



Development and Application of a Fast Method to Acquire the Accurate Whole-Genome Sequences of Human Adenoviruses

Shan Zhao¹, Wenyi Guan¹, Kui Ma², Yuqian Yan¹, Junxian Ou¹, Jing Zhang², Zhiwu Yu³, Jianguo Wu² and Qiwei Zhang^{1,2*}

¹ Guangdong Provincial Key Laboratory of Tropical Disease Research, School of Public Health, Southern Medical University, Guangzhou, China, ² Guangdong Provincial Key Laboratory of Virology, Institute of Medical Microbiology, Jinan University, Guangzhou, China, ³ Division of Laboratory Science, Affiliated Cancer Hospital & Institute of Guangzhou Medical University, Guangzhou, China

OPEN ACCESS

Edited by:

Chunfu Zheng,
Fujian Medical University, China

Reviewed by:

Ya-Fang Mei,
Umeå University, Sweden
Xingui Tian,
First Affiliated Hospital of Guangzhou
Medical University, China

*Correspondence:

Qiwei Zhang
zhangqw@jnu.edu.cn
orcid.org/0000-0002-2770-111X

Specialty section:

This article was submitted to
Virology,
a section of the journal
Frontiers in Microbiology

Received: 30 January 2021

Accepted: 06 April 2021

Published: 14 May 2021

Citation:

Zhao S, Guan W, Ma K, Yan Y,
Ou J, Zhang J, Yu Z, Wu J and
Zhang Q (2021) Development
and Application of a Fast Method
to Acquire the Accurate
Whole-Genome Sequences of Human
Adenoviruses.

Front. Microbiol. 12:661382.
doi: 10.3389/fmicb.2021.661382

The whole-genome sequencing (WGS) of human adenoviruses (HAdVs) plays an important role in identifying, typing, and mutation analysis of HAdVs. Nowadays, three generations of sequencing have been developed. The accuracy of first-generation sequencing is up to 99.99%, whereas this technology relies on PCR and is time consuming; the next-generation sequencing (NGS) is expensive and not cost effective for determining a few special samples; and the third-generation sequencing technology has a higher error rate. In this study, first, we developed an efficient HAdV genomic DNA extraction method. Using the complete genomic DNA instead of the PCR amplicons as the direct sequencing template and a set of walking primers, we developed the HAdV WGS method based on first-generation sequencing. The HAdV whole genomes were effectively sequenced by a set of one-way sequencing primers designed, which reduced the sequencing time and cost. More importantly, high sequence accuracy is guaranteed. Four HAdV strains (GZ01, GZ02, HK35, and HK91) were isolated from children with acute respiratory diseases (ARDs), and the complete genomes were sequenced using this method. The accurate sequences of the whole inverted terminal repeats (ITRs) at both ends of the HAdV genomes were also acquired. The genome sequence of human adenovirus type 14 (HAdV-B14) strain GZ01 acquired by this method is identical to the sequence released in GenBank, which indicates that this novel sequencing method has high accuracy. The comparative genomic analysis identified that strain GZ02 isolated in September 2010 had the identical genomic sequence with the HAdV-B14 strain GZ01 (October 2010). Therefore, strain GZ02 is the first HAdV-B14 isolate emergent in China (September 2010; GenBank acc no. MW692349). The WGS of HAdV-C2 strain HK91 and HAdV-E4 strain HK35 isolated from children with acute respiratory disease in Hong Kong were also determined by this sequencing method. In conclusion, this WGS method is fast, accurate, and universal for common human adenovirus species B, C, and E. The sequencing strategy may also be applied to the WGS of the other DNA viruses.

Keywords: whole-genome sequencing, genomic DNA, human adenovirus, human adenovirus type 14, human adenovirus type 2, human adenovirus type 4, Sanger sequencing, inverted terminal repeats (ITRs)

INTRODUCTION

Human adenoviruses (HAdVs) have seven species, A (HAdV-A) through G (HAdV-G), defined by various biological and morphological criteria, nucleic acid characteristics, and homologies (Zhang et al., 2017; Cheng et al., 2018). One-third of HAdVs are related to human diseases and are estimated to cause 8% of global clinically relevant viral diseases (Yu et al., 2016). Common adenoviral diseases include respiratory infections in children and military recruits, infantile gastroenteritis, and ocular infections among healthy individuals (Zhang et al., 2009, 2012b; Han et al., 2013). Less frequently, these pathogens can cause urinary tract infections, myocarditis, meningoencephalitis, and acute hemorrhagic cystitis. Meanwhile, in neonates and immunocompromised individuals, HAdVs have been reported to cause fulminant fatal pneumonia, hepatitis, and/or encephalitis (Adhikary et al., 2014; Scott et al., 2016; Cheng et al., 2018). In 2019, a local outbreak of adenovirus in southwest China killed more than 40 children.

Human adenoviruses are usually classified by serological criteria and hemagglutination-inhibition tests, both of which are associated with the three major capsid proteins, hexon, fiber, and penton base (Zhang et al., 2009; Seto et al., 2010). But now, the classification of adenovirus is usually based on more modern criteria, including genome and phylogenetic analysis (Robinson et al., 2013), which uses the whole-genome sequence to characterize and name human adenoviruses (Human Adenovirus Working Group (HAWG)¹ (Seto et al., 2011). The commonly used sequencing methods include the Sanger sequencing method and the next- and third-generation sequencing technology. The traditional first-generation sequencing technology represented by Sanger sequencing reads up as long as 1,000 bp, and the accuracy is as high as 99.99% (Sanger et al., 1977; Lu et al., 2016; Shendure et al., 2017). However, the first-generation sequencing technology takes much time because it depends on polymerase chain reactions (PCRs) and/or DNA electrophoresis separation technology. The next-generation sequencing (NGS) technology is expensive and not suitable for the common laboratories in developing countries to determine large clinical samples' whole-genome sequences (Koboldt et al., 2013; Lu et al., 2016; Shendure et al., 2017). Currently, more than 95% of next-generation sequencing (NGS) applications require the reads aligning to a reference genome or reference transcriptome sequence, termed "mapping." However, the accuracy of mapping varies considerably, and mutations or recombinations are often missed or ignored (Zhang et al., 2021); the third-generation sequencing technology has a high error rate (Lu et al., 2016; Shendure et al., 2017; van Dijk et al., 2018). Recently, a new sequencing technology, nanopore sequencing, had been developed, which was famous for low cost and little sample preparation, but irregular steps caused substantial errors by the motor enzymes (Rusk, 2019).

The next-generation whole-genome sequencing is classified into *de novo* sequencing and resequencing. *De novo* sequencing does not require any reference genome information or custom

primers. Bioinformatics analysis methods are used to splice and assemble the genome. However, *de novo* sequencing is easy to produce gaps, requiring primer design and PCR to obtain the gap sequences. Additionally, there are more mismatches than one generation. Resequencing is the sequencing of the genomes with reference genomic sequences from the same or similar species. However, this strategy works badly for recombinant viruses. For example, internal recombination within human adenovirus species B is so common that the reference strain is not easy to select correctly (Zhang et al., 2016). As a result, the recombination or insertions/deletions (indels) might be mistaken for mismatches and ignored, failing to identify the recombination or indels in clinical isolates.

