Original Article

Rapid and Accurate Sequencing of Enterovirus Genomes Using MinION Nanopore Sequencer*

WANG Ji[1,&], KE Yue Hua[2,&], ZHANG Yong[1,&], HUANG Ke Qiang[1], WANG Lei[3], SHEN Xin Xin[1], DONG Xiao Ping[1], XU Wen Bo[1,#], and MA Xue Jun[1,#]

1. National Institute for Viral Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing 102206, China; 2. Institute of Disease Control and Prevention, Beijing 100071, China; 3. Beijing Applied Biological Technologies Co., Ltd. Beijing 102206, China

Abstract

Objective Knowledge of an enterovirus genome sequence is very important in epidemiological investigation to identify transmission patterns and ascertain the extent of an outbreak. The MinION sequencer is increasingly used to sequence various viral pathogens in many clinical situations because of its long reads, portability, real-time accessibility of sequenced data, and very low initial costs. However, information is lacking on MinION sequencing of enterovirus genomes.

Methods In this proof-of-concept study using Enterovirus 71 (EV71) and Coxsackievirus A16 (CA16) strains as examples, we established an amplicon-based whole genome sequencing method using MinION. We explored the accuracy, minimum sequencing time, discrimination and high-throughput sequencing ability of MinION, and compared its performance with Sanger sequencing.

Results Within the first minute (min) of sequencing, the accuracy of MinION was 98.5% for the single EV71 strain and 94.12%-97.33% for 10 genetically-related CA16 strains. In as little as 14 min, 99% identity was reached for the single EV71 strain, and in 17 min (on average), 99% identity was achieved for 10 CA16 strains in a single run.

Conclusion MinION is suitable for whole genome sequencing of enteroviruses with sufficient accuracy and fine discrimination and has the potential as a fast, reliable and convenient method for routine use.

Key words: Nanopore sequencing; MinION; Enterovirus; Single molecule sequencing; Viral genome sequencing

Biomed Environ Sci, 2017; 30(10): 718-726 *doi: 10.3967/bes2017.097* *ISSN: 0895-3988*

www.besjournal.com (full text) *CN: 11-2816/Q*

INTRODUCTION

Enteroviruses (EVs) are associated with a great variety of manifestations, varying from mild respiratory and gastrointestinal infections, herpangina, and hand-foot-and-mouth disease (HFMD), to more severe diseases leading to mortality[1]. Enterovirus 71 (EV71) and Coxsackievirus A16 (CA16) are the most common EVs causing HFMD[2]. Knowledge of viral genome sequences is very important in epidemiologic investigations, to determine whether specific cases

*This work was supported by the National key research and development plan (2016TFC1202700, 2016YFC1200900); Beijing Municipal Science & Technology Commission project (grant numbers D151100002115003); and Guangzhou Municipal Science & Technology Commission project (grant numbers 2015B2150820).

&WANG Ji, KE Yue Hua, and ZHANG Yong contributed equally to the work and should be regarded as co-first authors.

#Correspondence should be addressed to XU Wen Bo, E-mail: wenbo_xu1@aliyun.com; MA Xue Jun, E-mail: maxj@ivdc.chinacdc.cn

Biographical notes of the first authors: WANG Ji, male, born in 1987, MM, majoring in pathogen biology; KE Yue Hua, male, born in 1982, PhD, majoring in microbial pathogenesis; ZHANG Yong, male, born in 1975, PhD&MD, majoring in medical virology.

are epidemiologically linked, to identify transmission patterns, and to ascertain the extent of an outbreak[3]. Considering the great variations among different serotypes/genotypes of EVs, whole genome sequencing is the most appropriate way to distinguish different serotypes/genotypes, to detect recent recombination events and to provide a high-resolution view of viral lineages[4].

