

RESEARCH

Open Access



Characterization of a novel papillomavirus identified from a whale (*Delphinapterus leucas*) pharyngeal metagenomic library

Xiang Lu^{1†}, Rong Zhu^{2†} and Ziyuan Dai^{2*}

Abstract

Here, using viral metagenomic method, a novel whale papillomavirus (temporarily named wPV, GenBank accession number OP856597) was discovered in a whale (*Delphinapterus leucas*) pharyngeal metagenomic library. The complete genome size of wPV is 7179 bp, with GC content of 54.4% and a nucleotide composition of 23.4% A, 22.3% T, 28.4% G, and 25.9% C. The viral genome has a typical papillomavirus organization pattern, and five ORFs were predicted, including two late genes encoding L1 and L2, and three early genes encoding E1, E2, and E6. Pairwise sequence comparison and phylogenetic analysis based on the L1 gene sequence indicated that wPV may be a novel species within genus *Dyodeltapapillomavirus*. In addition, the E2 region of wPV was predicted to have a potential recombination event. The discovery of this novel papillomavirus increases our understanding of the viral ecology of marine mammals, providing insights into possible future infectious diseases.

Keywords Whale, Virome, Papillomavirus, Gene recombination

Introduction

Due to the rapid development and wide application of sequencing technology in recent years, more and more novel viruses have been continuously identified, and the host species population of known viruses has also increased rapidly. Members of the family *Papillomaviridae* are non-enveloped, small, circular viruses with double-stranded (ds) DNA genomes of approximately 5.7 to 8.6 kilobases (kb) in size[1]. Papillomaviruses (PVs) are capable of infecting the skin and mucosal surfaces of mammals and many other vertebrates in a host-specific

manner and can cause papilloma or fibromyoidoma in humans, animals, and birds[2, 3]. However, there are several studies that have identified PVs in healthy skin[4] and mucous membranes[5]. At present, numerous identified non-human papillomaviruses (NHPs) have been identified in a wide range of animals including cattle[6], dogs[7], pigs[8], domestic cats[9], manatees[10], giant pandas[11], and cetaceans[12–14], and they are commonly reported to be associated with mucosal and cutaneous lesions.

PVs contain early genes (E1-E8) encoding nonstructural proteins and late genes (L1 and L2) encoding structural proteins, and the open reading frames (ORFs) arrangement of viruses in different animal species is different[15]. Among them, E1, E2 and E4 are regulatory genes involved in transcription and replication, E5, E6 and E7 are potential oncogenes, L1 and L2 are two genes encoding self-assembling proteins that produce viral capsids[16, 17]. There are also some reports that early

[†]Xiang Lu and Rong Zhu contributed equally.

*Correspondence:

Ziyuan Dai
michelle_dai99@126.com

¹Department of Microbiology, School of Medicine, Jiangsu University, Zhenjiang, China

²Department of Clinical Laboratory, Affiliated Hospital 6 of Nantong University, Yancheng Third People's Hospital, Yancheng, Jiangsu, China



genes such as E5 and E8 are involved in oncogenesis[18]. According to the International Committee for the Taxonomy of Viruses (ICTV), the family *Papillomaviridae* currently consists of 133 species grouped into 53 genera. Classification of papillomavirus types is now based on the nucleotide sequence of the L1 gene[19]. Members of the same genus of PVs share more than 60% nucleotide sequence identity in the L1 ORF. PV types within a species share 71–89% nucleotide identity in the complete L1 ORF. Sequences that share 90–98% nucleotide identity represent different subtypes within the same type, while sequences that share 98–100% nucleotide identity represent different variants within the same subtype[19–21].

Here, we report the complete genome characterization of a novel whale papillomavirus (temporarily named wPV) identified from a whale (*Delphinapterus leucas*) pharyngeal metagenomic library. The advent of this novel full-length genome adds to our understanding of the diversity of marine mammalian PVs. Our analysis suggests that wPV could be putatively a novel species in the genus *Dyodeltapapillomavirus*.

Materials and methods

Metagenome assembly

During the investigation of potential pathogenic viruses in mammals, one available library was downloaded from the SRA database, SRR12366696, uploaded by Du et al.[22], Hainan Medical University, corresponding to the host whale. The sample processing method has been described in the previous[22]. Pfastq-dump v0.1.6 (<https://github.com/inutano/pfastq-dump>) was used to convert SRA format file to fastq format file, and Bowtie2 v2.4.5[23, 24] was used to remove host sequences. Potential primer sequences of raw reads were trimmed using Trim Galore v0.6.5 (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore) and the resulting files were quality controlled with options ‘--phred33 --length 35 --stringency 3 --fastqc’. Duplicated reads were marked using PRINSEQ-lite v0.20.4 (-derep 1)[25]. This library was assembled in-house pipeline. Single-end reads were assembled with MEGAHIT v1.2.9[26] with default parameters. The results were then imported into Geneious Prime v2022.0.1[27] for sorting and manual confirmation. To reduce false negatives during sequence assembly, further semi-automatic assembly of those contigs or singlets with sequence length < 500 bp was performed and contigs with sequence length > 1500 bp after reassembly were retained, where the individual contig was used as reference for mapping to the raw data using the Low Sensitivity/Fastest parameter in Geneious Prime. In addition, mixed assembly was performed using MEGAHIT in combination with BWA v0.7.17[28] to search for unused reads for possible low abundance contigs.