The lack of accuracy is still a major weakness of NGS when compared with first-generation sequencing. This study improved the sequencing method based on the first-generation sequencing and applied it in the clinical sample sequencing. First, we developed an effective adenovirus genomic DNA extraction method. Using the adenovirus genomic DNA as a direct sequencing template instead of the PCR amplicons, we did the Sanger sequencing using a set of walking primers covering the whole genome. This method reduces the mismatches and time without any PCR procedure and also brings down the costs. Second, we obtained the whole-genomic sequences of four HAdV isolates [strains GZ01 (HAdV-B14), GZ02 (HAdV-B14), HK91 (HAdV-C2), and HK35 (HAdV-E4)], with both complete inverted terminal repeat (ITR) ends sequenced successfully using this method. The genome sequence of strain GZ01 acquired by this method is identical to the sequence released before (GenBank no. JQ824845), which indicates that this novel sequencing method has high accuracy. We also found that strain GZ02 isolated in September 2010 had the identical genomic sequence with strain GZ01 isolated in October 2010 (Zhang et al., 2012b, 2017), which suggests that both isolates derived from the same strain; HAdV-B14 strains GZ02, not strain GZ01, is the first HAdV-B14 isolate emergent in China. Finally, we successfully applied this method for the whole-genome sequencing of two additional clinical isolates, HAdV-C2 (strain HK91) and HAdV-E4 (strain HK35), and the complete genomic sequences were acquired and annotated.

MATERIALS AND METHODS

Cells and Virus Isolates

HAdV-B14 strain GZ01 was isolated from a throat swab of a 17-month-old child hospitalized with acute suppurative tonsillitis in Guangzhou, China (October 2010) (Zhang et al., 2017). HAdV-B14 GZ02 was isolated from a throat swab of a child hospitalized with bronchopneumonia in another hospital in Guangzhou (September 2010). Strains HK35 and HK91 were isolated from throat swabs of children with ARD in Hong Kong (2014). The throat swabs were inoculated into A549 cells (ATCC), respectively, and grown in Dulbecco's minimum essential medium supplemented with 100 IU/ml penicillin, 100 µg/ml streptomycin, and 2% (v/v) fetal calf serum, at an atmosphere of 5% (v/v) carbon dioxide. The cells were observed for 10 days

¹<http://hadvwm.gmu.edu>

for cytopathic effect (CPE) and harvested for viral genomic DNA extraction. This study was approved by the Institutional Review Board of Southern Medical University following the Declaration of Helsinki, with the patient consent for using leftover specimens waived.

Extraction of Adenoviral Genomic DNA as the Direct WGS Template

Viral genomic DNA was extracted from infected cells for genome sequencing using the Hirt viral genomic DNA extraction method modified by us (Hirt, 1967). In brief, a sample of each stock virus was grown in the A549 cells in a 75-cm² tissue culture dish (Corning) at 37°C and an atmosphere of 5% CO₂. The cells were infected following standard procedures at an approximate multiplicity of infection (MOI) of 10 and observed for CPEs every day. First, when an 80% CPE or more occurred, the supernatant was discarded, and 800 µl of the lysate [0.6% sodium dodecyl sulfate (SDS), 0.01 M ethylenediaminetetraacetic acid (EDTA), 0.01 M Tris-HCl (pH 7.4)] and 20 µl protease K were added to the cells. The mixture was incubated at 55°C for 1 h. If the cells detached before adding the lysate, the cells and medium were transferred to another centrifuge tube, centrifuged for 2 min at 1,000 × g, 4°C. Then, the supernatant was discarded, and 2.4 ml of lysate and 60 µl protease K (TAKARA, 20 mg/ml stock solution) was added to the centrifuge tube, then gently mixed by a pipet. The mixture was transferred to the original dish and incubated at 55°C for 1 h. Next, the cell lysis was transferred to 1.5-ml centrifuge tubes (800 µl per tube). Three hundred microliters of 5 M NaCl was added to each tube (the volume of NaCl vs. lysis: 3:8) and mixed gently by inversion (≥10 times). The tubes were placed at 4°C overnight (≥12 h) and then centrifuged at 1,7000 g, 4°C for 20 min. The supernatant was collected into another fresh 1.5-ml centrifuge tube, and viral DNA was isolated from the supernatant solution by standard phenol/chloroform/isoamyl alcohol extraction and ethanol precipitation. The precipitated DNA was resuspended in 50 µl of TE buffer, and the genomic DNA was quantified by electrophoresis gel.

Prior to sequencing, we performed restriction endonuclease analysis (REA) to check and identify the genomic DNA rapidly. *EcoRI* was used, and an *in silico* REA map was generated first by the Vector NTI 11.5.1 software (Invitrogen Corp., San Diego, CA, United States) (Zhao et al., 2014; Yu et al., 2016; Zhang et al., 2017; Pan H. et al., 2018; Yan et al., 2020a). Then, the genomic DNA was digested by *EcoRI* according to the instruction.

HAdV Genome Sequencing Strategy

The isolates' genome sequences were obtained and used as the sequencing template by the Sanger primer-walking sequencing method without PCR amplification. To verify the method's accuracy, the HAdV-B14 strain GZ01 (JQ824245) stored in our laboratory was sequenced again, previously sequenced by traditional first-generation sequencing (Zhang et al., 2017). Using the strain GZ01 sequence (JQ824245) as a template, all one-way sequencing primers for HAdV-B14 sequencing were designed.

The distance between the two sequencing primers was 500–600 bp. For other unknown isolates, to determine the genomic reference sequence for primer design, molecular typing of HadVs has been performed first using the universal PCR primers revised according to our previous study (Han et al., 2013): HexonF (5'GCCCCARTGGGCRATACATGCACATC3') and HexonR (5'AGCACSCCSCGRATGTCAAAG3'). The amplification length was 300 bp, and the amplification conditions were as follows: 94°C for 1 min, 34 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 20 s with a final extension of 72°C for 7 min. The products were subsequently sequenced and analyzed. By blast, the reference strains (GenBank nos. JQ824845, KX384951, and JX173077) were selected as the primer-design templates for HAdV-B14, HAdV-E4, and HAdV-C2, respectively. The sequencing primers covering the whole genomes were designed and synthesized for all three types (Supplementary Table 1).

