



## OPEN ACCESS

## EDITED BY

Albert Van Geelen,  
USDA-APHIS Center for Veterinary Biologics,  
United States

## REVIEWED BY

Yanpeng Li,  
Fudan University, China  
Mohamad S. Hakim,  
Gadjah Mada University, Indonesia

## \*CORRESPONDENCE

Wen Zhang  
✉ zhangwen@ujts.edu.cn

†These authors have contributed equally to this work

RECEIVED 01 July 2024

ACCEPTED 10 December 2024

PUBLISHED 03 January 2025

## CITATION

Zhao Y, Zhao M and Zhang W (2025)  
Identification of a novel papillomavirus in oral  
swabs from giant pandas (*Ailuropoda  
melanoleuca*). *Front. Vet. Sci.* 11:1457471.  
doi: 10.3389/fvets.2024.1457471

## COPYRIGHT

© 2025 Zhao, Zhao and Zhang. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

# Identification of a novel papillomavirus in oral swabs from giant pandas (*Ailuropoda melanoleuca*)

Yi Zhao<sup>†</sup>, Min Zhao<sup>†</sup> and Wen Zhang<sup>\*</sup>

Department of Laboratory Medicine, School of Medicine, Jiangsu University, Zhenjiang, Jiangsu, China

To fully characterize papillomavirus diversity in giant pandas (*Ailuropoda melanoleuca*), we identified a novel papillomavirus (named AmPV5, GenBank accession number MZ357114) in oral swabs from giant pandas with the help of viral metagenomics technology in this study. The complete circular genome of AmPV5 is 7,935 bp in length, with a GC content of 39.1%. It encodes five early genes (E1, E2, E4, E6, and E7), two late genes (L1 and L2), and features conserved zinc-binding domains (CXXC- (X)28/29-CXXC) in E6 and E7 genes. E7 protein has an LxCxE domain (pRB binding) in its N-terminal region. The nucleotide sequence of AmPV5 L1 gene shares <70% identity with other related sequences available in the GenBank database. Phylogenetic analysis indicated that AmPV5 fell within the *Lambdapapillomavirus* genus but formed a monophyletic branch away from other papillomaviruses found in *Ailuropoda melanoleuca*, *Canis*, *Felis catus*, *Panthera uncia*, *Enhydra lutris*, and *Procyon lotor*. According to the International Committee on Taxonomy of Viruses (ICTV) classification guidelines, AmPV5 is classified as a new species within the *Lambdapapillomavirus* genus. The discovery provides valuable insights into the viral diversity in giant pandas and highlights the need for continued surveillance of wildlife pathogens. Future studies should explore the potential role of AmPV5 in the health and disease ecology of this endangered species.

## KEYWORDS

papillomavirus, giant panda, phylogenetic analysis, virus genome, viral metagenomics

## 1 Introduction

With the advancement of viral metagenomics, an increasing number of complete genomes of papillomaviruses (PVs) have been identified in a wide range of organisms (1–3). PVs can infect the mucous membranes and squamous epithelium of vertebrates including fish, birds, reptiles, mammals, and even humans, often causing subclinical lesions and benign proliferations (4–6). In severe cases, some of them have potential to transform into malignant tumors, making PVs a continuing concern for both human and animal health even though it is highly specific (7). Various studies have established strong links between PV infections and several types of cancers, such as cervical cancer in humans, squamous cell carcinomas in cats, and papillomas in cows and horses (8–10). Furthermore, it has been shown that PVs have evolved multiple mechanisms to suppress the interferon response and evade host innate immunity effectively, delaying adaptive immune activation and enabling persistent infections that may last months or even years (11, 12). Given the limited genetic diversity of giant pandas and their potential vulnerability to diseases, the role of immune evasion by PVs warrants further investigation to understand its impact on the health and survival of this endangered species.

Papillomaviruses are small, un-enveloped, circular, double-stranded DNA viruses with a genome size of ~8 kb. They contain three functional regions: an early region, a late region, and a non-coding upstream regulatory region (URR) located between L1 and E6 open reading frames (ORFs) (13, 14). The early region encodes six proteins (E1, E2, E4, E5, E6, and E7) involved in viral replication, transcriptional regulation, and cellular transformation (15). The major coat protein L1 and the minor coat protein L2 in the late region play important roles in viral assembly (16). The URR, also known as the locus control region (LCR), contains binding sites for various cellular transcription factors, including E1 and E2 binding sites that regulate transcription and are essential for viral replication (17). The International Committee on Taxonomy of Viruses (ICTV) classifies the *Papillomaviridae* family into 2 subfamilies, over 50 genera, and 130 species based on the highly conserved nucleotide sequence of L1 ORF (13, 18). Members of the same subfamily share <45% nucleotide sequence identity in their L1 ORF, <60% for distinct genera, and between 60 and 70% for different species within the same genus (13, 18).

