

高通量基因测序技术在新型冠状病毒检测中的应用

李梦婷¹, 祝小荐², 万绍贵³

(赣南医学院 1. 2019 级硕士研究生; 2. 2018 级硕士研究生; 3. 基础医学院分子病理学中心, 江西 赣州 341000)

摘要: 2019 年 12 月份起, 不明原因的肺炎患者造成中国武汉疫情大暴发。通过对肺炎患者样本分析, 发现了一种新型冠状病毒, 目前被世界卫生组织命名为 2019-nCoV。高通量测序技术的发展促进了对微生物病原的快速分析能力, 当前暴发的新型冠状病毒的首次发现便是基于宏基因组高通量二代测序技术所实现的。本文就二代测序(NGS)和第三代纳米孔测序等高通量基因测序技术在新型冠状病毒中的应用进行简要阐述。

关键词: 高通量测序; 新型冠状病毒; 宏基因组学; 二代测序; 纳米孔测序

中图分类号: R392.11 **文献标志码:** A **文章编号:** 1001-5779(2020)03-

DOI: 10.3969/j.issn.1001-5779.2020.03.000

Application of high-throughput gene sequencing technology in detection of new coronavirus

LI Meng-ting¹, ZHU Xiao-jian², WAN Shao-gui³

(Gannan Medical University 1. Postgraduate student, Grade 2019; 2. Postgraduate student, Grade 2018; 3. Centre of Molecular Pathology, School of Basic Medicine, Ganzhou, Jiangxi 341000)

Abstract: From December 2019, unexplained pneumonia patients had caused an outbreak in Wuhan, China. Analysis of samples from patients with pneumonia revealed a new coronavirus named 2019-nCoV by World Health Organization. The development of high-throughput gene sequencing technology has improved the ability to analyze pathogens, especially in infectious diseases. At present, the detection of new coronaviruses from the pneumonia patients was performed by metagenome next generation sequencing technology. In this review article, it will be described about the applications of next generation sequencing and third generation sequencing based on nanopore platform, in the discovery of 2019-nCoV pathogens recently.

Key words: high-throughput sequencing; 2019-nCoV; metagenomics; next generation sequencing; Nanapore sequencing

冠状病毒是一种可以引起多种动物多系统感染的病毒, 在人类中主要导致呼吸系统相关疾病, 例如严重急性呼吸综合征(SARS)、中东呼吸综合征(MERS)^[1-2]。2019 年底武汉暴发的以发热、咳嗽等始发症状的病毒性肺炎被鉴定为一种新型冠状病毒引起, 该病毒属 B 系 β -冠状病毒, 还包括 2003 年在中国引起非典大流行的严重急性呼吸综合征冠状病毒(SARS-CoV)和蝙蝠 SARS 样冠状病毒^[3]。新型冠状病毒被世界卫生组织命名为 2019 新型冠状病毒(2019-nCoV)^[4-5], 国际病毒分类委员会将新型冠状病毒命名为 SARS-CoV-2, 存在人传人和人类严重感染情况。该病毒之所以被命名为冠状病毒是因为在电子显微镜下, 病毒边缘有一个类似冠状病

毒的突起, 且通过高通量测序手段发现该病原体基因序列和 SARS-like bat coronavirus 有着密切的进化关系, 与 SARS 相关冠状病毒密切相关^[6-8], 2019-nCoV 与 SARS 相似性约 79%^[9], 最终鉴定为 SARSr-CoV(SARS 相关冠状病毒)。新型冠状病毒的检测方法多样, 通过临床样本和冠状病毒细胞培养上清可用于初步分析评估^[7]; 实时逆转录聚合酶链式反应(RT-PCR)可用于检测病毒特有 RNA^[7,10-11]; 透射电子显微镜可观察病毒形态^[10]。此外, 高通量测序也在病毒检测中扮演了至关重要的角色。高通量测序平台可以获得病毒全基因组信息, 发现病原体^[12], 有利于检测新型冠状病毒, 同时能迅速发现变异增加人类对其的认识度。高通量测

基金项目: 江西省科技创新杰出青年人才培养计划(20192BCBL23017); 赣南医学院 COVID-19 应急项目(YJ202022)

通信作者: 万绍贵, 男, 博士, 教授, 研究方向: 生物标志物与分子诊断。E-mail: wansg@gmu.edu.cn

投稿网址: <http://gnyxyxb.gmu.cn>

序技术对疫苗和新疗法的开发、病毒进化的研究、疾病遗传、跟踪暴发具有重要作用,近期也有报道高通量测序应用于监测医疗环境中病毒的暴发^[13-15]。本篇综述针对不同高通量测序平台对新型冠状病毒的检测应用和其特点进行简要阐述。

