

纳米孔测序在功能基因组学中的潜在应用回顾

顾章媛 莊志刚

上海市第一妇婴保健院

DOI:10.12238/bmtr.v6i5.10082

[摘要] 分子生物学领域始终受到各种先进测量技术的推动,通量的不断提升为组学分析带来了飞跃性的发展。大规模并行测序技术的进步使得研究人员能够广泛获取分子数据,从而揭示基因组与转录组的深层特征。在这场技术革命中,纳米孔测序仪将测序能力推向了一个全新的高度。由Oxford Nanopore Technologies Inc.开发的纳米孔测序仪采用蛋白质纳米孔技术,能够直接读取单分子序列,而无需传统测序过程中常见的DNA合成或扩增步骤。该技术支持对超长核苷酸序列的测序,突破了传统技术对读长的限制,且还可以识别DNA或RNA修饰。近年来,这一技术在医学、流行病学、生态学以及教育等多个领域中取得了显著成果,并逐渐成为这些领域的核心工具。在这篇综述中,我们详细探讨了纳米孔测序仪的独特优势及其功能,阐述了该技术在各个领域中的具体应用实例,并展望了它在未来可能带来的更多突破与创新。

[关键词] 长读长; 纳米孔测序; 下一代测序

中图分类号: K477 **文献标识码:** A

Review of the potential applications of nanopore sequencing in functional genomics

Zhangyuan Gu Zhigang Zhuang

Shanghai First Maternity and Infant Hospital

[Abstract] The field of molecular biology has consistently been driven by various advanced measurement technologies, with increasing throughput leading to significant advancements in omics analyses. The development of high-throughput sequencing technologies has enabled researchers to extensively acquire molecular data, revealing the deeper characteristics of genomes and transcriptomes. In this technological revolution, nanopore sequencing has elevated sequencing capabilities to a new level. Developed by Oxford Nanopore Technologies Inc., the nanopore sequencer employs protein nanopores to directly read single-molecule sequences without the need for DNA synthesis or amplification, commonly required in traditional sequencing processes. This technology supports sequencing of ultra-long nucleotide sequences, overcoming the read-length limitations of conventional techniques, and it can also identify DNA or RNA modifications. In recent years, this technology has achieved remarkable results in fields such as medicine, epidemiology, ecology, and education, gradually becoming a core tool in these domains. In this review, we thoroughly explore the unique advantages and functions of nanopore sequencers, present specific case studies of their application in various fields, and anticipate further breakthroughs and innovations they may bring in the future.

[Key words] long read length; nanopore sequencing; next generation sequencing

引言

纳米孔测序技术标志着基因组学研究的一个重要转折点。随着DNA测序技术的不断进步,测序通量和读长的提升大大扩展了我们对基因组的理解,推动了功能基因组学的深入研究。Oxford Nanopore Technologies (ONT)开发的纳米孔测序仪是一种无需DNA合成或扩增的单分子测序技术,这一特点使

其在测序领域独树一帜。与传统的测序方法相比,纳米孔测序具有显著的优势,包括超长读长、便携性、实时测序以及对DNA/RNA修饰的直接检测能力^[1]。

近年来,纳米孔测序的应用范围迅速扩大,从基因组学、转录组学到医学、生态学等多个领域。其便携式的MinION测序设备更是推动了基因组学研究从实验室走向野外及临床的实时应

用。本综述将重点介绍纳米孔测序技术的特点、其在功能基因组学中的应用现状以及未来的发展前景。

1 超长读长

纳米孔测序技术以其超长读长的能力而闻名,这使其在基因组组装和复杂区域解析中表现尤为突出。相比其他受限于短读长的测序技术,纳米孔测序可以生成超过2Mb的单个读长,大大简化了基因组从头组装的过程,尤其在长重复序列或结构变异的区域,这种超长读长的优势使其能跨越这些复杂区域,更准确地拼接出完整的基因组。此外,纳米孔技术无需DNA扩增或合成,避免了可能引入的偏差^[2]。

为了优化纳米孔测序的效果,ONT提供的文库制备方案建议可选的片段化步骤,使输入DNA平均长度在8kbp左右,从而获得最佳的接头连接效果并提高通量。然而,纳米孔测序并没有严格限制待测序DNA的长度,Loman等研究人员提出通过纳米孔技术实现“发现鲸鱼”的目标,即捕获超长DNA片段。这一过程对DNA提取和处理步骤提出了更高要求,传统的提取方法常会导致DNA断裂,而通过改进的苯酚-氯仿-异戊醇方法,研究人员成功获取了长度超过882kbp的读长,进一步展示了纳米孔测序的潜力。

随着纳米孔技术的便携性和实时性的提高,超长读长在基因组医学和临床测试中的应用潜力也日益凸显。结构变异的检测、人类白细胞抗原基因(Human Leukocyte Antigen genes, HLA)分型以及微卫星重复序列的识别都得到了显著改善。尤其是在癌症研究中,纳米孔测序已能检测大规模的结构变异,如胰腺癌中的肿瘤抑制基因失活^[3]。