Searching for viruses in the whale library

The contigs were aligned with the non-redundant protein (nr) database using blastx program built in DIAMOND v2.0.15[29], with a cut-off E-value of $< 10^{-5}$. In addition, the sequences of proteins such as RdRp (RNA-directed RNA polymerase), Rep (replication-associated protein) and NS1 (non-structural protein) were also downloaded from the RefSeq database to align contigs with sequence length > 1500 bp. The rma2info program built into MEGAN6[30] was used to perform taxonomic identification. Putative open reading frames (ORFs) were predicted by Geneious Prime with built-in parameters (Minimum size: 100)[27], and were further checked through comparing to related viruses. The annotations of these ORFs were based on comparisons to the Conserved Domain Database (CDD).

Phylogenetic analysis

To infer phylogenetic relationships, nucleotide and protein sequences of reference strains belonging to different group of corresponding viruses were downloaded from the NCBI GenBank database. Related nucleotide and protein sequences were aligned using alignment program implemented in the CLC Genomics Workbench 10.0 (<https://digitalinsights.qiagen.com>), and the resulting alignment was further optimized using MUSCLE in MEGA v7.0[31] and MAFFT v7.3.1 employing the E-INS-I algorithm[32]. Sites containing more than 50% gaps were temporarily removed from alignments. Bayesian inference trees were then constructed using MrBayes v3.2[33]. The Markov chain was run for a maximum of 1 million generations, in which every 50 generations were sampled and the first 25% of Markov chain Monte Carlo (mcmc) samples were discarded as burn-in. Maximum Likelihood trees were also constructed to confirm all the Bayesian inference trees using software MEGA v7.0[31]. Colorcoded distance matrix analysis between novel papillomavirus and other members of *Papillomaviridae* were performed with Sequence Demarcation Tool v1.2[34].

Prediction of potential genome recombination events

Genomic alignments of reference strains and the wPV strain were analyzed using the algorithms (RDP, GENECONV, Chimaera, MaxChi, BootScan, SiScan) of the Recombination Detection Program v4.39 (RDP4) software to screen for potential recombination events[35].

Prediction of spatial structure

ColabFold[36] was used to predict the three-dimensional structure of the viral structural protein identified in this study, and SWISS-MODEL[37] was used to compare and screen models with similar spatial structures from the PDB database. PyMOL v2.0 (www.pymol.org) was used for visualization.

Data availability

The novel papillomavirus sequence obtained in this study have been deposited in GenBank database under accession number OP856597.

Results

Viral metagenomic overview

This library generated a total of 9,167,717 raw reads on the Illumina HiSeq platform, the number of clean reads obtained after quality control is 9,166,522. After queried of the clean reads against the nr database, a total of 581,875 reads had the best matches with viral proteins (Supplementary Table 1), accounting for 6.35% of the total clean reads. About 17 viral families were detected, the most abundant viral family was *Siphoviridae* (68.27% of the total clean viral reads), followed by *Picobirnaviridae* (13.74%), *Myoviridae* (10.64%), *Papillomaviridae* (3.31%), *Podoviridae* (1.34%), *Inoviridae* (1.04%), *Ackermannviridae* (0.52%), and *Mimiviridae* (0.30%) (Supplementary Table 1). In addition, there are 20,152 reads that have not been clearly assigned to viruses at any taxonomic level ($E\text{-value} > 10^{-5}$) and potentially novel viruses may be present in these reads.

Identification of a novel papillomavirus

In this study, one circular complete genome (temporarily named wPV) related to *Papillomaviridae* was obtained using the assemble sequences program in Geneious Prime v2022.0.1. The complete genome size of wPV is 7179 bp with GC content of 54.4% and a nucleotide composition of 23.4% A, 22.3% T, 28.4% G, and 25.9%

C. The characteristic organizational pattern of this complete viral genome is shown in Fig. 1. Five ORFs were well predicted, including two late genes encoding L1 and L2, and three early genes encoding E1, E2 and E6. One ATP-dependent helicase motif (GPPDTGKS; aa 445–452) and five cyclin A interaction motifs (RXL; aa 62–64, aa 108–110, aa 247–249, aa 258–260, aa 505–507) were predicted in the E1 gene. The long control region (LCR) is 569 bp long and is located between the L1 gene and the E6 gene of the circular genome (nt 6611–7179). In LCR region, there are one E1 binding site (E1BS; ATNGTTN₃AACNAT; nt 7084–7098), three E2 binding sites (E2BS; ACCN₆GGT; nt 6927–6938, nt 7046–7057, nt 7125–7136) and two nuclear factor 1 binding sites (NF1BS; TTGGC; nt 6726–6730, nt 6886–6890). In the protein encoded by the E6 gene, there are two conserved metal-binding domains (CX₂C-X₂₉-CX₂C) separated by 36 aa (aa 28–64 and aa 101–137). Furthermore, there is a PDZ-binding motif (XS/TXV/L; YSDL; aa 182–185)[21].