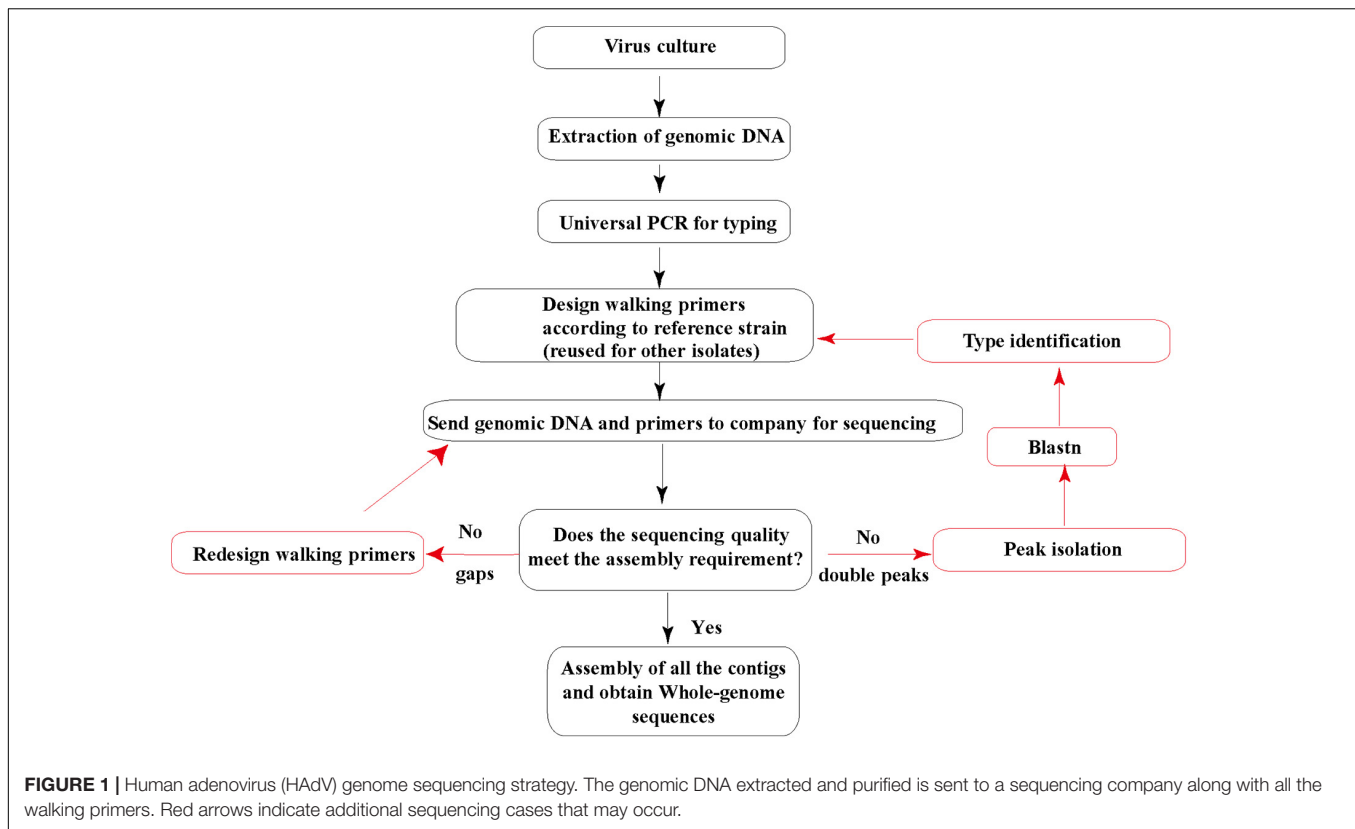
Then, the genomic DNA and sequencing primers were submitted to Invitrogen (Guangzhou) for Sanger sequencing. All the sequencing was repeated three times. DNA sequence fragments were assembled using the SeqMan software from the Lasergene package (DNA Star; Madison, WI) into a single contig for each strain. The default parameters of SeqMan were used during the assembly: trim sequence ends, removed contaminant sequences, optimized sequence assembly order, medium quality stringency, and sequence percent match higher than 90% in contigs. Each sequence from Sanger sequencing assembled into contigs should be more than 600 nucleotide acid long with clear single peaks. Most of the base QV should be higher than 40, and the signal/noise is higher than 30. Suppose the sequences from Sanger sequencing contain double peaks with similar and high signals. In that case, we will do the peak isolation and blast the two separated sequences and confirm if they are mixed adenovirus isolates, i.e., coinfection with two types of HadVs. If there are gaps among contigs after assembly, a couple of walking primers will be redesigned and sent for additional sequencing (Figure 1).

Direct Sequencing of the Inverted Terminal Repeat Ends

The sequencing primers for the 5' and 3' ends of the linear human adenovirus genome [inverted terminal repeat (ITR)] are designed according to the obtained contig. Forward and reverse sequencing primers were 200–300 bp away from both genome ends, respectively. Both 5' and 3' ends, including inverted terminal repeats, were sequenced directly using genomic DNA as a template by Sanger primer-walking sequencing method described earlier (Zeng et al., 2016; Cheng et al., 2018).

Genome Annotation and Sequence Analysis

The fragment sequences were collected and assembled with SeqMan software from the Lasergene package (DNASTar). Genome annotation provided an additional layer of sequence quality control. Unresolved and ambiguous sequences were resequenced with primers close to the regions in question.



Phylogenetic Analysis

Phylogenetic analysis of the whole genomes was performed using MEGA X², and maximum composite likelihood method was applied to generate neighbor-joining and bootstrapped phylogenetic trees with 1000 bootstrap replications; all other parameters were set by default (Jing et al., 2019; Zhang et al., 2019; Yan et al., 2020b).

RESULTS

Extraction of Adenoviral Genomic DNA

The whole genomic DNA of four isolates (HK35, HK91, GZ01, and GZ02) were extracted by the improved genomic DNA extraction method successfully. Viral genomic DNA (40–75 μg) was extracted from a 10-cm cell culture dish, and the OD₂₆₀/OD₂₈₀ and OD₂₆₀/OD₂₃₀ ratios were around 1.9 and 2.2, respectively. This indicated that the DNA extracted by this method was not contaminated by proteins, carbohydrates (sugars), salts, or organic solvents. The quality and quantity is high enough to meet our sequencing requirements. To verify if the viral genomic DNA obtained is correct or not, the genomes were digested with *Eco*RI restriction endonuclease. The restriction maps generated were consistent with the *in silico* restriction maps predicted by the Vector NTI 11.5.1 software (Figure 2).

²<http://www.megasoftware.net>

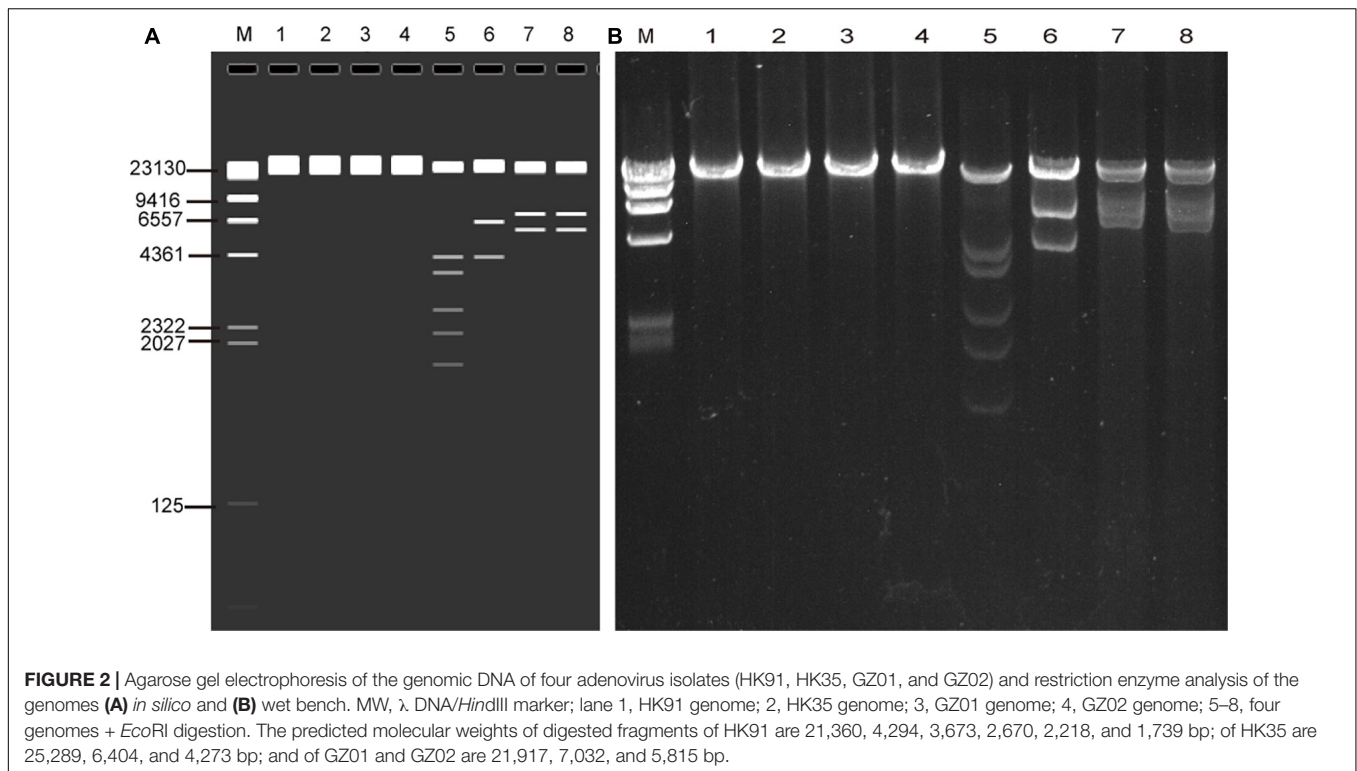
Sequencing and Analysis of Both Inverted Terminal Repeats

Compared with the standard PCR and Sanger sequencing method, it is hard to get the complete ITR sequences; our novel sequencing method can easily acquire the accurate adenoviral end sequences (Figure 3). HK91 ITR is 104 bp, both GZ01 and GZ02 ITRs are 137 bp, and HK35 ITR is 208 bp. All the ITR sequences are about the same length as the reference strains after the alignment (Figures 3A–C). However, we still found two nt deletions in the ITR sequence of HAdV-C2 reference strain ARG/A8649/2005. The sequencing error of the ARG/A8649/2005 strain was speculated, and it could be further confirmed using genomic DNA as a direct template for Sanger sequencing.