Low rates of reproduction and susceptibility to infectious diseases are significant factors hindering the population growth of giant pandas, an endangered species (19). Although PVs have been detected in giant pandas, virological studies on this species remain limited. To date, only four complete PV genomes (AmPV1 - 4) have been identified in giant pandas, among which AmPV1 and AmPV2 belong to the *Omegapapillomavirus* genus, AmPV4 represents a new species within the *Lambdapapillomavirus* genus, and AmPV3 defines a novel genus, and all their genomes encode the early proteins (E6, E1, and E2) and the late proteins (L2 and L1), typically ranging from 7,000 to 8,000 base pairs (20). Given the importance of this species in global conservation efforts, studying PVs in giant pandas offers a unique opportunity to uncover potential viral threats to their health and contribute to broader wildlife conservation strategies. Understanding PV diversity and its implications in giant pandas could provide insights into disease management and highlight the importance of virological surveillance in endangered species. In our study, we collected nine oral swab samples from captive giant pandas in Chengdu and characterized a novel complete PV in these samples. This discovery enhances our understanding of the PV diversity affecting this endangered species and establishes a foundation for further research into related diseases.

## 2 Materials and methods

### 2.1 Sample collection and preparation

In April 2020, nine oral swab samples from nine different individual giant pandas were collected from giant pandas in Chengdu, Sichuan Province. The collection process involved using disposable sterile gloves, swabs, and tubes. All the samples were transported on dry ice and stored at  $-80^{\circ}\text{C}$ . Each swab was placed in a 1.5 mL tube with the head broken off and shaken vigorously for 5 min with 1 mL Dulbecco's Phosphate Buffered Saline (DPBS). All collection tools were sterilized, and the sampling was conducted in a clean environment. To release the virus from the suspension and remove large particles, the samples were incubated for 30 min at

$4^{\circ}\text{C}$ , followed by centrifugation for 10 min at  $4^{\circ}\text{C}$  and  $12,000 \times g$ . Finally, the supernatant from each sample was collected into new 1.5 mL tubes.

The sampling procedures were conducted with utmost care to minimize risks to the giant pandas, other associated animals, and the environment. All experiments in this study were approved by the Institutional Animal Care and Use Committee of the Chengdu Research Base of Giant Panda Breeding (Protocol No. 2019013) and the Ethics Committee of Jiangsu University (Protocol Nos. 2015011 and 2017021).

### 2.2 DNA library construction

The shaken and mixed library containing 500  $\mu\text{L}$  sample supernatant was centrifuged at  $12,000 \times g$  and  $4^{\circ}\text{C}$  for 5 min before being filtered through a  $0.45 \mu\text{m}$  pore-size filter (Millipore) to remove eukaryotic cells and cell-sized particles (21). Free nucleic acids not protected by viral capsid proteins were degraded using an RNase/DNase digestive enzyme system at  $37^{\circ}\text{C}$  for 60 min (22). This step was performed to degrade both RNA and DNA, targeting host-derived contamination while preserving encapsidated viral nucleic acids. Total nucleic acids (RNA/DNA) were extracted using QIAamp Viral RNA Mini Kit (QIAGEN) following the instructions of the manufacturer. After extraction, the nucleic acids were subjected to reverse transcription using SuperScript III Reverse Transcriptase, and the second strand of cDNA was synthesized with the Klenow enzyme. The DNA libraries were constructed using Nextera XT DNA Sample Preparation Kit (Illumina) to further increase the concentration of viral nucleic acids. Subsequently, the constructed libraries were sequenced on the Illumina NovaSeq 6000 platform, with each sample pool generating 250 bp paired-end reads with double barcodes.