Phylogenetic analysis

Representative sequences of the current 53 genera of the family *Papillomaviridae* were included in the subsequent analysis, phylogenetic analysis based on the nucleotide sequence of the L1 gene revealed that wPV clustered in a sister clade with *Sus scrofa* papillomavirus type 1 (NC_011280) belonging to the genus *Dyodeltapapillomavirus* (Fig. 2). Currently, the genus *Dyodeltapapillomavirus* contains only one species, *Dyodeltapapillomavirus 1*.

Distance matrix analysis of the L1 gene showed that wPV shared <70% identity with the nucleotide sequence of *Sus scrofa* papillomavirus type 1 (Fig. 3 and Supplementary Table 2). According to Taxonomy Guidelines of THE PAPILOMAVIRUS EPISTEME (PaVE) (https://pave.niaid.nih.gov/explore/taxonomy/taxonomy_concept), wPV should be considered a novel species belonging to the genus *Dyodeltapapillomavirus*.

Recombination analysis

Reference sequences belonging to the family *Papillomaviridae* closest to wPV were screened using blastx program. Using the suite of six recombination detection methods implemented in RDP4, four algorithms (RDP, Chimaera, BootScan, SiScan) supported the detection of recombination signal in the E2 region (wPV; nt 2767–2938), with P values of $1.871 \cdot 10^{-3}$, $2.678 \cdot 10^{-2}$, $4.397 \cdot 10^{-2}$, $1.253 \cdot 10^{-4}$. The major parent of the wPV was the *Sus scrofa* papillomavirus 1 (NC_011280; *Dyodeltapapillomavirus*), the minor parent was the *Phocoena phocoena* papillomavirus 4 (NC_018076; *Dyopipapillomavirus*) (Fig. 4). These two possible parents are both closely related to wPV in the above phylogenetic analysis (Fig. 2).

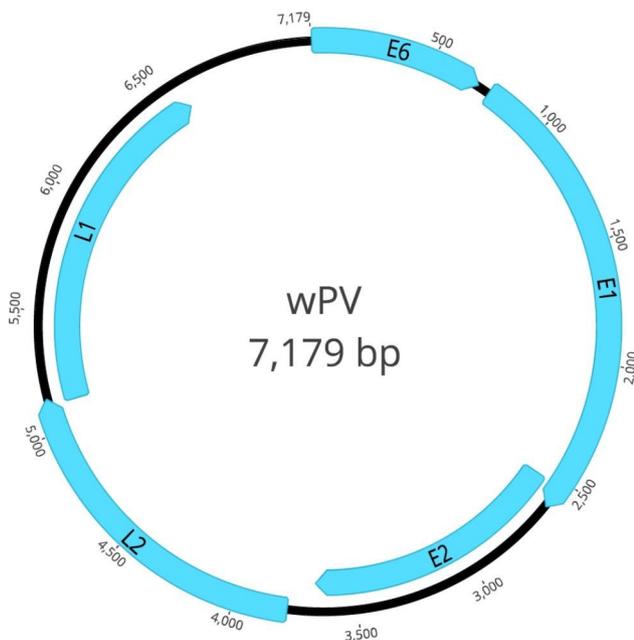


Fig. 1 Genomic organization of wPV identified in a whale pharyngeal metagenomic library