Comparative Genomic Analysis of the Four HAdV Isolates

The previous study has found that some adenovirus isolates had partial gene deletion (Marinheiro et al., 2011; Su et al., 2011; Pan Q. et al., 2018; Shen et al., 2020). For example, both Marinheiro et al. (2011) and Su et al. (2011) reported HAdV-7 variants in which E3 regions were partially deleted. Therefore, we analyzed the transcriptional maps and genome organizations of HAdV-C2 (HK91), HAdV-B14 isolates (GZ01/GZ02), and HAdV-E4 (HK35) (Figure 4). They could be used to visually identify if there were gene deletions or insertions in the genomes.

The map of GZ02 was the same as GZ01. A total of 38 coding sequences were identified (Figure 4A). These genome



data of strain GZ02 were deposited in GenBank (accession number MF062484) under the formal name preferred by the National Center for Biotechnology Information (NCBI) (Seto et al., 2011): “Human adenovirus 14 isolate HAdV-B/CHN/GZ02/2010/14[P14H14F14].” The genome comprises 34,767 bp, with a GC content of 48.83%, consistent with the other subspecies B2 (mean of 49%). These lengths are very similar to that of other HAdV-B14p1 strains 303600 (Lackland Air Force Base, United States; 2007) and CHN2012 (Beijing, China; 2012), as well as the prototype from 1955: genome sizes of 34,763, 34,760, and 34,764 bp, respectively. HAV-B14 strain GZ02 was isolated in September 2010, 1 month earlier than strain GZ01. Surprisingly, both isolates had identical genomic sequences, indicating that both isolates have the same origin. Given that both strains were isolated from two children from different hospitals in Guangzhou, the transmission chain of HAdV-B14 needs to be further explored.

The genomic sequence of HAdV-C2 strains HK91 was annotated and deposited into GenBank under accession number MF044052. Like other genera Mastadenovirus, the genome of strain HK91 was organized into early, intermediate, and late transcription regions (**Figure 4B**). The genome was 35,954 bp in length and had an overall base composition of 23.17% A, 27.95% C, 27.24% G, 21.61% T, and 55.2% GC. Thirty-nine protein-coding sequences and two RNA-coding sequences were identified in the genomic sequence.

The genomic sequence of HAdV-E4 strains HK35 was annotated and deposited into GenBank under accession number MW692349. The genome was 35,966 bp in length and had an overall base composition of 22.65% A, 28.35% C, 27.93%

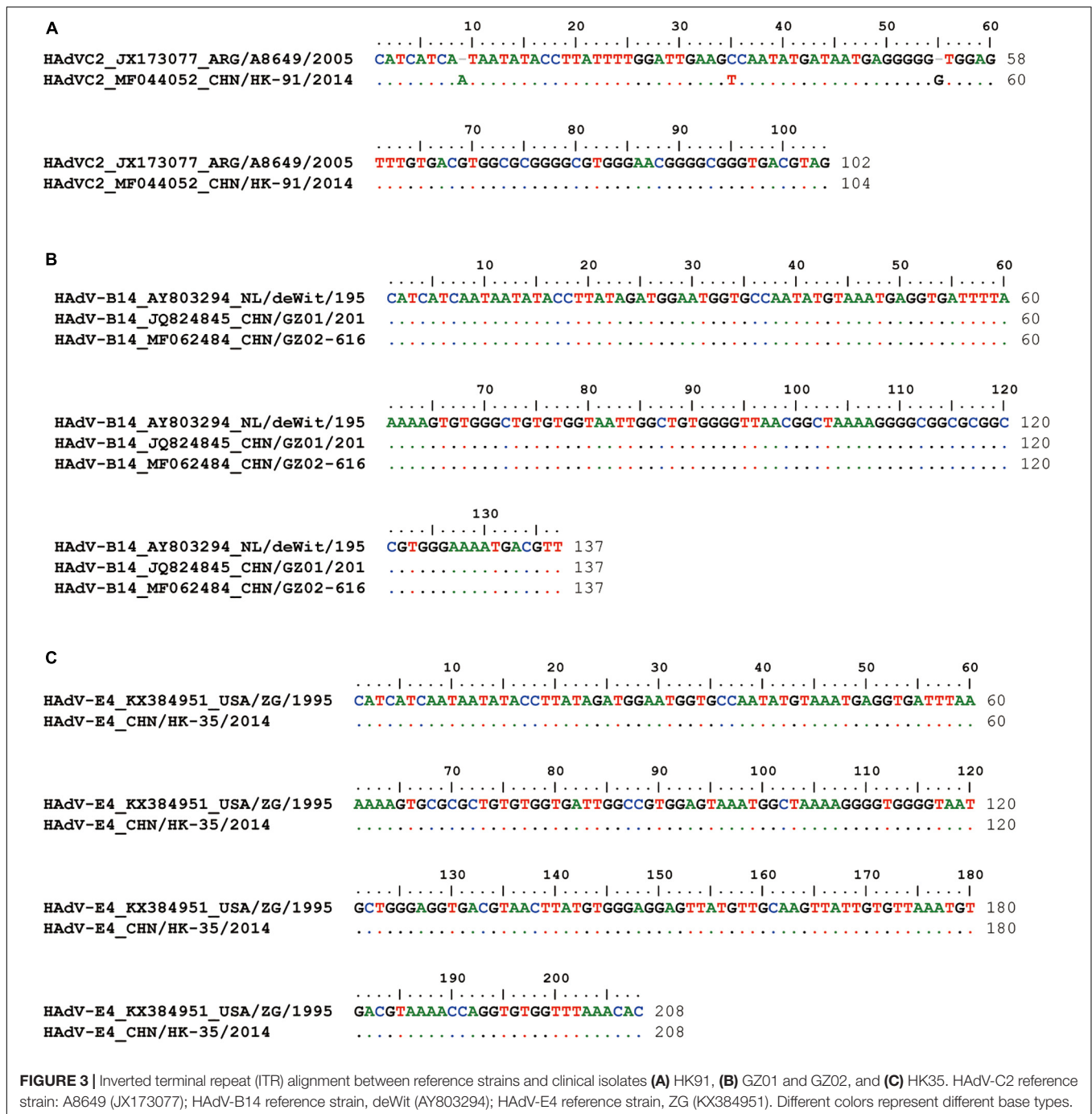
G, 21.07% T, and 56.28% GC. The genome was divided into early, intermediate, and late transcription regions (**Figure 4C**). A total of 75 features are annotated, including gene, CDS, RNA, poly(A), or poly(T).