### 2.3 Bioinformatic analysis

The raw data were split and exported based on the Index combination of each library after the downstream machine. Data preprocessing was performed using Geneious Prime and involved debarcoding, trimming, and *de novo* assembling (23). The assembled contigs were analyzed using BLASTx to search for matches against a viral protein database created in our lab. A cut-off E-value of  $<10^{-5}$  was applied for significant hits. Single-stranded reads were also subjected to the same BLASTx analysis (24). These identified sequences were then compared against the in-house non-virus non-redundant (NVNR) protein database to exclude potential false positives. Format conversion was subsequently performed in Megan, where sequence classification information was presented visually (25). Our previous study found that before the giant panda (*Ailuropoda melanoleuca*) had mild diarrhea in Chengdu Research Base of Giant Panda Breeding (April, 2020), *Papillomaviridae* were observed in the highest abundance of all the detected viral reads at the family level (26). Here, to fully characterize papillomavirus diversity, contigs belonging to the *Papillomaviridae* family were selected and assembled using Geneious Prime. Open reading

frames (ORFs) and genome annotations were also predicted in Geneious Prime based on the comparisons to the Conserved Structural Domain Database (CDD) (27). SWISS-MODEL was used to predict the tertiary structure of viral proteins (28, 29). Finally, a complete papillomavirus genome sequence with low nucleotide similarity to other papillomaviruses was obtained and included in subsequent analyses.

## 2.4 Phylogenetic analysis

The amino acid sequence of the L1 protein predicted in our study was compared with related protein sequences, including those of the closest viral relatives identified through BLASTp searches in GenBank, as well as the reference sequences from different genera and species of *Papillomaviridae*. All these sequences were downloaded and aligned using MUSCLE with the E-INS-I algorithm in MEGA (30, 31). A phylogenetic tree was constructed using the “prset aamodelpr=mixed” model in MrBayes, employing the Markov chain Monte Carlo (MCMC) method (32). Additionally, a Maximum Likelihood tree was constructed to validate the Bayesian inference tree using MEGA (33).

## 2.5 Quality control

To prevent cross-contamination of samples and nucleic acid degradation, all experimental procedures were strictly conducted in accordance with the relevant guidelines. All materials in direct contact with nucleic acid samples were sterile and free of enzymes. Nucleic acid samples were dissolved in DEPC-treated water to maintain integrity, and all extraction and processing steps were performed under sterile conditions using a biological safety cabinet to ensure sample purity.

# 3 Results

## 3.1 Overview of the virome

All nine oral swab samples from giant pandas were pooled into a library (gporal003) for next-generation sequencing (NGS), generating a total of 1,063,578 raw sequence reads, among which 8.73% ( $n = 92,889$ ) reads showed similarity to known eukaryotic viruses. The remaining 91.27% ( $n = 970,689$ ) of the sequencing data aligned with sequences from eukaryotes, prokaryotes, bacteriophages, or sequences with no significant similarity to any amino acid sequences in the NR database. At the family level, almost all eukaryotic viral reads ( $n = 92,654$ ) were assigned to three families of *Papillomaviridae* ( $n = 92,185$ ), *Parvoviridae* ( $n = 240$ ), *Herpesviridae* ( $n = 229$ ) (Figure 1). Reads related to *Papillomaviridae* were overwhelmingly dominant. At the genus level, these reads were classified into 10 distinct papillomavirus genera, including *Alphapapillomavirus*, *Deltapapillomavirus*, *Dyonupapillomavirus*, *Dyoomikronpapillomavirus*, *Dyopipapillomavirus*, *Gammapapillo*

*mavirus*, *Lambdapapillomavirus*, *Omegapapillomavirus*, *Treisiota papillomavirus*, and *Upsilonpapillomavirus*.

## 3.2 Genomic characteristics

The circular genome of AmPV5 is 7,935 bp in length with a GC content of 39.1%. It contains three main regions: an early region encoding five proteins (E1, E2, E4, E6, and E7), a late region encoding two proteins (L1 and L2), and a non-coding URR region located between the E2 and L2 ORFs. Their lengths are as follows: E1 is 1,791 bp, E2 is 1,170 bp, E4 is 396 bp, E6 is 417 bp, E7 is 279 bp, L1 is 1,506 bp, and L2 is 1,545 bp. Additionally, there is a small non-coding region named NCR2 positioned between the L1 and E6 ORFs at positions from 3,589 to 4,428bp (Figure 2A). Non-coding regions like the URR and NCR2 play key roles in transcription regulation of papillomaviruses. These regions often contain promoter and enhancer elements that influence viral gene expression. They may also mediate interactions with host factors, potentially affecting viral replication and pathogenicity. Future functional studies could provide more insight into their specific roles in AmPV5. The first nucleotide was defined as the initiation codon of the E6 gene in AmPV5.