1 基于二代基因测序的宏基因组测序(mNGS)

NGS是目前实验室进行生物学及临床基因组学研究的有力工具^[16],应用十分广泛。宏基因组二代测序(mNGS)是指对微生物群体进行高通量测序,无须分离培养微生物,且超高深度的mNGS具有较高的病原体鉴定灵敏度,同时在传染病的发现和监测中极具潜力^[17]。基于Illumina平台的mNGS技术对冠状病毒的测序分析有助于了解新型冠状病毒的变异情况、有更加深刻的认识病毒基因序列,同时可以提高对新型冠状病毒的检出复核效能和病毒传播的追踪分析。

任丽丽等利用Illumina NGS的测序平台,对取自2019年12月18日武汉市的5例肺炎患者肺泡灌洗液(BAL)临床标本进行实验研究。对实验结果分析发现,5例临床样本中的一例样本,约97%的病毒阅读数(reads)归类于冠状病毒科,其余4例临床样本的病毒阅读数也被归类于 β -CoV,且未发现其他病原体感染(如细菌病原体)。通过Sanger测序确定样本病原体基因组序列,发现5例临床样本病原体全基因组序列的核苷酸相似度为99.8%~99.9%,基因5'-ORF1ab-S-E-M-N-3'与著名的bat SARS-like(SL)-CoV相似。对病毒全长基因组测序结果进行同源性分析,发现新型冠状病毒与bat SARS-like-CoVs最为相似,但在不同CoVs中高度保守的RNA依赖的RNA聚合酶序列只显示86.3%~86.5% nt同源性,提示5例临床样本中的病原体并非是bat SARS-like-CoVs感染,确认为新发现的病原体,即2019-nCoV^[18]。

同时,CHEN等也使用Illumina Miseq平台对2例临床肺炎样本进行研究。测序结果显示,2例样本的reads数与冠状病毒相似度分别为99.9%和99.7%。虽然2例样本存在一些单核苷酸多样性,但是从2例样本获得的一致性的基因序列是相同的,提示2名患者是由同一种冠状病毒感染的。对病毒全基因组进行分析比较发现,发现样本中的冠状病毒株与BtCoV/4991株的核苷酸同源性为98.7%,与蝙蝠冠状病毒bat-SL-CoVZC45株和bat-SL-CoVZXC21株的核苷酸同源性为87.9%,与

目前人类已知的冠状病毒包括SARS冠状病毒的核苷酸同源性为79.7%。CHEN等进一步分析发现,该2名肺炎患者的病毒株与其他bat-SL-CoVs在不同基因上的系统发育不一致,存在重组情况,且未发现除冠状病毒外其他病原体感染。提示这2名患者为同一种新型冠状病毒感染^[19]。

mNGS可以直接从原始临床样本中调查感染微生物^[20],2019-nCoV已被确认为RNA病毒,但在未知病原体是何类型病毒时,基于RNA的mNGS方法可以揭示生物体内的整个感染情况,包括DNA病毒、RNA病毒、细菌、真菌等^[21],Illumina的错误率为0.1%^[22],故针对此次新型冠状病毒的暴发,使用基于RNA的mNGS方法可以快速准确鉴定临床病例中潜在病原体和鉴定临床发热病例是否为2019-nCoV感染,新型冠状病毒基因组分析和鉴定在5天内可以完成^[23],明确病毒的起源和进化,对疾病的控制、预防和针对性治疗有重大意义。

基于Thermo Fisher公司Ion Torrent平台的NGS的基因组测序,可以直接利用临床样本提取病毒RNA,通过RT-PCR构建cDNA文库,之后进行测序、结果比对分析。靶向扩增获得全长信息RNA文库中包含了病毒信息和大量的背景信息,不仅浪费数据量,还干扰数据分析。而进行靶向扩增病毒RNA可以完美的解决上述问题。将分离提纯的RNA进行逆转录,构建cDNA文库,通过增加特异性识别靶向序列的启动子进行特异性扩增,去除人为添加序列,连接barcode和adapter后进行质量控制,完成文库构建,完成后进行测序。某省疾控中心使用该平台成功获得新型冠状病毒序列,与BetaCoV-Wuhan-IVDC-HB-01-2019片段一致性为99%,体现了实验结果的可靠性,增加结果可信度^[24]。

2 第三代纳米孔高通量测序

以牛津纳米孔公司为例,牛津纳米孔公司的设备可以对DNA和RNA进行测序并提供快速分析,纳米孔测序技术直接对核酸片段进行全长读取,从而进行长读长序列的高通量基因测序^[12],同时具有极高的精确度,高样本检测效率。该技术已被应用于许多科学方面的研究,包括人类遗传学、癌症研究和环境应用等,还被应用于多种疾病暴发情况,目前已有多个团队采用纳米孔基因测序技术对新型冠状病毒进行高通量基因组测序检测。