Phylogenetic Analysis of Fiber, Hexon, Penton Base Genes, and the Whole Genomes

To find potential recombination that is usually emergent among the hexon, fiber, and penton base genes, the phylogenetic analysis of the fiber, hexon, penton base genes, and the whole genomic sequences was performed using MEGA X. Phylogenetic analysis of the HAdV-B14 whole genomes (**Figure 5A**) shows that the recent HAdV-B14 isolates (strains GZ01 and GZ02) are closely related to each other as well to the earlier reported HAdV-B14p genome (strain de Wit, 1955). By comparing the phylogenetic trees of fiber (**Figure 5B**), hexon (**Figure 5C**), and penton base genes (**Figure 5D**) of strains GZ01/GZ02, HK91, and HK35, no recombination was identified in these isolates. All the four isolates formed subclades with their prototype strains. However, new recombinant variants could be identified by applying this method to screen the clinical samples during the ARD surveillance (see section “Discussion”).

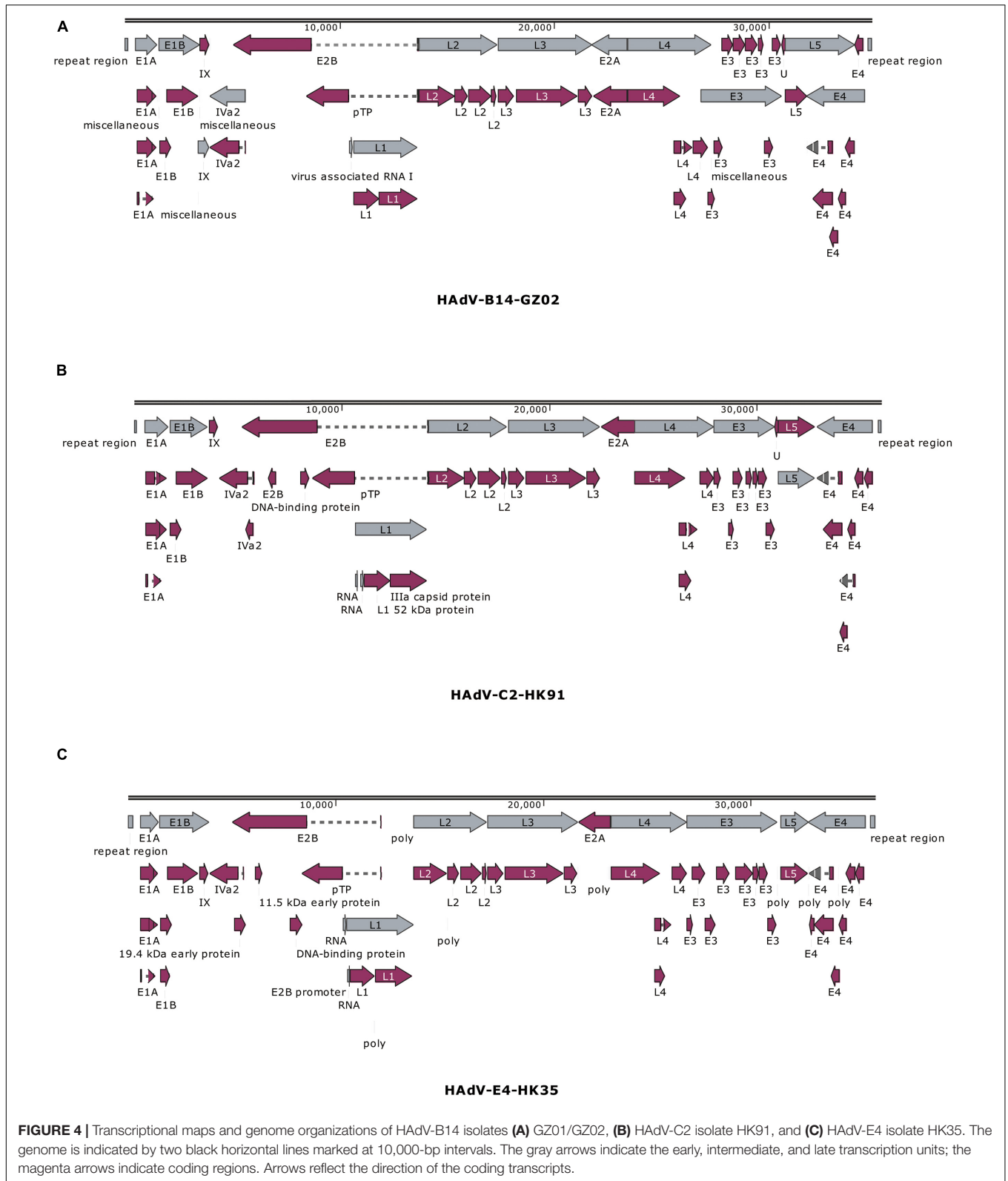
DISCUSSION

In this study, three HAdV isolates (HK35, HK91, and GZ02) were isolated and cultured from throat swabs of clinical children with respiratory diseases in Hong Kong



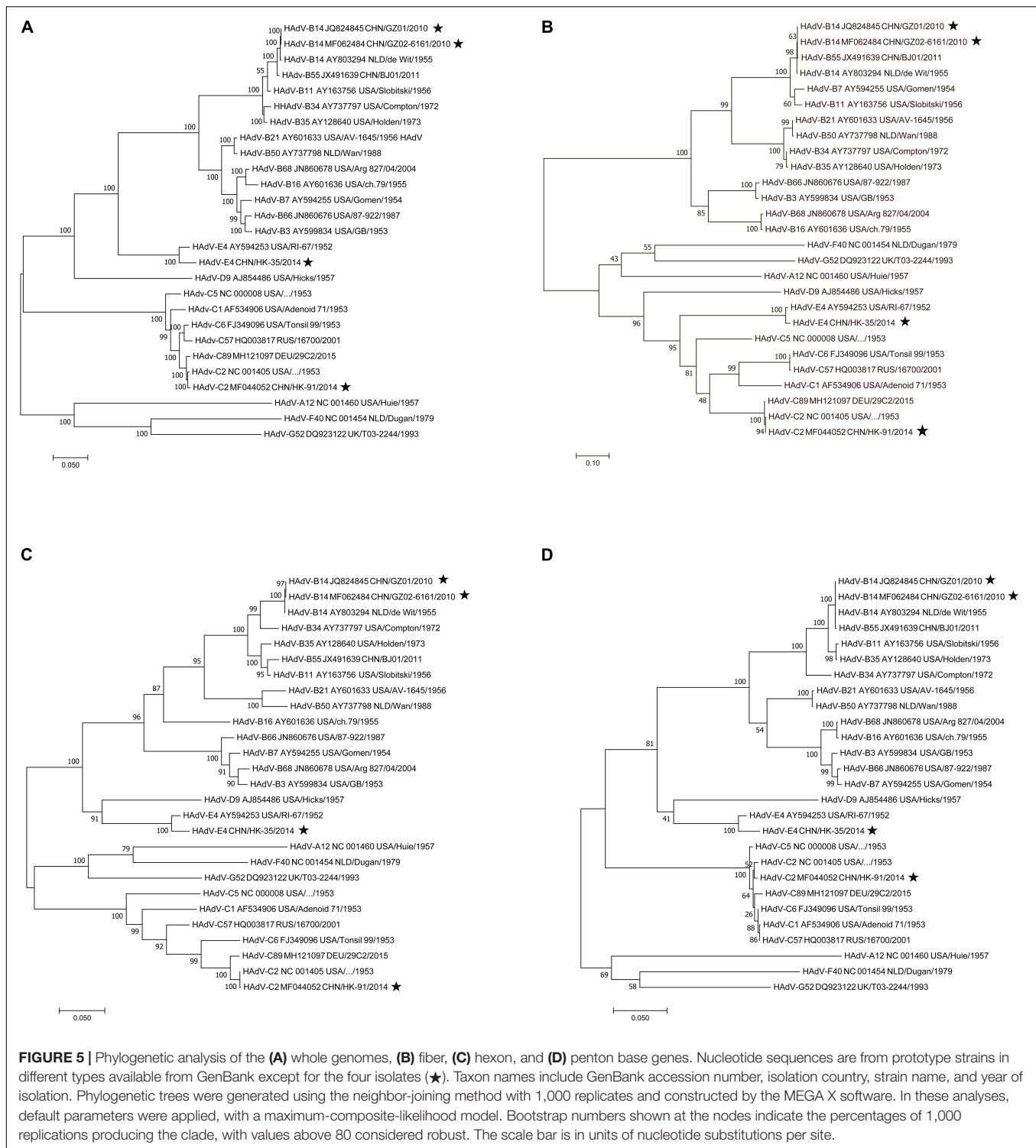
and Guangzhou. CPE were all found in A549 cells 5 days postinoculation. The complete genomes were extracted and sequenced by Sanger primer-walking method without any PCR amplification, which ensures the sequencing accuracy. The complete sequences of the ITRs were also obtained without PCR. The whole genomic sequence of HAdV-B14 strain GZ01 was resequenced using this method. It was identical to that sequenced previously with PCR plus Sanger method (GenBank no. JQ824845), which suggested that our HAdV sequencing method is highly accurate.

