



# Complete Genome Sequences, Derived by Next-Generation Sequencing, of JC Polyomavirus Strains Isolated from Vietnamese Renal Transplant Recipients

Trang Dinh Van,<sup>a,b</sup> Rebecca J. Rockett,<sup>b,c</sup> An Phan Hai Ha,<sup>d,e</sup> Trung Vu Nguyen,<sup>a,f</sup> Dominic E. Dwyer<sup>b,c</sup>

<sup>a</sup>Clinical Laboratory, National Hospital of Tropical Diseases, Hanoi, Vietnam

<sup>b</sup>Sydney Medical School, The University of Sydney, Camperdown, NSW, Australia

<sup>c</sup>Centre for Infectious Diseases and Microbiology Laboratory Services, NSW Health Pathology-Institute of Clinical Pathology and Medical Research, Westmead, NSW, Australia

<sup>d</sup>International Cooperation Department, Hanoi Medical University, Hanoi, Vietnam

<sup>e</sup>Kidney Diseases and Dialysis Department, Viet Duc Hospital, Hanoi, Vietnam

<sup>f</sup>Medical Microbiology Department, Hanoi Medical University, Hanoi, Vietnam

**ABSTRACT** JC polyomavirus (JCPyV) may cause clinical syndromes such as progressive multifocal leukoencephalopathy in immunocompromised patients. Here, we report seven complete genome sequences of JCPyV genotype 7A, generated directly from urine samples from Vietnamese renal transplant recipients by using rolling-circle amplification and next-generation sequencing.

JC polyomavirus (JCPyV) is a circular double-stranded DNA virus from the *Betapolyomavirus* genus of the *Polyomaviridae* family. In highly immunosuppressed individuals, JCPyV can cause severe disease, most commonly progressive multifocal leukoencephalopathy (1). JCPyV genotypes have been used to trace human migration, with genotypes 2, 4, and 7 predominating in Southeast Asia (2, 3). This study used a targeted metagenomic method to produce draft full-length JCPyV genome sequences from seven renal transplant recipients from Hanoi, Vietnam.

Seven urine samples were collected from Vietnamese renal transplant recipients between 2015 and 2017. Viral DNA was extracted from 250  $\mu$ l of urine (already identified as BK polyomavirus positive using the artus BK virus RG PCR kit [Qiagen GmbH, Hilden, Germany]) with the EZ1 DSP virus kit (Qiagen). To ensure sufficient polyomavirus DNA reads from the clinical specimens, a primer-directed rolling-circle amplification method was used to enrich DNA extracts for JCPyV (4). After enrichment, DNA libraries prepared by employing the Nextera XT library preparation kit v2 were sequenced on a NextSeq 500 system (Illumina, Australia) using paired-end 150-bp chemistry.

Raw sequencing reads were trimmed with Trimmomatic v0.36 (sliding window; minimum Phred score, 20) (5). Human reads were removed by mapping reads to a JCPyV reference genome (GenBank accession number [J02226](https://www.ncbi.nlm.nih.gov/nuccore/J02226)) using Burrows-Wheeler alignment (BWA-MEM v0.7.12) (6). Mapped reads were converted to fastq files using SAMtools v1.6 (7) (the numbers of mapped reads [percentage of total reads] are as follows: VN-24, 3,609,902 [99%]; VN-67, 217,170 [52%]; VN-68, 1,346,904 [34%]; VN-214, 5,346,640 [99%]; VN-247, 3,447,166 [99%]; VN-291, 4,586,564 [99%]; and VN-347, 3,557,524 [77%]) (8). Reference mapping, *de novo* assembly, and alignments were conducted using the software package CLC Bio Genomics Workbench v9.0 (Qiagen).

Read mapping to the reference sequence and *de novo* assembly were performed in parallel to confirm the draft genome sequences, because indels and duplications are known to occur within the noncoding control region. Reference mapping

**Citation** Van TD, Rockett RJ, Ha APH, Nguyen TV, Dwyer DE. 2020. Complete genome sequences, derived by next-generation sequencing, of JC polyomavirus strains isolated from Vietnamese renal transplant recipients. *Microbiol Resour Announc* 9:e01350-19. <https://doi.org/10.1128/MRA.01350-19>.