DNA sequencing technology has made remarkable progress since Sanger sequencing was established in 1977[5]. Since 2005, 454 technology (Roche), Solexa technology (Illumina) and SOLiD technology (Applied Biosystems) have been developed successively as Second-Generation Sequencing (SGS) systems, leading to a new age of high-throughput sequencing[6]. Single molecule sequencing technology, referred to as Third-Generation Sequencing (TGS), represented by SMRT (Pacific Bioscience) technology, offers longer read lengths than the SGS technologies, making it suitable for full-length viral sequencing and many other biological/medical researches[7-8].

The first commercial nanopore sequencer is MinION, developed by Oxford Nanopore Technologies (ONT). It can be defined as a TGS platform considering its single molecule sequencing ability[9], but its technical principles and properties are very different compared with the previous platforms. Thus, this nanopore sequencer is also described as a Fourth-Generation Sequencing (FGS) platform in some publications[10-11]. MinION is a palm-sized device that drives individual DNA/RNA molecules through a nanopore; only a single strand nucleic acid can pass through the pore. Because the electrical properties of the bases A, T, G, and C are different, electrical signals with base specificity can be detected by MinION and sequence information can thus be collected continuously using the MinKNOW software.

MinION is capable of generating reads as long as 882 kb[12], which improves the scaffolding of prokaryotic and eukaryotic genomes and sequencing of bacterial and viral isolates[13-16]. During the Ebola outbreak in west Africa, a set of portable viral MinION-based sequencing systems was transported *via* standard airline luggage to Guinea[17-18]. In addition, MinION has been used to detect a variety of viruses using amplicons[16-17,19-20], hybridization capture[21-22] or unbiased metagenomics approaches[23-24]. Though the read accuracy of MinION is lower than that of short-read sequencers such as MiSeq or HiSeq, the deeper genome coverage depths can be used to generate an accurate consensus sequence, achieving > 99% accuracy post-data analysis[16].

MinION may provide new opportunities in EV diagnostics by rapid whole genome sequencing and determination of EV serotypes/genotypes, for example during an outbreak of HFMD. In this study, we established a whole genome sequencing method for EV71 and CA16 using MinION. We explored the accuracy, minimum sequencing time, discrimination and high-throughput sequencing ability of MinION, and compared the data with Sanger sequencing results.

MATERIAL AND METHODS

Viral Strains and Amplicon Preparation

One EV71 and 10 CA16 isolates used in this study were from swab or stool samples collected from HFMD patients < 5-year-old. The samples were inoculated onto rhabdomyosarcoma cells provided by the WHO Global Poliovirus Specialized Laboratory, USA, and originally purchased from the American Type Culture Collection (Manassas, VA, USA). Cell cultures were harvested after complete cytopathic effect was observed. The titers of CA16 and EV71 were $10^{5.5}$ $CCID_{50}$/mL and $10^{6.5}$ $CCID_{50}$/mL, respectively. All the CA16 strains belonged to the B1 genotype with nucleotide similarity ranges 90.7%-98.5%, and the EV71 strain belonged to the C4a evolutionary branch in the C4 genotype. This work was approved by the Ethics Review Committee of the National Institute for Viral Disease Control and Prevention and was supported by a long-term surveillance program for HFMD.

Viral RNA was extracted from the viral isolates using a QIAamp Viral RNA Mini Kit (Qiagen) and stored at -80 °C until further use. The EV71 and CA16 amplicons were amplified using the RT-PCR protocol established by the National Polio Laboratory of China. Briefly, the viral RNA was converted to cDNA by a random-priming strategy. The cDNA was added to 4-10 separate PCR tubes with a SuperScript One-Step RT-PCR System for Long Templates Kit (Invitrogen) using gene specific primers. The PCR program used was: 2 min at 94 °C, then 40 cycles of 15 s at 94 °C, 30 s at 56 °C, and 1 min at 68 °C, with a final extension at 68 °C for 7 min.

Library Preparation

The Ligation Sequencing Kit 1D R9 Version (ONT,