2 直接RNA测序与修饰检测

纳米孔测序的一项关键优势在于其直接测序RNA分子的能力。传统的RNA测序方法通常需要通过逆转录将RNA转化为cDNA,再通过聚合酶链式反应(Polymerase Chain Reaction, PCR)扩增进行测序。这一过程可能会遗漏或错误识别RNA中的碱基修饰。而纳米孔测序则能够直接测量通过孔的RNA分子,从而准确地检测RNA的修饰状态。RNA修饰,如N6-甲基腺苷(m6A)和5-甲基胞嘧啶(5-mC),在基因表达调控和RNA代谢中起到重要作用。传统测序技术无法直接观察这些修饰,但纳米孔测序通过读取核糖核酸的电流信号变化,可以直接识别这些修饰,并为RNA修饰的功能研究提供了新的可能性^[4]。

不仅如此,纳米孔测序通过其超长的读长,可以在不需要拼接步骤的情况下完整测序整个转录本。这对于研究复杂的剪接异构体尤为重要。纳米孔测序能够揭示多种外显子和内含子的剪接形式,帮助研究人员深入了解基因表达的复杂调控机制^[5]。

3 实时测序与便携性

纳米孔测序设备的便携性和实时测序能力使其在多种应用场景中展现了独特优势。特别是MinION设备,其重量仅为90克,大小如同手机,可以通过USB连接至笔记本电脑供电。这一便携特性使MinION成为现场测序的理想工具,广泛应用于野外研究、临床诊断、应急监测等场景^[6]。例如,在埃博拉病毒和寨卡病毒爆发期间,研究人员利用MinION实现了病毒的现场基因组测序,

缩短了从样本采集到结果报告的时间,为公共卫生部门迅速应对疫情提供了关键支持^[7]。

MinION的便携性不仅仅体现在野外研究的便利性,它还能在极端环境中运行,例如南极洲和国际空间站。为了进一步提升其现场应用能力,科学家开发了与MinION兼容的移动实验室设备,如BentoLab,一体化DNA实验室,以及VoITRAX V2等自动化文库制备设备。这些设备简化了DNA提取和文库制备的流程,甚至可以通过替代传统PCR的LAMP方法(环介导等温扩增)进行DNA扩增,从而消除了对热循环仪的依赖。LAMP只需简单的水浴即可进行扩增,使得现场的基因组分析更加简便高效^[8]。

纳米孔测序技术的实时性同样具有重要意义。在测序过程中,系统可以即时读取DNA或RNA分子并进行碱基调用,使研究人员能够实时获取数据并快速分析。相比传统的测序技术,这种实时性大大加快了数据获取的速度,尤其在临床诊断和应急响应中极为重要。例如,MetaPORE是一种基于网络的实时分析管道,研究人员利用其从人体血液样本中同时检测基孔肯雅病毒、埃博拉病毒和丙型肝炎病毒,大幅缩短了诊断时间。同样,Mitsubishi的团队在样本采集后仅用两小时便完成了细菌成分的实时识别。

结合MinION设备的便携性和实时测序能力,它在资源有限的地区病毒爆发监测中起到了重要作用。2016年,世界卫生组织(World Health Organization, WHO)宣布寨卡病毒爆发为国际公共卫生紧急事件,随即启动了巴西寨卡实时分析(Zika Brazilian Real-Time Analysis, ZiBRA)项目。该项目利用MinION测序仪,对1000个寨卡病毒基因组进行了测序,实现了快速获取完整基因组并提供了流行病学信息。在西非埃博拉疫情中,纳米孔测序设备同样展现了其强大的应急响应能力,研究人员通过现场测序系统在短短15-60分钟内获得了病毒基因组序列,为疫情监测和控制提供了及时的数据支持。

4 长读长组装策略

纳米孔测序的超长读长能力极大简化了基因组的从头组装过程。与短读长测序依赖de Bruijn图算法不同,纳米孔测序更适合使用重叠布局一致性(Overlap-Layout-Consensus, OLC)算法来进行组装。由于OLC算法能够灵活处理长读长并对错误具有更强的鲁棒性,纳米孔测序可以更好地解决复杂基因组的组装难题,尤其是在含有长重复序列或结构变异的区域^[9]。

此外,为了提高组装的精度,研究人员通常采用混合组装策略,将纳米孔的长读长与Illumina短读长结合,利用短读长的高准确性来纠正长读长中的错误,从而生成更为完整、准确的基因组序列。随着生物信息学工具的不断优化,纳米孔测序在基因组组装中的应用将更加广泛^[10]。

5 使用案例: 功能基因组学中的应用

纳米孔测序技术在功能基因组学中的应用取得了显著进展,特别是在研究复杂基因组结构和解析基因功能方面。例如,在蜘蛛丝蛋白基因的研究中,研究人员利用纳米孔测序成功解析了其复杂的基因结构。蜘蛛丝蛋白基因具有高度重复的序列,这对