**Editor** Steven R. Gill, University of Rochester School of Medicine and Dentistry

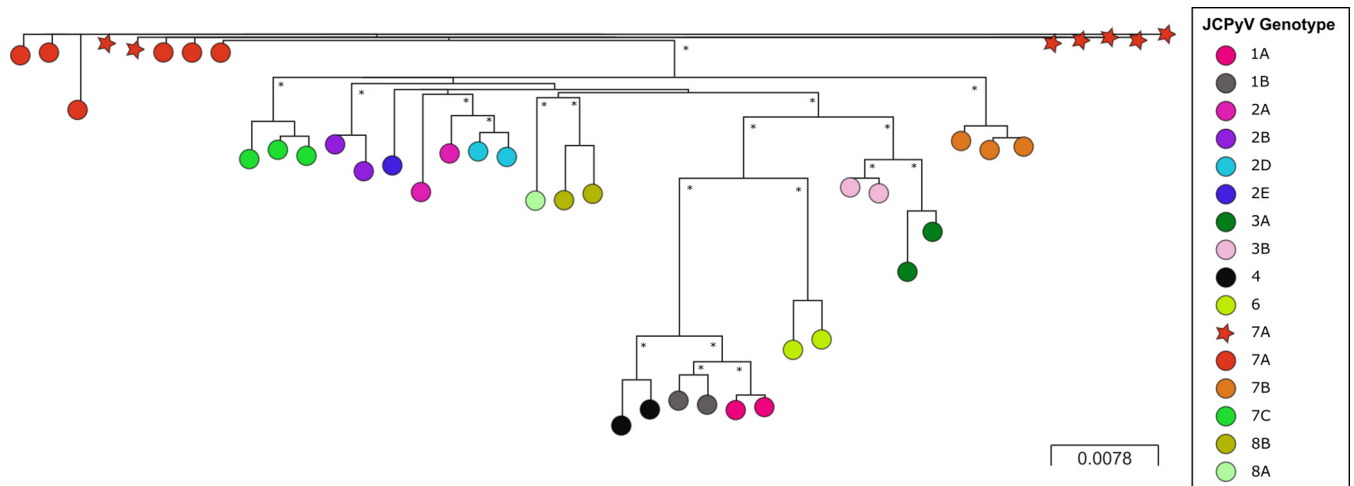
**Copyright** © 2020 Van et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Trang Dinh Van, [vandinhtrang.nhtd@gmail.com](mailto:vandinhtrang.nhtd@gmail.com).

**Received** 26 October 2019

**Accepted** 4 December 2019

**Published** 9 January 2020



**FIG 1** Unrooted phylogenetic tree constructed from the consensus sequences of 43 JCPyV genomes with the noncoding control region removed (alignment length, 4,854 bp). The seven new JCPyV genomes (ENA accession numbers [LR215999](#) to [LR216005](#)) produced in this study are depicted as red stars. Previously characterized JCPyV reference sequences (GenBank accession numbers [AB038249](#), [U61771](#), [AB048547](#), [AB048550](#), [JUU73501](#), [AF295735](#), [AF295737](#), [AF295738](#), [AF300960](#), [AF300965](#), [AF300963](#), [AF363833](#), [AF363832](#), [AF300950](#), [AF300951](#), [AF300957](#), [AF300959](#), [AF015536](#), [AF004350](#), [AF015533](#), [AF396423](#), [AF295732](#), [AF396432](#), [AF281623](#), [AB038252](#), [AF015537](#), [AF015534](#), [AB038251](#), [AB048576](#), [AB048574](#), [AB048563](#), [J02226](#), [AF281625](#), and [AF015528](#)), representing a range of JCPyV genomes, were also included in the phylogeny (3, 10). The tree was constructed using IQ-TREE software (11) (ModelFinder; substitution model, HKY+F+R2; number of bootstrap replicates, 1,000). The phylogeny was visualized using Microreact and annotated using Inkscape (12). Node colors represent each of the JCPyV genotypes; all of the draft JCPyV genomes produced in this study are genotype 7A. These genomes have high homology (3 to 7 single-nucleotide polymorphisms) to other JCPyV genotype 7A reference sequences (GenBank accession numbers [AF295737](#), [AF295738](#), [AF300960](#), [AF300965](#), [AF300963](#), and [U61771](#)) and also are closely related to each other (0 to 3 single-nucleotide polymorphisms) (13). The scale bar represents the number of nucleotide substitutions per site. Asterisks indicate branches with bootstrap support of >85%.