A HAdV-B14 variant, 14p1, emerged in the United States since 2005 and caused multiple outbreaks in both civilian and military settings and led to 76% hospitalization rate and 18% fatality rate (Binn et al., 2007; Centers for Disease Control and Prevention [CDC], 2007; Metzgar et al., 2007; Louie et al., 2008; Lewis et al., 2009). Since 2009, this genotype has caused at least two ARD outbreaks in Europe and one outbreak in Canada (Carr et al., 2011; O’Flanagan et al., 2011; Parcell et al., 2014). There were no HAdV-B14 cases in China until October 2010, when HAdV-B14 (strain GZ01) emerged in Guangzhou



(Zhang et al., 2012b, 2017). Subsequently, there were at least three additional HA Δ V-B14-related ARD outbreaks in China: 43 students with ARD in an elementary school in Gansu Province

(2011) (Huang et al., 2013), 30 adults presented with severe symptoms that required hospitalization in Beijing (2012) (Mi et al., 2013; Tang et al., 2013), and 24 students in a middle



school in Liaoning Province (2012) (Yu et al., 2014). In our study, strain GZ02 isolated in September 2010 had an identical genomic sequence with strain GZ01, which was identified as the first HAdV-B14 strain emergent in China (Zhang et al., 2012b, 2017). However, strain GZ02 was isolated 1 month earlier than strain GZ01. Therefore, we have identified that the first

HAdV-B14 isolate emergent in China is strain GZ02 (September 2010), not strain GZ01.

HAdV-C2 was previously reported as an uncommon type that causes asymptomatic or mild and self-limiting disease. However, it could lead to a severe outcome in immunocompromised hosts (Lion, 2014). Recently, more and more type 2 adenoviruses

were identified in pediatric infectious disease (Ma et al., 2015; Dhingra et al., 2019), some of which were recombinants with other members in HAdV-C (Wang et al., 2016; Dhingra et al., 2019; Zhang and Huang, 2019; Zhao et al., 2019). In this study, the complete genomic sequence of HAdV-C2 strain HK91 was obtained by our method. The genome was highly similar to that of another strain isolated in Xizang, China (99.56%).

Given that more and more respiratory cases caused by HAdV-C are emergent, further surveillance in HAdV-C pathogens is highly necessary and recommended. Although no recombination was found in these genomes, the sequencing and analysis procedure will reference clinicians and clinical laboratories and be used for other isolates in the future. Our continuous surveillance found a recombinant HAdV strain by sequencing the whole-genome sequence using this method and the phylogenetic analysis of the penton base, hexon, and fiber genes (data not shown). The penton base gene is closer to HAdV-1, but both the hexon and fiber genes are closer to HAdV-2. This strain's genomic sequence has been submitted to NCBI and HAWG to identify as a new genotype.

HAdV-E4 is a zoonotic pathogen that contains the chimpanzee adenovirus genome. It was typically restricted to United States military populations with occasional infections in civilian populations (Kajon et al., 2018). The "old" HAdV-E4 lack a critical replication motif in ITRs, nuclear factor I (NF-I), found in all HAdV respiratory pathogens, which is a host transcription factor that binds to 23–36 nt of the HAdV origin of replication (Mul et al., 1990; Hatfield and Hearing, 1991) and is recruited by the human adenoviral replication complex (Nagata et al., 1983). However, in our earlier study, we found that the re-emergent HAdV-E4 pathogens (strain HK35 and other recent HAdV-E4 isolates) circulating in children obtained NF-I in ITRs by recombination with other human adenoviruses (Zhang et al., 2019), which may facilitate HAdV-E4 to adapt hosts more efficiently. Attention should be paid to this recombinant HAdV-E4 circulating in civilian populations.

Our sequencing method improves the sequencing accuracy greatly because the genomic DNA is used as a direct sequencing template, and no PCR amplification is involved, which will avoid the base mismatch. Additionally, many ITR sequences released in GenBank are inconsistent at the left and right ends due to sequencing errors. Our method could be used to obtain the complete and accurate ITR sequences at the 5' and 3' ends of the genomes (Zhang et al., 2012a,b, 2017, 2019; Zhao et al., 2014; Zeng et al., 2016; Jing et al., 2019). The sequencing primers can be reused for the same type. Compared with the NGS method, our sequencing method could be applied for one or several clinical strains. More importantly, the gap that could not be sequenced by the second-generation sequencing but can be sequenced by our method; no PCR amplification needed.

Using the improved genomic DNA extraction protocol in our study, we obtained the HAdV genomic DNA with high quantity and quality to be used as a direct sequencing template. Li et al. (2016) compared the three methods for the adenoviral genomic DNA extraction. They found that when compared to the traditional Hirt's method, Viral DNA/RNA Extraction Kits A and B could be used to extract high-quality adenovirus genomic

DNA. However, usually if kits are used, the cellular genomic DNA is not easy to be eliminated, which may interfere our WGS in this study. Further verification would be necessary. In our study, the sequencing data were specific, and no background noise from hosts was found. The genomic DNA is also suitable for genotyping HAdVs by restriction enzyme analysis (Zhang et al., 2016). When the sequencing primers could not work due to the genome variation and poor primer binding, new primers are needed for primer-walking sequencing, which may be a little time consuming but guarantee genomic accuracy.

In conclusion, in this study, the sequencing method for the complete or partial genomes is universal, accurate, and convenient. It is an option for the genome sequencing of common human adenoviruses, especially when the samples include recombinant adenoviruses and accurate sequences are in need. The sequencing strategy may also be applied to the WGS of the other DNA viruses.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: GenBank, MF062484, MF044052, and MW692349.

AUTHOR CONTRIBUTIONS

QZ conceived and designed the experiments. JZ and ZY collected the clinical samples. SZ, WG, KM, YY, JO, and JZ performed the experiments. SZ and QZ analyzed the data. SZ, JW, and QZ contributed to the preparation of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by grants from the National Key Research and Development Program of China (2018YFE0204503), National Natural Science Foundation of China (81701995), and Natural Science Foundation of Guangdong Province (2018B030312010 and 2021A1515010788), as well as from the Guangzhou Healthcare Collaborative Innovation Major Project (201803040004 and 201803040007).