Analysis of the nucleotide and amino acid sequences revealed several classical signature sites in AmPV5 compared to other PVs. E1 and E2 genes of PVs play essential roles in viral DNA replication and gene regulation. An ATP-binding site of the ATP-dependent DNA helicase (GPPNTGKS) is found in AmPV5 E1 at position 424–431. The presence of this motif led to the speculation that this protein might exhibit ATPase activity and helicase activity similar to those of the simian virus 40 (SV40) and polyomavirus T antigens (34). The E2 gene consists of a N-terminal transactivation domain and a C-terminal DNA binding domain. The N-terminal domain is essential for regulating viral gene transcription, and influence the expression of early genes such as E6 and E7 (35). The C-terminal domain is responsible for the dimerization of E2 proteins (36). Recent studies have highlighted the critical interactions between E1 and E2, and their combined functions are crucial for maintaining the viral genome in a stable episomal form and ensuring efficient replication. There are some differences in the structure of the AmPV5 genome compared to other papillomaviruses. For instance, AmPV5 lacks the E5 gene, which in high-risk PVs like Human papillomavirus type 16 (HPV16) is associated with immune modulation and oncogenesis (37). However, compared with HPV16 and bovine PV, AmPV5 shows a longer URR area, which is located between the L1 and E6 ORFs and contains the viral origin of replication as well as binding sites for viral and cellular transcription factors. E6 ORF features four CXXC motifs, forming two conserved zinc-binding domains: CXXC-X29-CXXC and CXXC-X28-CXXC, separated by 36 amino acids (Figure 2B). A similar zinc-binding domain (CXXC-X28-CXXC) is located at the C-terminus of E7 ORF, which plays a key role in the formation of E7 dimer (38, 39). Furthermore, the E7 ORF contains a conserved motif (LXCXE), known as the core binding region for E7 and the MCV (Merkel cell polyomavirus). Large T oncoprotein targeting the retinoblastoma tumor suppressor protein

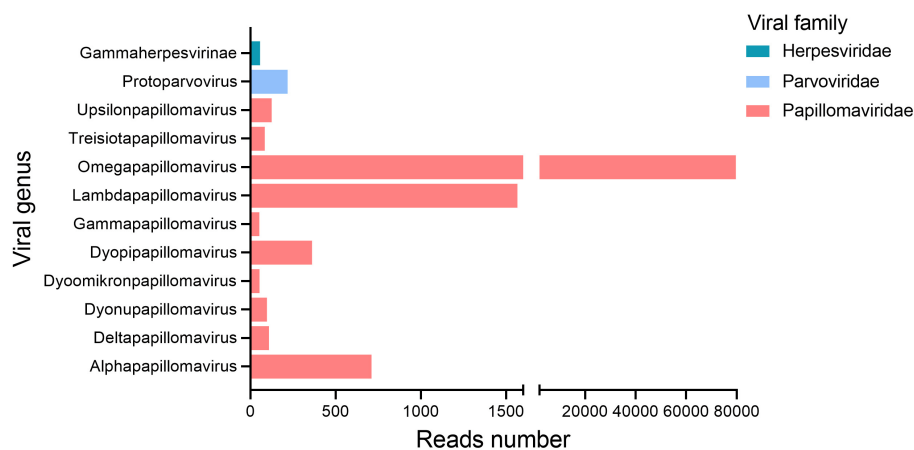


FIGURE 1  
Bar plots showing taxonomic category and viral abundance in the gporal003 library generated from nine giant panda oral swab samples.

(pRB) (40, 41). These results provide valuable insights for future studies on the interactions of AmPV5 with its host.

### 3.3 Projections of conservative structure

Given that the transformation of host cancer cells by HPV is primarily mediated by E6 and E7 genes, we conducted tertiary structure predictions of the characteristic conserved regions (CXXC and LXCXE) on AmPV5 E6 and E7 genes (Figure 3) (6, 41). Our analysis revealed that E6 protein of AmPV5 contains four CXXC conserved regions, forming two stable zinc-binding domains, whereas E7 protein has two CXXC regions, forming a single strong zinc-binding domain. Remarkably, this domain drives the formation of a unique dimeric structure, similar to what has been observed in high-risk genital HPVs involved in the inactivation of human cancer-related proteins such as histone deacetylase (42). The CXXC motif may enhance protein stability through zinc binding, while the LXCXE motif facilitates interactions with the retinoblastoma protein (pRB), disrupting cell cycle control. These features suggest AmPV5 may have enhanced host adaptation and potential oncogenic properties. Functional validation is an important next step to confirm these effects.