resulted in high average coverages ( $83,802 \times$  [VN-24];  $56,675 \times$  [VN-67];  $35,023 \times$  [VN-68];  $143,065 \times$  [VN-214];  $87,133 \times$  [VN-247];  $116,532 \times$  [VN-291]; and  $90,135 \times$  [VN-347]) and a GC content of 40.2%. *De novo* assembly produced a single contig spanning the length of the JCPyV genome for specimens VN-24 ( $96,285 \times$ , 5,116 bp), VN-214 ( $142,644 \times$ , 5,115 bp), VN-247 ( $89,920 \times$ , 5,116 bp), and VN-291 ( $119,699 \times$ , 5,116 bp). The remaining specimens produced a maximum of 3 contigs with BLAST+ identity to JCPyV (9) (VN-67, 2 contigs,  $53,880 \times$ , 3,933 bp; VN-68, 3 contigs,  $33,746 \times$ , 3,952 bp; VN-347, 2 contigs,  $88,362 \times$ , 3,937 bp).

Phylogenetic analysis of draft assemblies indicated that all seven JCPyV genomes were genotype 7A (Fig. 1), which is consistent with reports that this genotype predominates in Southeast Asia. Currently, only a small number of genotype 7 JCPyV genomes are publicly available. This result also demonstrates that JCPyV can be clearly differentiated using metagenomics and that high levels of JCPyV in the viromes of renal transplant recipients may be common.

**Data availability.** Raw fastq files and final consensus draft JCPyV genomes have been deposited in the European Nucleotide Archive (ENA) under accession numbers [ERR3561211](#) to [ERR3561215](#) (fastq files) and [LR215999](#) to [LR216005](#) (genomes).

## ACKNOWLEDGMENTS

We acknowledge the Sydney Informatics Hub and Artemis, the University of Sydney high-performance computing cluster, for providing computing resources. We thank the staff of the NSW Health Pathology-Institute of Clinical Pathology and Medical Research (NSWHP-ICPMR) at Westmead Hospital for technical assistance and Chau Minh Tran from Hanoi Medical University (Hanoi, Vietnam) for DNA extraction and polyomavirus reverse transcription-PCR.

This work was funded by an Australian Agency for International Development scholarship and supported with a grant from the Centre for Infectious Diseases and Microbiology-Public Health, NSWHP-ICPMR, Westmead Hospital (Westmead, Australia).

## REFERENCES

1. Hirsch HH, Kardas P, Kranz D, Leboeuf C. 2013. The human JC polyomavirus (JCPyV): virological background and clinical implications. *APMIS* 121:685–727. <https://doi.org/10.1111/apm.12128>.
2. Stoner GL, Jobes DV, Cobo MF, Agostini HT, Chima SC, Ryschkewitsch CF. 2000. JC virus as a marker of human migration to the Americas. *Microbes Infect* 2:1905–1911. [https://doi.org/10.1016/S1286-4579\(00\)01339-3](https://doi.org/10.1016/S1286-4579(00)01339-3).
3. Hu C, Huang Y, Su J, Wang M, Zhou Q, Zhu B. 2018. Detection and analysis of variants of JC polyomavirus in urine samples from HIV-1-infected patients in China's Zhejiang Province. *J Int Med Res* 46:1024–1032. <https://doi.org/10.1177/0300060517746297>.
4. Rockett R, Barraclough KA, Isbel NM, Dudley KJ, Nissen MD, Sloots TP, Bialasiewicz S. 2015. Specific rolling circle amplification of low-copy human polyomaviruses BKV, HPyV6, HPyV7, TSPyV, and STLPyV. *J Virol Methods* 215–216:17–21. <https://doi.org/10.1016/j.jviromet.2015.02.004>.
5. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
6. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>.
7. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25:2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
8. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUILT: quality assessment tool for genome assemblies. *Bioinformatics* 29:1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>.
9. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. <https://doi.org/10.1186/1471-2105-10-421>.
10. Ou W, Tsai R, Wang M, Fung C, Hseu T, Chang D. 1997. Genomic cloning and sequence analysis of Taiwan-3 human polyomavirus JC virus. *J Formos Med Assoc* 96:511–516.
11. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 32:268–274. <https://doi.org/10.1093/molbev/msu300>.
12. Argimón S, Abudahab K, Goater RJ, Fedosejev A, Bhai J, Glasner C, Feil EJ, Holden MT, Yeats CA, Grundmann H, Spratt BG, Aanensen DM. 2016. Microreact: visualizing and sharing data for genomic epidemiology and phylogeography. *Microb Genom* 2:e000093. <https://doi.org/10.1099/mgen.0.000093>.
13. Page AJ, Taylor B, Delaney AJ, Soares J, Seemann T, Keane JA, Harris SR. 2016. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microb Genom* 2:e000056. <https://doi.org/10.1099/mgen.0.000056>.