ACKNOWLEDGMENTS

We thank Patrick C. Y. Woo and Susanna K. P. Lau at the University of Hong Kong for providing the sample.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2021.661382/full#supplementary-material>

REFERENCES

- Adhikary, A. K., Hanaoka, N., and Fujimoto, T. (2014). Simple and cost-effective restriction endonuclease analysis of human adenoviruses. *BioMed. Res. Int.* 2013:363790. doi: 10.1155/2014/363790
- Binn, L. N., Sanchez, J. L., and Gaydos, J. C. (2007). Emergence of adenovirus Type 14 in US Military recruits—a new challenge. *J. Infect. Dis.* 196, 1436–1437. doi: 10.1086/522969
- Carr, M. J., Kajon, A. E., Lu, X., Dunford, L., O'Reilly, P., Holder, P., et al. (2011). Deaths associated with human adenovirus-14p1 infections, Europe, 2009–2010. *Emerg. Infect. Dis.* 17, 1402–1408. doi: 10.3201/eid1708.101760
- Centers for Disease Control and Prevention [CDC] (2007). Acute respiratory disease associated with adenovirus serotype 14—four states, 2006–2007. *MMWR Morb. Mortal. Wkly. Rep.* 56, 1181–1184.
- Cheng, Z., Yan, Y., Jing, S., Li, W.-G., Chen, W.-W., Zhang, J., et al. (2018). Comparative genomic analysis of re-emergent human adenovirus Type 55 pathogens associated with adult severe community-acquired pneumonia reveals conserved genomes and capsid proteins. *Front. Microbiol.* 9:1180. doi: 10.3389/fmicb.2018.01180
- Dhingra, A., Hage, E., Ganzenmueller, T., Bottcher, S., Hofmann, J., Hamprecht, K., et al. (2019). Molecular evolution of human adenovirus (HAdV) species C. *Sci. Rep.* 9:1039. doi: 10.1038/s41598-018-37249-4
- Han, G., Niu, H., Zhao, S., Zhu, B., Wang, C., Liu, Y., et al. (2013). Identification and typing of respiratory adenoviruses in Guangzhou, Southern China using a rapid and simple method. *Virology* 184, 103–108. doi: 10.1007/s12250-013-3308-7
- Hatfield, L., and Hearing, P. (1991). Redundant elements in the adenovirus type 5 inverted terminal repeat promote bidirectional transcription in vitro and are important for virus growth in vivo. *Virology* 184, 265–276.
- Hirt, B. (1967). Selective extraction of polyoma DNA from infected mouse cell cultures. *J. Mol. Biol.* 26, 365–369.
- Huang, G., Yu, D., Zhu, Z., Zhao, H., Wang, P., Gray, G. C., et al. (2013). Outbreak of febrile respiratory illness associated with human adenovirus type 14p1 in Gansu Province, China. *Influenza Other Respir. Viruses* 7, 1048–1054. doi: 10.1111/irv.12118
- Jing, S., Zhang, J., Cao, M., Liu, M., Yan, Y., Zhao, S., et al. (2019). Household transmission of human adenovirus Type 55 in case of fatal acute respiratory disease. *Emerg. Infect. Dis.* 25, 1756–1758. doi: 10.3201/eid2509.181937
- Kajon, A. E., Lamson, D. M., Bair, C. R., Lu, X., Landry, M. L., Menegus, M., et al. (2018). Adenovirus Type 4 respiratory infections among civilian adults, Northeastern United States, 2011–2015. *Emerg. Infect. Dis.* 24, 201–209. doi: 10.3201/eid2402.171407
- Koboldt, D. C., Steinberg, K. M., Larson, D. E., Wilson, R. K., and Mardis, E. R. (2013). The next-generation sequencing revolution and its impact on genomics. *Cell* 155, 27–38. doi: 10.1016/j.cell.2013.09.006
- Lewis, P. F., Schmidt, M. A., Lu, X., Erdman, D. D., Campbell, M., Thomas, A., et al. (2009). A community-based outbreak of severe respiratory illness caused by human adenovirus serotype 14. *J. Infect. Dis.* 199, 1427–1434. doi: 10.1086/598521
- Li, X., Zhou, R., Ma, Q., Jiang, Z., Zhou, Z., Liao, X., et al. (2016). The impact analysis of three different genome extraction methods for adenovirus research. *China Medical Herald* 13, 4–7.
- Lion, T. (2014). Adenovirus infections in immunocompetent and immunocompromised patients. *Clin. Microbiol. Rev.* 27, 441–462.
- Louie, J. K., Kajon, A. E., Holodniy, M., Guardia-LaBar, L., Lee, B., Petru, A. M., et al. (2008). Severe pneumonia due to adenovirus serotype 14: a new respiratory threat? *Clin. Infect. Dis.* 46, 421–425. doi: 10.1086/525261
- Lu, H., Giordano, F., and Ning, Z. (2016). Oxford nanopore MinION sequencing and genome assembly. *Genomics Proteomics Bioinformatics* 14, 265–279. doi: 10.1016/j.gpb.2016.05.004
- Ma, G., Zhu, Y., Xiao, Y., Ji, Y., Wu, J., Bao, J., et al. (2015). Species C is predominant in Chinese children with acute respiratory adenovirus infection. *Pediatr. Infect. Dis. J.* 34:1042. doi: 10.1097/INF.0000000000000791
- Marinho, J. C., Dos Santos, T. G., Siqueira-Silva, J., Lu, X., Carvalho, D., da Camara, A. A., et al. (2011). A naturally occurring human adenovirus type 7 variant with a 1743 bp deletion in the E3 cassette. *J. Gen. Virol.* 92(Pt. 10), 2399–2404. doi: 10.1099/vir.0.029181-0
- Metzgar, D., Osuna, M., Kajon, A. E., Hawksworth, A. W., Irvine, M., and Russell, K. L. (2007). Abrupt emergence of diverse species B adenoviruses at US military recruit training centers. *J. Infect. Dis.* 196, 1465–1473.
- Mi, Z., Butt, A. M., An, X., Jiang, T., Liu, W., Qin, C., et al. (2013). Genomic analysis of HAdV-B14 isolate from the outbreak of febrile respiratory infection in China. *Genomics* 102, 448–455. doi: 10.1016/j.ygeno.2013.09.001
- Mul, Y. M., Verrijzer, C. P., and van der Vliet, P. C. (1990). Transcription factors NFI and NFIII/oct-1 function independently, employing different mechanisms to enhance adenovirus DNA replication. *J. Virol.* 64, 5510–5518.
- Nagata, K., Guggenheimer, R. A., and Hurwitz, J. (1983). Specific binding of a cellular DNA replication protein to the origin of replication of adenovirus DNA. *Proc. Natl. Acad. Sci. U.S.A.* 80, 6177–6181.
- O'Flanagan, D., O'Donnell, J., Domegan, L., Fitzpatrick, F., Connell, J., Coughlan, S., et al. (2011). First reported cases of human adenovirus serotype 14p1 infection, Ireland, October 2009 to July 2010. *Euro Surveill.* 16:19801.
- Pan, H., Yan, Y., Zhang, J., Zhao, S., Feng, L., Ou, J., et al. (2018). Rapid construction of a replication-competent infectious clone of human adenovirus Type 14 by gibson assembly. *Viruses* 10:568. doi: 10.3390/v10100568
- Pan, Q., Wang, J., Gao, Y., Cui, H., Liu, C., Qi, X., et al. (2018). The natural large genomic deletion is unrelated to the increased virulence of the novel genotype fowl adenovirus 4 recently emerged in China. *Viruses* 10:494. doi: 10.3390/v10090494
- Parcell, B. J., McIntyre, P. G., Yirrell, D. L., Fraser, A., Quinn, M., Templeton, K., et al. (2014). Prison and community outbreak of severe respiratory infection due to adenovirus type 14p1 in Tayside, UK. *J. Public Health* 37, 64–69. doi: 10.1093/pubmed/ftu009
- Robinson, E. R., Walker, T. M., and Pallen, M. J. (2013). Genomics and outbreak investigation: from sequence to consequence. *Genome Med.* 5:36. doi: 10.1186/gm440
- Rusk, N. (2019). More accurate nanopore sequencing. *Nat. Methods* 16:460. doi: 10.1038/s41592-019-0449-0
- Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* 74, 5463–5467.
- Scott, M. K., Chommanard, C., Lu, X., Appelgate, D., Grenz, L., Schneider, E., et al. (2016). Human adenovirus associated with severe respiratory infection, Oregon, USA, 2013–2014. *Emerg. Infect. Dis.* 22, 1044–1051. doi: 10.3201/eid2206.151898
- Seto, D., Chodosh, J., Brister, J. R., Jones, M. S., and Members of the Adenovirus Research Community (2011). Using the whole-genome sequence to characterize and name human adenoviruses. *J. Virol.* 85, 5701–5702. doi: 10.1128/jvi.00354-11
- Seto, J., Walsh, M., Mahadevan, P., Zhang, Q., and Seto, D. (2010). Applying genomic and bioinformatic resources to human adenovirus genomes for use in vaccine development and for applications in vector development for gene delivery. *Viruses* 2, 1–26.
- Shen, Y., Liu, J., Zhang, Y., Ma, X., Yue, H., and Tang, C. (2020). Prevalence and characteristics of a novel bovine adenovirus type 3 with a natural deletion fiber gene. *Infect. Genet. Evol.* 83:104348. doi: 10.1016/j.meegid.2020.104348
- Shendure, J., Balasubramanian, S., Church, G. M., Gilbert, W., Rogers, J., Schloss, J. A., et al. (2017). DNA sequencing at 40: past, present and future. *Nature* 550, 345–353. doi: 10.1038/nature24286
- Su, X., Tian, X., Zhang, Q., Li, H., Li, X., Sheng, H., et al. (2011). Complete genome analysis of a novel E3-partial-deleted human adenovirus type 7 strain isolated in Southern China. *Virology* 418:8–91. doi: 10.1016/j.virol.2011.07.011
- Tang, L., An, J., Xie, Z., Dehghan, S., Seto, D., Xu, W., et al. (2013). Genome and bioinformatic analysis of a HAdV-B14p1 virus isolated from a baby with pneumonia in Beijing, China. *PLoS One* 8:e60345. doi: 10.1371/journal.pone.0060345
- van Dijk, E. L., Jaszczyszyn, Y., Naquin, D., and Thermes, C. (2018). The third revolution in sequencing technology. *Trends Genet.* 34, 666–681. doi: 10.1016/j.tig.2018.05.008
- Wang, Y., Li, Y., Lu, R., Zhao, Y., Xie, Z., Shen, J., et al. (2016). Phylogenetic evidence for intratypic recombinant events in a novel human adenovirus C that causes severe acute respiratory infection in children. *Sci. Rep.* 6:23014. doi: 10.1038/srep23014
- Yan, Y., Jing, S., Feng, L., Zhang, J., Zeng, Z., Li, M., et al. (2020a). Construction and characterization of a novel recombinant attenuated and replication-deficient