### 3.4 Identification of the novel species of *Lambdapapillomavirus* and phylogenetic analysis

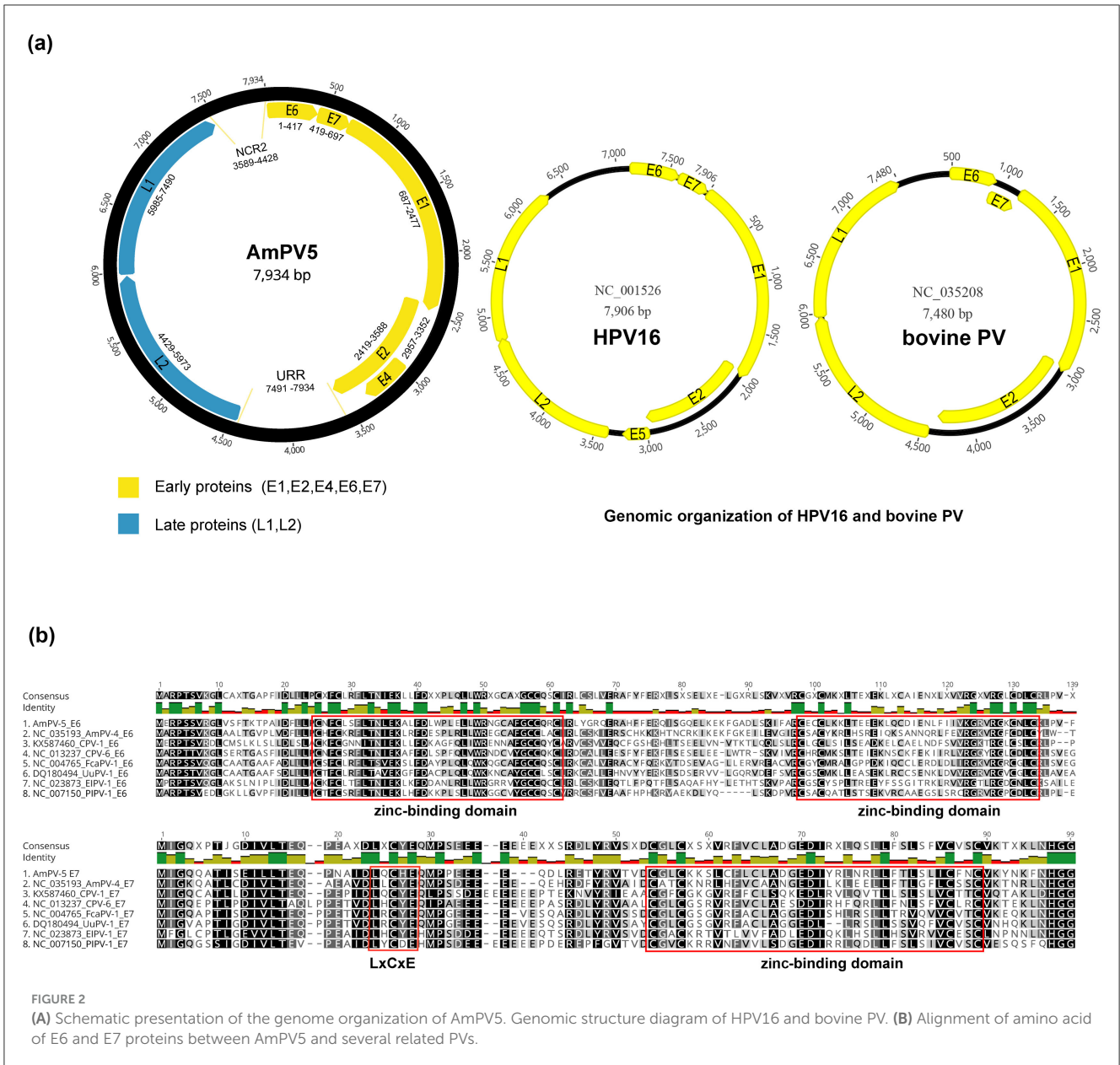
The L1 gene of AmPV5 showed 69.34% nucleotide sequence identity to that of the best-matched papillomavirus strain ElkPV-2 (GenBank accession number OQ746290) in BLASTx searching. Based on the ICTV classification criteria for PVs, AmPV5 represents a new species. A previous study has identified four complete PVs in giant pandas (AmPV1, AmPV2, AmPV3, and AmPV4) (20). Therefore, the new papillomavirus identified in this study was named as AmPV5. Multiple sequence alignment revealed