JASPER FUK-WOO CHAN等使用纳米孔测序

技术对新型冠状病毒进行研究。样本来源于确诊的 2 名患者及其他 5 名家族成员,采用全基因组测序方法。样本中提取核酸,去除宿主 DNA 后,应用不依赖序列的单引物扩增方法,用特异性引物 A(5'-GTTTCCCACTGGAGGATA-N9-3')将 DNA 酶处理的 RNA 进行逆转录为 cDNA,利用 Klenow 片段(3'→5' exo-)进行 cDNA 第二条链的合成(新英格兰生物实验室,马萨诸塞州伊普斯维奇)。用引物 B(5'-GTTTCCCACTGGAGGATA-3')进行 PCR 扩增,获得扩增的 cDNA 文库。进行纳米孔测序文库构建,文库准备完成即可进行测序,采用 R9.4.1 进行 12~48 小时测序。利用 MEGA X 软件构建进化树,进行数据分析。他们发现了一种与蝙蝠 SARS 样冠状病毒 bat-SL-CoVZXC21(NCBI 登录号 MG772934)和 bat-SL-CoVZC45(NCBI 登录号 MG772933)关系最为密切的新型冠状病毒,且其中 2 名患者的病毒株基因组序列完全一致,提示为同一病毒株感染。且两名患者家属也存在同一病毒株感染情况,故提示该病毒具有传染性。实验发现,痰样本来源的循环阈值比咽拭子样本来源早 8~13 个循环,表明下呼吸道病毒载量较高,提示临床初筛阴性患者可以重复检测上呼吸道样本或下呼吸道样本^[25]。

为了更好地对新型冠状病毒进行测序研究,纳米孔公司发布了一个关于 2019-nCoV 测序的实验流程^[26],利用针对新型冠状病毒特异性的 98 对引物对患者核酸样本的 cDNA 进行 PCR 扩增,靶向富集病毒基因组序列片段,进行测序文库构建和高通量测序分析^[26]。

纳米孔具有的长读长优势,对未知新型肺炎病原体检测、鉴定具有重要意义。利用纳米孔技术,在建库时不同样本添加不同的 barcode 以区分样本信息,对于不同样品之间的异同点,对检测新型肺炎家族聚集性或人传人现象具有重大意义^[25]。不同样本来源提示病毒载量的多少,有利于临床初筛阴性样本针对病毒载量高的部位重新取样、检测,降低漏诊率,对疫情控制具有重要作用。

3 展 望

针对新型冠状病毒肺炎的疫情暴发,目前临床上病原检测存在假阴性、病毒分离鉴定时间较长,病毒变异的潜在性等问题,而高通量测序技术很好地解决了这些问题。当前高通量测序技术飞速发展,具有极高的准确率,提高了对病原体检测的准确率,对于新型冠状病毒,高通量测序技术能更好地鉴定

和分析其来源、变异情况,减少临床检测的盲目性,且检出率较高。以纳米孔测序技术为例,牛津纳米孔公司与纳米孔社区科学家开发实时快速的病毒基因组测序方案来有效的监测和实时掌握病毒变异动态。杭州市疾控中心使用纳米孔基因测序平台对新型冠状病毒基因组进行高通量测序和组装分析,发现其与参考基因组一致性为 100%。在纳米孔社区开通的新型冠状病毒的检测方案流程,可在 8 小时内完成从样本到结果的获取。进一步体现了高通量测序在新型冠状病毒应用中的优势。疫情尚未结束,对新型冠状病毒的实时监测尤为重要,高通量测序在了解病毒来源、了解病毒变异情况具有积极作用。高通量测序的快速检测对于新型冠状病毒的控制、了解新型冠状病毒传播情况进而采取有效措施极其重要。故高通量测序对提高处理重大突发事件的应急能力,对保障人类生命安全有积极意义。

参考文献:

- [1] YIN Y, WUNDERINK R G. MERS, SARS and other coronaviruses as causes of pneumonia[J]. *Respirology*, 2018,23(2):130-137.
- [2] GUO L, WEI D, ZHANG X, et al. Clinical features predicting mortality risk in patients with viral pneumonia: The MuLBSTA Score[J]. *Front Microbiol*, 2019,10:2752.
- [3] JIANG S, DU L, SHI Z. An emerging coronavirus causing pneumonia outbreak in Wuhan, China: calling for developing therapeutic and prophylactic strategies [J]. *Emerg Microbes Infect*, 2020,9(1):275-277.
- [4] ZHANG N, WANG L, DENG X, et al. Recent advances in the detection of respiratory virus infection in humans [J]. *J Med Virol*, 2020,92(4):408-417.
- [5] HUANG C, WANG Y, LI X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China[J]. *Lancet*, 2020,395(10223):497-506.
- [6] PEIRIS J S, YUEN K Y, OSTERHAUS A D, et al. The severe acute respiratory syndrome [J]. *N Engl J Med*, 2003,349(25):2431-2441.
- [7] CORMAN V M, LANDT O, KAISER M, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR[J]. *Euro Surveill*, 2020,25(3):2000045.
- [8] ZHOU P, YANG X L, WANG X G, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin[J]. *Nature*, 2020,579:270-273.
- [9] LU R J, ZHAO X, LI J, Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for

