



Complete Genome Sequences of Six Novel *Macaca mulatta* Papillomavirus Types Isolated from Genital Sites of Rhesus Monkeys in Hong Kong SAR, China

 Teng Long,^a Po Yee Wong,^a Wendy C. S. Ho,^a  Robert D. Burk,^{b,c,d,e} Paul K. S. Chan,^{a,f}  Zigui Chen^{a,f}

^aDepartment of Microbiology, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong Special Administrative Region, China

^bDepartment of Pediatrics, Albert Einstein College of Medicine, Bronx, New York, USA

^cDepartment of Microbiology and Immunology, Albert Einstein College of Medicine, Bronx, New York, USA

^dDepartment of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, New York, USA

^eDepartment of Gynecology and Woman's Health, Albert Einstein College of Medicine, Bronx, New York, USA

^fCentre for Gut Microbiota Research, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong Special Administrative Region, China

ABSTRACT The complete genomes of six *Macaca mulatta* papillomavirus types isolated from genital sites of rhesus monkeys were characterized, and less than 72% identity with the complete L1 genes of known papillomaviruses was found. *Macaca mulatta* papillomavirus type 2 (MmPV2), MmPV3, and MmPV6 cluster into the genus *Alphapapillomavirus*, and MmPV4, MmPV5, and MmPV7 cluster into the genus *Gammapapillomavirus*.

Papillomaviruses (PVs) are a heterogeneous family of viruses with circular double-stranded DNA genomes about 8 kb long that infect the epithelial surfaces of skin and oral and anogenital sites from a large spectrum of vertebrates, including reptiles, fish, birds, mammals, and humans. Persistent infection by high-risk human papillomaviruses (HPVs) has been linked to cervical cancer. PV genomes usually contain 6 early genes (E6, E7, E1, E2, E4, and E5), 2 late genes (L2 and L1), and an upstream regulatory region (URR) between the L1 and E6 genes. A distinct PV type is defined when the L1 nucleotide sequence is at least 10% dissimilar from that of any other type (1, 2). Currently, the complete genomes for more than 200 types of HPVs and 160 types of nonhuman animal PVs have been characterized (3).

In this report, we characterize the complete genomes of six novel *Macaca mulatta* PV types that are distributed in Hong Kong SAR, China, and were isolated from exfoliated vaginal cells of one adult female rhesus monkey and swabs from the penile surfaces of two adult male monkeys (Table 1). Total DNA was purified using a QIAamp DNA minikit (Qiagen, USA) and was pre-enriched for circular viral DNA using a TempliPhi amplification kit (GE Healthcare, USA). Next-generation sequencing libraries prepared with a KAPA HTP library preparation kit (Roche, USA) underwent metagenomic sequencing on a HiSeq 4000 platform (Illumina, USA) using paired-end 100-bp reads. A total of 92,908,115 raw sequence reads were trimmed for adaptors and low-quality bases using Trimmomatic v0.38 (4) and were further filtered for host genome contamination using Bowtie 2 v2.3.4 (5) against a rhesus chromosome reference (rheMac8, University of California Santa Cruz [UCSC]). High-quality short reads were assembled *de novo* to build contigs using MEGAHIT v1.1.3 (6). Whole viral genomes were validated using type-specific PCR amplification in three or four overlapping fragments and Sanger sequencing using a primer-walking strategy. Open reading frames (ORFs) were predicted using Geneious v9 and searched for homology using a BLASTX search against the closely related HPV types.

Received 11 October 2018 **Accepted** 8 November 2018 **Published** 6 December 2018

Citation Long T, Wong PY, Ho WCS, Burk RD, Chan PKS, Chen Z. 2018. Complete genome sequences of six novel *Macaca mulatta* papillomavirus types isolated from genital sites of rhesus monkeys in Hong Kong SAR, China. *Microbiol Resour Annot* 7:e01414-18. <https://doi.org/10.1128/MRA.01414-18>.

Editor Julie C. Dunning Hotopp, University of Maryland School of Medicine

Copyright © 2018 Long 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 Paul K. S. Chan, paulkschan@cuhk.edu.hk, or Zigui Chen, zigui.chen@cuhk.edu.hk.