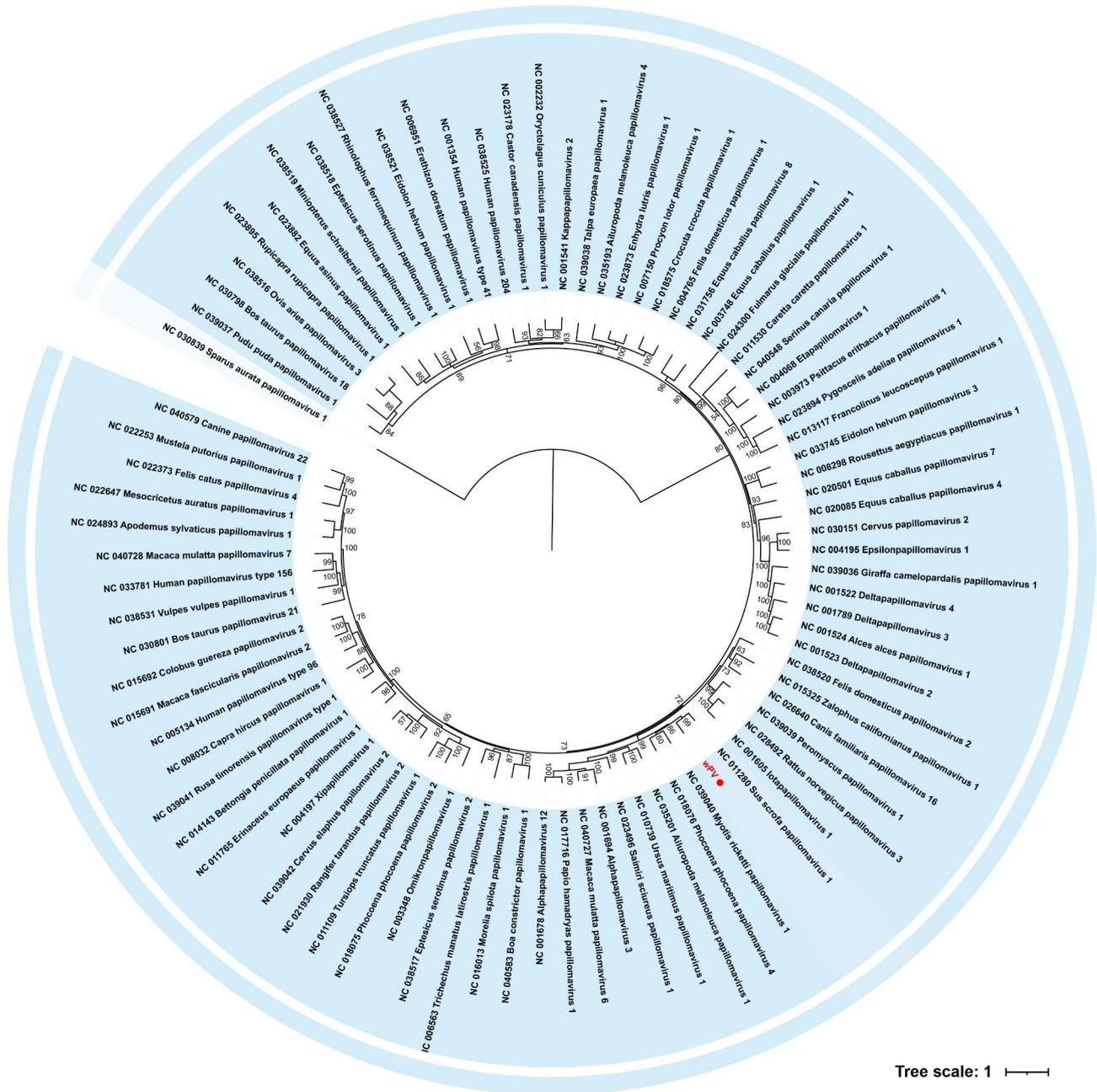


Fig. 2 Phylogenetic analysis of wPV. Bayesian inference tree based on nucleotide sequence of the L1 region of virus belonging to *Papillomaviridae* identified here. Within tree the virus found in this study is marked with red

Prediction of spatial structure of wPV structural protein L1

The major capsid protein L1 of papillomavirus is a protein of about 55kD that has the ability to spontaneously self-assemble into virus-like particles (VLP). In order to predict and compare the similarity between the spatial structure of wPV structural protein L1 and the structure encoded by the currently known sequences, the sequence with the highest degree of identity to wPV structural protein L1, *Sus scrofa papillomavirus type 1* (NC_011280), was downloaded from the GenBank database. ColabFold

was used to predict the spatial structure of sequences encoding structural protein L1. In addition, the most similar virus model (Bovine papillomavirus type 1; 3iyj.1.A) to wPV was searched and downloaded in SWISS-MODEL. All PDB files were imported into PyMOL software and do pairwise comparisons, an RMSD of less than about 2 Å would generally be considered very close. In this research, wPV has good similarity with *Sus scrofa papillomavirus type 1* and *Bovine papillomavirus type 1* in L1 spatial structure (Fig. 5). This suggests that wPV