- virus origins and receptor binding[J]. *Lancet*, 2020, 395: 10224.
- [10] ZHU N, ZHANG D, WANG W, et al. A novel coronavirus from patients with pneumonia in China, 2019[J]. *N Engl J Med*, 2020, 382(8):727–733.
- [11] HUANG P, LIU T, HUANG L, et al. Use of Chest CT in combination with negative RT-PCR assay for the 2019 novel coronavirus but high clinical suspicion[J]. *Radiology*, 2020, 295(1):22–23.
- [12] SAKAMOTO Y, SEREEWATTANAWOOT S, SUZUKI A. A new era of long-read sequencing for cancer genomics[J]. *J Hum Genet*, 2020, 65(1):3–10.
- [13] GARVEY M I, BRADLEY C W, HOLDEN K L, et al. Use of genome sequencing to identify hepatitis C virus transmission in a renal healthcare setting[J]. *J Hosp Infect*, 2017, 96(2):157–162.
- [14] HOULIHAN C F, FRAMPTON D, FERNS R B, et al. Use of whole-genome sequencing in the investigation of a nosocomial influenza virus outbreak[J]. *J Infect Dis*, 2018, 218(9):1485–1489.
- [15] VAUGHAN G, FORBI J C, XIA G L, et al. Full-length genome characterization and genetic relatedness analysis of hepatitis A virus outbreak strains associated with acute liver failure among children[J]. *J Med Virol*, 2014, 86(2):202–208.
- [16] VAN DIJK E L, JASZCZYSZYN Y, NAQUIN D, et al. The third revolution in sequencing technology[J]. *Trends Genet*, 2018, 34(9):666–681.
- [17] MIAO Q, MA Y, WANG Q, et al. Microbiological diagnostic performance of metagenomic next-generation sequencing when applied to clinical practice[J]. *Clin Infect Dis*, 2018, 67(suppl_2):S231–S240.
- [18] REN L L, WANG Y M, WU Z Q, et al. Identification of a novel coronavirus causing severe pneumonia in human: a descriptive study[J]. *Chin Med J (Engl)*, 2020, 10. 1097/CM9. 0000000000000722. doi: 10. 1097/CM9. 0000000000000722.
- [19] CHEN L, LIU W, ZHANG Q, et al. RNA based mNGS approach identifies a novel human coronavirus from two individual pneumonia cases in 2019 Wuhan outbreak[J]. *Emerg Microbes Infect*, 2020, 9(1):313–319.
- [20] GU W, MILLER S, CHIU CY. Clinical metagenomic next-generation sequencing for pathogen detection[J]. *Annu Rev Pathol*, 2019, 14:319–338.
- [21] SHI M, ZHANG Y Z, HOLMES E C. Meta-transcriptomics and the evolutionary biology of RNA viruses[J]. *Virus Res*, 2018, 243:83–90.
- [22] GOODWIN S, MCPHERSON J D, MCCOMBIE W R. Coming of age: ten years of next-generation sequencing technologies[J]. *Nat Rev Genet*, 2016, 17(6):333–351.
- [23] 陶悦, 傅启华, 莫茜. 病原宏基因组测序在新型冠状病毒检测中的应用与挑战[J]. *中华检验医学杂志*, 2020, 43(2020–02–16). [2020–03–13]. <http://rs.yiigle.com/yufabiao/1181377.htm>. doi:10.3760/cma.j.issn.1009–9158.2020.0008.
- [24] 赛默飞 Ion Torrent 高通量测序. 新冠疫情面前, NGS 如何助力[EB/OL]. [2020–03–13]. https://mp.weixin.qq.com/s/O7IlhQO1770rKLIQv_k75w.
- [25] CHAN J F, YUAN S, KOK K H, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster[J]. *Lancet*, 2020, 395(10223):514–523.
- [26] JOSH QUICK. nCoV-2019 sequencing protocol[EB/OL]. [2020–03–13]. <https://www.protocols.io/view/ncov-2019-sequencing-protocol-bbmui6w>.
- (收稿日期:2020–03–13)(责任编辑:敖慧斌)