TABLE 1 Novel papillomaviruses isolated from rhesus monkeys (*Macaca mulatta*)

<i>Macaca mulatta</i> papillomavirus type	Isolate	Sex	Collection site	NCBI accession no.	Genus	Length in bp (% GC content)	Closest HPV type (L1 nt identity [%]) ^a
2	PM06953_c168082	Female	Vagina	MG837557	<i>Alphapapillomavirus</i>	7,876 (50.4)	HPV54 (70.7)
3	PM08453_c177403	Male	Penis	MG837558	<i>Alphapapillomavirus</i>	8,051 (52.0)	HPV117 (70.3)
4	PM08453_c176982	Male	Penis	MG837559	<i>Gammapapillomavirus</i>	7,424 (46.2)	HPV130 (65.7)
5	PM01953_c169203	Male	Penis	MH745747	<i>Gammapapillomavirus</i>	7,408 (40.6)	HPV134 (71.7)
6	PM08453_c170482	Male	Penis	MH745748	<i>Alphapapillomavirus</i>	7,935 (51.8)	HPV54 (71.6)
7	PM08453_c160986	Male	Penis	MH745749	<i>Gammapapillomavirus</i>	7,292 (38.7)	HPVw34c04a (72.7) ^b

^a nt, nucleotide.

^b NCBI accession number [MF588693](#).

Six novel PV genomes were named *Macaca mulatta* papillomaviruses 2, 3, 4, 5, 6, and 7 (MmPV2 to MmPV7, respectively) following the previously reported MmPV1 (7). The L1 nucleotide sequences of these novel PV types share 55.8% to 75.6% pairwise identity with each other and are less than 72% identical to those of any other PVs. A RAxML phylogenetic tree inferred from the L1 nucleotide sequences clustered MmPV2, -3, and -6 into the genus *Alphapapillomavirus*, and MmPV4, -5, and -7 cluster into the genus *Gammapapillomavirus*. Identification of novel papillomaviruses from nonhuman primate animals provides important insights into the complex phylodynamic interactions between viruses and primate hosts.

Data availability. The complete genome sequences of MmPV2 to MmPV7 were submitted to NCBI/GenBank under the accession numbers [MG837557](#) to [MG837559](#) and [MH745747](#) to [MH745749](#). The raw Illumina reads were deposited in the Sequence Read Archive under the accession numbers [SRX4994337](#), [SRX4994338](#), and [SRX4994339](#). The phylogenetic tree was deposited in figshare (<https://doi.org/10.6084/m9.figshare.7301432>).

ACKNOWLEDGMENTS

We thank the Agriculture, Fisheries, and Conservation Department of the Government of the Hong Kong Special Administrative Region for access to the macaques and the Ocean Park Conservation Foundation Hong Kong for assistance with sample collection. We also thank Samuel Tong and Jeffrey Xu for assistance with sample transportation and processing.

This study was supported in part by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (project number CUHK 24100716 to Z.C.) and the Seed Fund for Gut Microbiota Research by the Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong Special Administrative Region, China (to P.K.S.C.). The funder had no role in study design, data collection, analysis, interpretation, or writing of the report.

REFERENCES

- Bernard HU, Burk RD, Chen Z, van Doorslaer K, Zur Hausen H, de Villiers EM. 2010. Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology* 401:70–79. <https://doi.org/10.1016/j.virol.2010.02.002>.
- de Villiers EM, Fauquet C, Broker TR, Bernard HU, Zur Hausen H. 2004. Classification of papillomaviruses. *Virology* 324:17–27. <https://doi.org/10.1016/j.virol.2004.03.033>.
- Van Doorslaer K, Chen Z, Bernard HU, Chan PKS, DeSalle R, Dillner J, Forslund O, Haga T, McBride AA, Villa LL, Burk RD, ICTV Report Consortium. 2018. ICTV virus taxonomy profile: Papillomaviridae. *J Gen Virol* 99:989–990. <https://doi.org/10.1099/jgv.0.001105>.
- 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>.
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359. <https://doi.org/10.1038/nmeth.1923>.
- Li D, Luo R, Liu CM, Leung CM, Ting HF, Sadakane K, Yamashita H, Lam TW. 2016. MEGAHIT v1.0: a fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods* 102:3–11. <https://doi.org/10.1016/j.ymeth.2016.02.020>.
- Ostrow RS, LaBresh KV, Faras AJ. 1991. Characterization of the complete RhPV 1 genomic sequence and an integration locus from a metastatic tumor. *Virology* 181:424–429. [https://doi.org/10.1016/0042-6822\(91\)90519-H](https://doi.org/10.1016/0042-6822(91)90519-H).