传统测序技术构成了巨大的挑战,而纳米孔测序凭借其超长读长能力,帮助研究人员从头组装了蜘蛛丝蛋白的全长序列,为揭示其独特的物理特性和潜在的工业应用打下了基础^[1]。

蜘蛛丝蛋白作为一种多功能蛋白质材料,具有非凡的物理特性,如高韧性、高拉伸强度和热稳定性,因而在工业领域中具有广泛的应用潜力。蜘蛛基因组的复杂性和规模,加之蜘蛛丝蛋白基因长度通常超过10kbp,且几乎完全由重复序列组成,使得解析这些基因变得极具挑战性。传统的短读长测序方法常常难以完整拼接基因序列,特别是在涉及重复单元的情况下。而纳米孔测序能够通过长读长直接读取这些复杂区域,并提供基因的完整结构信息。

在具体研究中,研究人员通过纳米孔测序对球织蜘蛛的七个蜘蛛丝蛋白基因家族进行了深入研究。小壶腹蛛丝蛋白基因通过低覆盖率的纳米孔DNA测序获得了完整的序列,结果表明其基因长度超过8.3kbp,比之前的短读长测序结果(约5,000bp)大幅延伸。此外,腺泡状蜘蛛蛋白基因的结构也通过纳米孔测序得到了详细解析,不仅揭示了其基因长度,还准确解析了外显子和内含子的结构。

6 总结与展望

纳米孔测序技术的快速发展为功能基因组学研究提供了强大的工具。其超长读长、便携性、实时测序以及直接检测RNA修饰的能力,使其在基因组学、转录组学、医学、生态学等多个领域展现了广泛的应用前景。尽管当前纳米孔测序仍存在一些技术瓶颈,如相对较高的错误率,但随着孔化学和碱基识别算法的不断优化,纳米孔测序的准确性和效率将进一步提高。

未来,纳米孔测序有望在基因组医学、结构变异分析、生态环境监测等方面发挥更加重要的作用。特别是在临床应用中,纳米孔测序的便携性和实时性将成为推动个性化医学发展的重要技术支持。通过进一步优化测序技术和生物信息学工具,纳米孔测序将在未来的功能基因组学研究中发挥更大的潜力。

[课题]

PIPK1 α 调节SWI/SNF染色质重塑酶构象促进三阴性乳腺癌转移的分子机制研究(21ZR1451000)。

[参考文献]

[1] Dorey A, Howorka S. Nanopore DNA sequencing technologies and their applications towards single-molecule proteomics [J]. Nature Chemistry, 2024, 16(3): 314–34.

[2] Rand A C, Jain M, Eizenga J M, et al. Mapping DNA methylation with high-throughput nanopore sequencing [J]. Nature methods, 2017, 14(4): 411.

[3] Rang F J, Kloosterman W P, De Ridder J. From squiggle to basepair: computational approaches for improving nanopore sequencing read accuracy [J]. Genome biology, 2018, 19(1): 90.

[4] Patel A, Belykh E, Miller E J, et al. MinION rapid sequencing: Review of potential applications in neurosurgery [J]. Surgical neurology international, 2018, 9.

[5] Hoang P N T, Michael T P, Sarah G, et al. Generating a high-confidence reference genome map of the Greater Duckweed by integration of cytogenomic, optical mapping, and Oxford Nanopore technologies [J]. Plant Journal, 2018.

[6] Gautier M, Yamaguchi J, Foucaud J, et al. The Genomic Basis of Color Pattern Polymorphism in the Harlequin Ladybird [J]. Current biology: CB, 2018, 28(20): 3296–302.e7.

[7] Tan M H, Austin C M, Hammer M P, et al. Finding Nemo: hybrid assembly with Oxford Nanopore and Illumina reads greatly improves the clownfish (*Amphiprion ocellaris*) genome assembly [J]. GigaScience, 2018, 7(3): 1–6.

[8] Ameer A, Kloosterman W P, Hestand M S. Single-Molecule Sequencing: Towards Clinical Applications [J]. Trends in biotechnology, 2019, 37(1): 72–85.

[9] Kolmogorov M, Billingsley K J, Mastoras M, et al. Scalable Nanopore sequencing of human genomes provides a comprehensive view of haplotype-resolved variation and methylation [J]. Nature methods, 2023, 20(10): 1483–92.

[10] Mitsuhashi S, Kryukov K, Nakagawa S, et al. A portable system for rapid bacterial composition analysis using a nanopore-based sequencer and laptop computer [J]. Scientific reports, 2017, 7(1): 5657.

[11] Ashton P M, Nair S, Dallman T, et al. MinION nanopore sequencing identifies the position and structure of a bacterial antibiotic resistance island [J]. Nat Biotechnol, 2015, 33(3): 296–300.

作者简介:

顾章媛(1995--),女,江苏扬中人,博士研究生,住院医师,研究方向:基于纳米孔测序的DNA/RNA修饰检测在肿瘤标志物筛查中的应用研究。

*通讯作者:

莊志刚(1971--),男,上海人,主任医师,博士生导师,研究方向:纳米孔测序技术在遗传病基因诊断中的应用及优化策略研究。