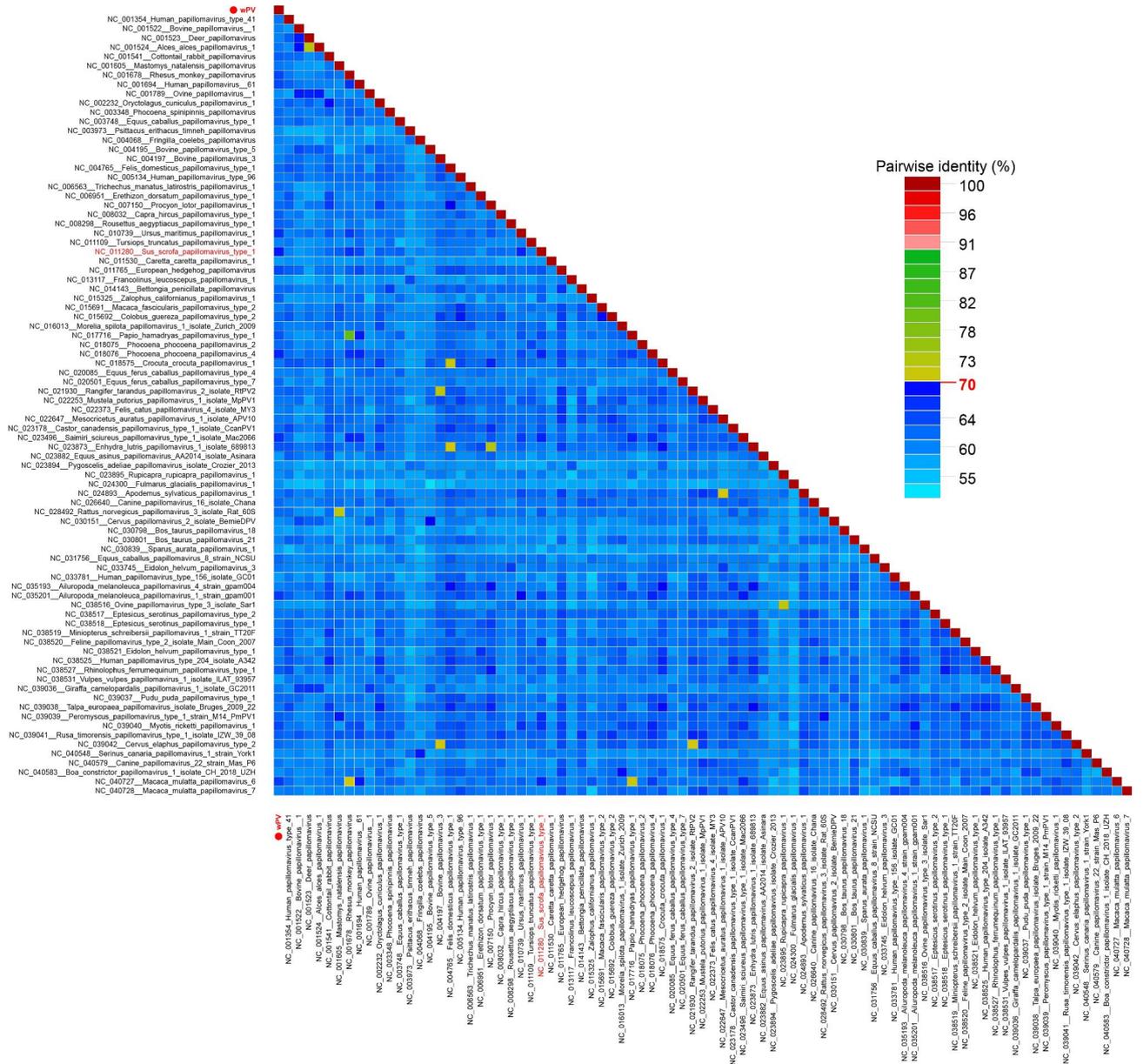


Fig. 3 Distance matrix analysis of wPV. Pairwise sequence comparison produced with L1 nucleotide sequences within Bayesian consensus tree

and these viruses may share similar properties in mediating cell attachment during infectious entry.

Discussion

In recent years, an increasing number of novel viruses are being discovered and characterized in a variety of settings, especially novel PVs of human and animal sources[38], largely due to the rise and advancement of next-generation sequencing, which allows researchers to easily detect viruses on the skin[39] or in vivo[40]. Some of these newly discovered viruses tend to accompany the emergence and progression of disease, but in most cases, these viruses are harmless to humans or animals. Nonetheless, coexistence of viruses with healthy humans and

animals does not completely rule out their pathogenic potential. Understanding how these newly identified viruses affect humans or animals will allow us to determine aggressive prevention and treatment strategies.

Members of PVs infect the skin and mucosal surfaces of diverse animal hosts[15], typically resulting in benign mucosal and skin lesions[3]. The core ORFs of papillomavirus include E1, E2, L1 and L2. Previous studies have pointed out that cetacean PVs lack E5 and E7 genes and the E6 protein is larger than other species, and the identification of wPV further supports this view[41, 42]. But we did not predict other ORFs in the wPV genome.

Members of PVs have the characteristics of double-stranded DNA genome stability, species specificity, and