- candidate human adenovirus Type 3 vaccine: “adenovirus vaccine within an adenovirus vector”. *Viol. Sin.* doi: 10.1007/s12250-020-00234-1 [Epub ahead of print].
- Yan, Y., Ou, J., Zhao, S., Ma, K., Lan, W., Guan, W., et al. (2020b). Characterization of Influenza A and B viruses circulating in southern china during the 2017–2018 season. *Front. Microbiol.* 11:1079. doi: 10.3389/fmicb.2020.01079
- Yu, W., Sun, H., Tian, J., Wang, L., Wang, B., Chen, Y., et al. (2014). Molecular epidemic characteristics analysis of an adenovirus infection outbreak. *Chin. J. Public Health* 7, 972–974.
- Yu, Z., Zeng, Z., Zhang, J., Pan, Y., Chen, M., Guo, Y., et al. (2016). Fatal community-acquired pneumonia in children caused by re-emergent human adenovirus 7d associated with higher severity of illness and fatality rate. *Sci. Rep.* 6:37216. doi: 10.1038/srep37216
- Zeng, Z., Zhang, J., Jing, S., Cheng, Z., Bofill-Mas, S., Maluquer de Motes, C., et al. (2016). Genome sequence of a cynomolgus macaque adenovirus (CynAdV-1) isolate from a primate colony in the United Kingdom. *Genome Announc.* 4, e01193–16. doi: 10.1128/genomeA.01193-16
- Zhang, G., Zhang, Y., and Jin, J. (2021). The ultrafast and accurate mapping algorithm FANSe3: mapping a human whole-genome sequencing dataset within 30 Minutes. *Phenomics* 1, 22–30.
- Zhang, J., Kang, J., Dehghan, S., Sridhar, S., Lau, S., Ou, J., et al. (2019). A survey of recent adenoviral respiratory pathogens in hong kong reveals emergent and recombinant human adenovirus type 4 (HAdV-E4) circulating in civilian populations. *Viruses* 11:129. doi: 10.3390/v11020129
- Zhang, Q., Dehghan, S., and Seto, D. (2016). Pitfalls of restriction enzyme analysis in identifying, characterizing, typing, and naming viral pathogens in the era of whole genome data, as illustrated by HAdV type 55. *Viol. Sin.* 31, 448–453. doi: 10.1007/s12250-016-3862-x
- Zhang, Q., Jing, S., Cheng, Z., Yu, Z., Dehghan, S., Shamsaddini, A., et al. (2017). Comparative genomic analysis of two emergent human adenovirus type 14 respiratory pathogen isolates in China reveals similar yet divergent genomes. *Emerg. Microbes Infect.* 6:e92. doi: 10.1038/emi.2017.78
- Zhang, Q., Seto, D., Cao, B., Zhao, S., and Wan, C. (2012a). Genome sequence of human adenovirus type 55, a re-emergent acute respiratory disease pathogen in China. *J. Virol.* 86, 12441–12442. doi: 10.1128/JVI.02225-12
- Zhang, Q., Seto, D., Zhao, S., Zhu, L., Zhao, W., and Wan, C. (2012b). Genome sequence of the first human adenovirus type 14 isolated in China. *J. Virol.* 86, 7019–7020. doi: 10.1128/jvi.00814-12
- Zhang, Q., Su, X., Seto, D., Zheng, B.-J., Tian, X., Sheng, H., et al. (2009). Construction and characterization of a replication-competent human adenovirus type 3-based vector as a live-vaccine candidate and a viral delivery vector. *Vaccine* 27, 1145–1153. doi: 10.1016/j.vaccine.2008.12.039
- Zhang, W., and Huang, L. (2019). Genome analysis of a novel recombinant human adenovirus type 1 in China. *Sci. Rep.* 9:4298. doi: 10.1038/s41598-018-37756-4
- Zhao, M.-C., Guo, Y.-H., Qiu, F.-Z., Wang, L., Yang, S., Feng, Z.-S., et al. (2019). Molecular and clinical characterization of human adenovirus associated with acute respiratory tract infection in hospitalized children. *J. Clin. Virol.* 123:104254.
- Zhao, S., Wan, C., Ke, C., Seto, J., Dehghan, S., Zou, L., et al. (2014). Re-emergent human adenovirus genome type 7d caused an acute respiratory disease outbreak in Southern China after a twenty-one year absence. *Sci. Rep.* 4:7365. doi: 10.1038/srep07365

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Zhao, Guan, Ma, Yan, Ou, Zhang, Yu, Wu and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.