui WY, Hou QC, Cao CX, et al. Identification of microbial profile of *Koji* using single molecule, real-time sequencing technology[J]. Journal of Food Science, 2017, 82(5): 1193-1199
- [46] Chan KG, Tan KH, Yin WF, et al. Complete genome sequence of *Cedeceaneteri* strain SSMD04, a bacterium isolated from pickled mackerel sashimi[J]. Genome Announcements, 2014, 2(6): e01339-14
- [47] Wibberg D, Luque-Almagro VM, Igeño MI, et al. Complete genome sequence of the cyanide-degrading bacterium *Pseudomonas pseudoalcaligenes* CECT5344[J]. Journal of Biotechnology, 2014, 175: 67-68
- [48] Handelsman J, Rondon MR, Brady SF, et al. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products[J]. Chemistry & Biology, 1998, 5(10): R245-R249
- [49] Rondon MR, August PR, Bettermann AD, et al. Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms[J]. Applied and Environmental Microbiology, 2000, 66(6): 2541-2547
- [50] Howe AC, Chain PSG. Challenges and opportunities in understanding microbial communities with metagenome assembly (accompanied by IPython Notebook tutorial)[J]. Frontiers in Microbiology, 2015, 6: 678
- [51] Frank JA, Pan Y, Tooming-Klunderud A, et al. Improved metagenome assemblies and taxonomic binning using long-read circular consensus sequence data[J]. Scientific Reports, 2016, 6: 25373
- [52] Driscoll CB, Otten TG, Brown NM, et al. Towards long-read metagenomics: complete assembly of three novel genomes from bacteria dependent on a diazotrophic cyanobacterium in a freshwater lake co-culture[J]. Standards in Genomic Sciences, 2017, 12: 9
- [53] Wang JC, Zheng Y, Xi XX, et al. Application of PacBio single molecule real-time (SMRT) sequencing in bacterial source tracking analysis during milk powder production[J]. Food Control, 2018, 93: 226-234