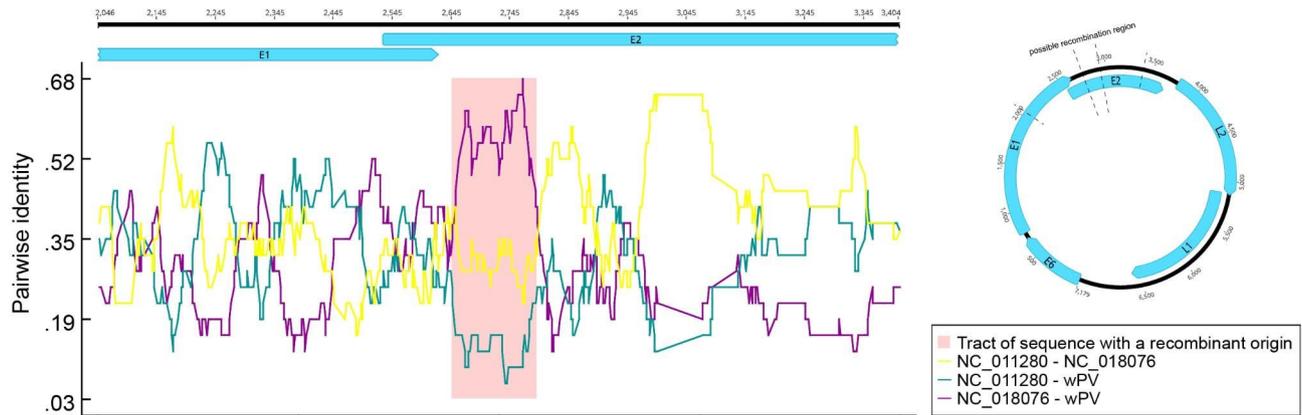
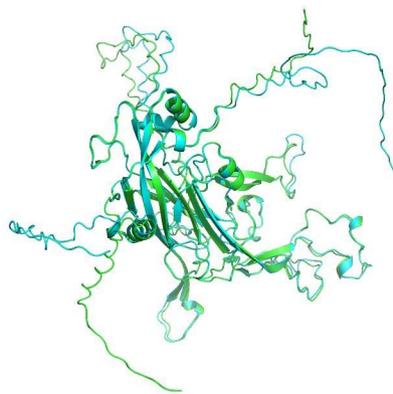


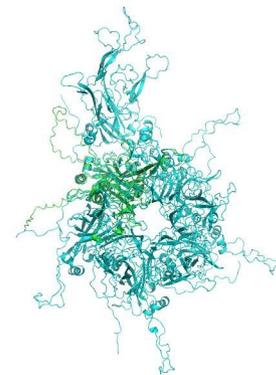
Fig. 4 Putative recombination events in the wPV. The region with a pink background represents the potential recombination region

RMSD = 0.334

RMSD = 1.160



wPV vs *Sus scrofa* papillomavirus type 1 (NC_011280)



wPV vs Bovine papillomavirus type 1 (3iyj.1.A)

Fig. 5 Viral L1 protein structural model visualization. PDB files were visualized and pairwise aligned using PyMOL v2.0 software

close skin or mucous membrane contact as the prerequisite for transmission, which makes it difficult to explain the spread of the virus between species on a global scale. It is argued that this is due to the co-existence and evolution of PVs with the early ancestors of various hosts through a host-linked co-divergence pattern. This suggests that PVs may be one of the oldest and most widespread viral families[15]. But there is also evidence that evolutionary mechanisms such as interspecific transmission, adaptive radiation, recombination, and positive selection have also played an important role in papillomavirus evolution[13]. Evidence for interspecies and intraspecies recombination of PVs has been explained[17, 43–50]. However, to date only cetacean PVs have reported conclusive evidence of specific recombination events[13, 17, 48]. Previous studies have shown that the first possible recombination event in Cetaceans is localized near the end of E2/beginning of L2, and the second possible recombination site is estimated to be within the viral LCR region[48, 51]. In this study, the possible

recombination event of wPV occurred at the beginning of the E2 protein. The two parents of wPV predicted in this study are domestic pig and Harbor porpoise. Combined with the above-mentioned characteristics of PVs infection, we cautiously speculate that there may be recombination between domestic pigs and amphibians. However, considering that the signal of recombination is not very obvious, the results of this study cannot be regarded as conclusive evidence for the existence of recombination in wPV.

In conclusion, the identification of wPV increases our knowledge of marine mammalian papillomavirus diversity and provides limited assistance in improving taxonomy. Although there is no direct evidence of the impact of wPV on the host, further dynamic monitoring of it is warranted in the future.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12985-023-02009-y>.

Supplementary Table 1. Summary of information for a whale pharyngeal metagenomic library.

Supplementary Table 2. Supplementary information for distance matrix analysis of wPV.

Acknowledgements

Not applicable.

Authors' contributions

Xiang Lu and Ziyuan Dai collected data and wrote this article, Rong Zhu and Ziyuan Dai assisted in data processing and proofreading. All authors reviewed the manuscript.

Funding

This work was supported by the Special Funds for Science Development of the Clinical Teaching Hospitals of Jiangsu Vocational College of Medicine no. 20229152.

Data Availability

The novel papillomavirus sequence obtained in this study have been deposited in GenBank database under accession number OP856597.

Declarations

Competing interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Written informed consent for publication was obtained from all participants.

Received: 24 November 2022 / Accepted: 13 March 2023

Published online: 20 March 2023

References

1. Van Doorslaer K, Chen Z, Bernard HU, Chan PKS, DeSalle R, Dillner J, Forslund O, Haga T, McBride AA, Villa LL, et al. ICTV Virus Taxonomy Profile: Papillomaviridae. *J Gen Virol*. 2018;99:989–90.
2. McBride AA. Oncogenic human papillomaviruses. *Philos Trans R Soc Lond B Biol Sci* 2017, 372.
3. Sobhy NM, Singh V, El Damaty HM, Mor SK, Youssef CRB, Goyal SM. Molecular characterization of a novel *Camelus dromedarius* papillomavirus. *Comp Immunol Microbiol Infect Dis*. 2020;73:101561.
4. Antonsson A, Erfurt C, Hazard K, Holmgren V, Simon M, Kataoka A, Hossain S, Hakangard C, Hansson BG. Prevalence and type spectrum of human papillomaviruses in healthy skin samples collected in three continents. *J Gen Virol*. 2003;84:1881–6.
5. Liu Z, Yang S, Wang Y, Shen Q, Yang Y, Deng X, Zhang W, Delwart E. Identification of a novel human papillomavirus by metagenomic analysis of vaginal swab samples from pregnant women. *Virol J*. 2016;13:122.
6. Yildirim Y, Kale M, Ozmen O, Cagircan AA, Hasircioglu S, Kucuk A, Usta A, Sokel S. Phylogenetic analysis and searching bovine papillomaviruses in teat papillomatosis cases in cattle by performing histopathology, immunohistochemistry, and transmission electron microscopy. *Microb Pathog*. 2022;170:105713.
7. Munday JS, Gedye K, Knox MA, Ravens P, Lin X. Genomic Characterisation of *Canis Familiaris* Papillomavirus Type 24, a Novel Papillomavirus Associated with Extensive Pigmented Plaque Formation in a Pug Dog. *Viruses* 2022, 14.
8. Link EK, Hoferer M, Strobel B, Rigbers K, Langenmayer MC, Sutter G, Fux R. *Sus scrofa* papillomavirus 2 - genetic characterization of a novel suid papillomavirus from wild boar in Germany. *J Gen Virol*. 2017;98:2113–7.
9. Carrai M, Van Brussel K, Shi M, Li CX, Chang WS, Munday JS, Voss K, McLuckie A, Taylor D, Laws A et al. Identification of A Novel Papillomavirus Associated with Squamous Cell Carcinoma in A Domestic Cat. *Viruses* 2020, 12.
10. Bossart GD, Ewing RY, Lowe M, Sweat M, Decker SJ, Walsh CJ, Ghim SJ, Jensen AB. Viral papillomatosis in Florida manatees (*Trichechus manatus latirostris*). *Exp Mol Pathol*. 2002;72:37–48.
11. Zhang W, Yang S, Shan T, Hou R, Liu Z, Li W, Guo L, Wang Y, Chen P, Wang X, et al. Virome comparisons in wild-diseased and healthy captive giant pandas. *Microbiome*. 2017;5:90.
12. Smeele ZE, Burns JM, Van Doorslaer K, Fontenele RS, Waits K, Stainton D, Shero MR, Beltran RS, Kirkham AL, Bergnatt R, et al. Diverse papillomaviruses identified in Weddell seals. *J Gen Virol*. 2018;99:549–57.
13. Robles-Sikisaka R, Rivera R, Nollens HH, St Leger J, Durden WN, Stolen M, Burchell J, Wellehan JF Jr. Evidence of recombination and positive selection in cetacean papillomaviruses. *Virology*. 2012;427:189–97.
14. De Guise S, Lagace A, Beland P. Gastric papillomas in eight St. Lawrence beluga whales (*Delphinapterus leucas*). *J Vet Diagn Invest*. 1994;6:385–8.
15. Rector A, Van Ranst M. Animal papillomaviruses. *Virology*. 2013;445:213–23.
16. Munger K, Howley PM. Human papillomavirus immortalization and transformation functions. *Virus Res*. 2002;89:213–28.
17. Gottschling M, Bravo IG, Schulz E, Bracho MA, Deaville R, Jepson PD, Van Bresse MF, Stockfleth E, Nindl I. Modular organizations of novel cetacean papillomaviruses. *Mol Phylogenet Evol*. 2011;59:34–42.
18. Han R, Cladel NM, Reed CA, Christensen ND. Characterization of transformation function of cottontail rabbit papillomavirus E5 and E8 genes. *Virology*. 1998;251:253–63.
19. Bernard HU, Burk RD, Chen Z, van Doorslaer K, zur Hausen H, de Villiers EM. Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology*. 2010;401:70–9.
20. de Villiers EM, Fauquet C, Broker TR, Bernard HU, zur Hausen H. Classification of papillomaviruses. *Virology*. 2004;324:17–27.
21. Cortes-Hinojosa G, Subramaniam K, Wellehan JFX Jr, Ng TFF, Delwart E, McCulloch SD, Goldstein JD, Schaefer AM, Fair PA, Reif JS, et al. Genomic sequencing of a virus representing a novel type within the species *Dyopipapillomavirus 1* in an indian river lagoon bottlenose dolphin. *Arch Virol*. 2019;164:767–74.
22. Zhang P, Su H, Peng R, Chan JF, Bai S, Wang G, Huang Y, Hu X, Luo J, Liu S, et al. Identification of a Novel Astrovirus in Pinnipeds. *Front Microbiol*. 2022;13:845601.
23. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9:357–9.
24. Langmead B, Wilks C, Antonescu V, Charles R. Scaling read aligners to hundreds of threads on general-purpose processors. *Bioinformatics*. 2019;35:421–32.
25. Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics*. 2011;27:863–4.
26. Li D, Luo R, Liu CM, Leung CM, Ting HF, Sadakane K, Yamashita H, Lam TW. MEGAHIT v1.0: a fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods*. 2016;102:3–11.
27. Kearsley M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 2012;28:1647–9.
28. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2010;26:589–95.
29. Buchfink B, Reuter K, Drost HG. Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat Methods*. 2021;18:366–8.
30. Gautam A, Felderhoff H, Bagci C, Huson DH. Using AnnoTree to get more assignments, faster, in DIAMOND + MEGAN Microbiome Analysis. *mSystems*. 2022;7:e0140821.
31. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for bigger datasets. *Mol Biol Evol*. 2016;33:1870–4.
32. Kuraku S, Zmasek CM, Nishimura O, Katoh K. aLeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity. *Nucleic Acids Res*. 2013;41:W22–28.
33. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Hohns S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. *Syst Biol*. 2012;61:539–42.

34. Muhire BM, Varsani A, Martin DP. SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. *PLoS ONE*. 2014;9:e108277.
35. Martin DP, Murrell B, Golden M, Khoosal A, Muhire B. RDP4: detection and analysis of recombination patterns in virus genomes. *Virus Evol*. 2015;1:vev003.
36. Mirdita M, Schutze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. Colab-Fold: making protein folding accessible to all. *Nat Methods*. 2022;19:679–82.
37. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res*. 2018;46:W296–W303.
38. Varsani A, Kraberger S, Jennings S, Porzig EL, Julian L, Massaro M, Pollard A, Ballard G, Ainley DG. A novel papillomavirus in Adelie penguin (*Pygoscelis adeliae*) faeces sampled at the Cape Crozier colony, Antarctica. *J Gen Virol*. 2014;95:1352–65.
39. Truchado DA, Williams RAJ, Benitez L. Natural history of avian papillomaviruses. *Virus Res*. 2018;252:58–67.
40. Liu P, Qiu Y, Xing C, Zhou JH, Yang WH, Wang Q, Li JY, Han X, Zhang YZ, Ge XY. Detection and genome characterization of two novel papillomaviruses and a novel polyomavirus in tree shrew (*Tupaia belangeri chinensis*) in China. *Virology*. 2019;16:35.
41. Rehtanz M, Ghim SJ, Rector A, Van Ranst M, Fair PA, Bossart GD, Jensen AB. Isolation and characterization of the first american bottlenose dolphin papillomavirus: *Tursiops truncatus* papillomavirus type 2. *J Gen Virol*. 2006;87:3559–65.
42. Van Bresseem MF, Cassonnet P, Rector A, Desaintes C, Van Waerebeek K, Alfaro-Shigueto J, Van Ranst M, Orth G. Genital warts in Burmeister's porpoises: characterization of *Phocoena spinipinnis* papillomavirus type 1 (PsPV-1) and evidence for a second, distantly related PsPV. *J Gen Virol*. 2007;88:1928–33.
43. Angulo M, Carvajal-Rodriguez A. Evidence of recombination within human alpha-papillomavirus. *Virology*. 2007;4:33.
44. Bravo IG, Alonso A. Phylogeny and evolution of papillomaviruses based on the E1 and E2 proteins. *Virus Genes*. 2007;34:249–62.
45. Carvajal-Rodriguez A. Detecting recombination and diversifying selection in human alpha-papillomavirus. *Infect Genet Evol*. 2008;8:689–92.
46. Garcia-Vallve S, Alonso A, Bravo IG. Papillomaviruses: different genes have different histories. *Trends Microbiol*. 2005;13:514–21.
47. Narechania A, Chen Z, DeSalle R, Burk RD. Phylogenetic incongruence among oncogenic genital alpha human papillomaviruses. *J Virol*. 2005;79:15503–10.
48. Rector A, Stevens H, Lacave G, Lemey P, Mostmans S, Salbany A, Vos M, Van Doorslaer K, Ghim SJ, Rehtanz M, et al. Genomic characterization of novel dolphin papillomaviruses provides indications for recombination within the Papillomaviridae. *Virology*. 2008;378:151–61.
49. Shah SD, Doorbar J, Goldstein RA. Analysis of host-parasite incongruence in papillomavirus evolution using importance sampling. *Mol Biol Evol*. 2010;27:1301–14.
50. Varsani A, van der Walt E, Heath L, Rybicki EP, Williamson AL, Martin DP. Evidence of ancient papillomavirus recombination. *J Gen Virol*. 2006;87:2527–31.
51. Van Doorslaer K. Evolution of the papillomaviridae. *Virology*. 2013;445:11–20.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.