that AmPV5 shared < 70% identity in both nucleotide and amino acid sequences of L1 gene with other PVs discovered in giant pandas (Tables 1, 2). Among them, AmPV5 and AmPV4 (GenBank accession number NC\_035193) shared the highest nucleotide sequence identity of 68.79% in L1 genes and belonged to the same genus, *Lambdapapillomavirus*. Conversely, AmPV5 and the other three giant panda PVs exhibited < 60% nucleotide identity, placing them in different viral genera. Alignment of the L1 protein amino acid sequences from these five giant panda-origin PV strains confirmed these findings.

To elucidate the genetic relationship of this novel papillomavirus with others in mammals, we conducted amino acid sequence alignment of L1 proteins and constructed a Bayesian inference tree (Figure 4). The alignment involved 62 sequences, including AmPV5, top matches identified through BLASTp searches of L1 protein, and representative papillomavirus reference sequences from various genera and species in GenBank. Results revealed that AmPV5 shares the highest amino acid sequence similarity (70.38%) with the L1 protein of a papillomavirus isolated from *Enhydra lutris* (GenBank accession number NC\_023873). Phylogenetically, AmPV5 were grouped together with other 6PVs from *Ailuropoda melanoleuca*, *Canis*, *Felis catus*, *Panthera uncia*, *Enhydra lutris*, and *Procyon lotor* within the *Lambdapapillomavirus* genus, amino acid similarities of 60–70%. The Maximum-likelihood tree was constructed using MEGA supported these above findings, confirming that AmPV5 is indeed a new species of *Lambdapapillomavirus* genus.

## 4 Discussion

Many viruses have evolved to establish a homeostatic equilibrium with their hosts over extended periods, often resulting in persistent asymptomatic infections (43). However, the spread and mutation of viruses between species can significantly alter their pathogenicity and virulence, particularly in immunocompromised hosts who may develop health complications post-infection (44, 45). Our findings indicate that AmPV5 infects giant pandas without



**FIGURE 2** (A) Schematic presentation of the genome organization of AmPV5. Genomic structure diagram of HPV16 and bovine PV. (B) Alignment of amino acid of E6 and E7 proteins between AmPV5 and several related PVs.

causing concurrent disease. Conversely, Zhang et al., reported that the number of *Papillomaviridae* reads and the papillomavirus positivity rate in healthy pandas were much lower compared to those in the deceased giant pandas, suggesting a potential causative role of papillomaviruses in fatal cases (20). Furthermore, instances of papillomavirus pathogenicity in humans and other species are well-documented, and cross-species transmission events have been reported (44, 46). Therefore, there is an urgent need to employ viral metagenomics approaches to explore and characterize potential pathogenic viruses in humans, animals, and the environment, expanding our understanding of candidate pathogenic agents.

In this study, we successfully assembled and annotated a complete *Lambdapapillomavirus* in oral swabs from giant pandas, named AmPV5 MZ357114. The genome of AmPV5 is 7,935 bp in length with a GC content of 39.1%, encoding both the early region (E1, E2, E4, E6, E7) and the late region

(L1 and L2). A BLASTx search in GenBank based on the conserved L1 sequence revealed a 69.34% identity with the closest match (ElkPV2, GenBank accession number OQ746290), while the BLASTn search showed the highest amino acid similarity (70.38%) with a papillomavirus from *Enhydra lutris* (EIPV1, GenBank accession number NC\_023873). A Bayesian inference tree was constructed based on the amino acid sequences of the L1 ORF, and the viruses from *Enhydra lutris*, *Ailuropoda melanoleuca*, *Canis*, *Felis catus*, *Panthera uncia*, and *Procyon lotor* were clustered in a branch within *Lambdapapillomavirus* genus, suggesting potential interspecies transmission. Additionally, AmPV5 was classified as a new species within the genus *Lambdapapillomavirus* according to ICTV classification guidelines. Our analysis identified key regions on AmPV5 E6 and E7 proteins: CXXC and LXCXE motifs. These conserved structures are crucial in papillomavirus interactions with host physiology and

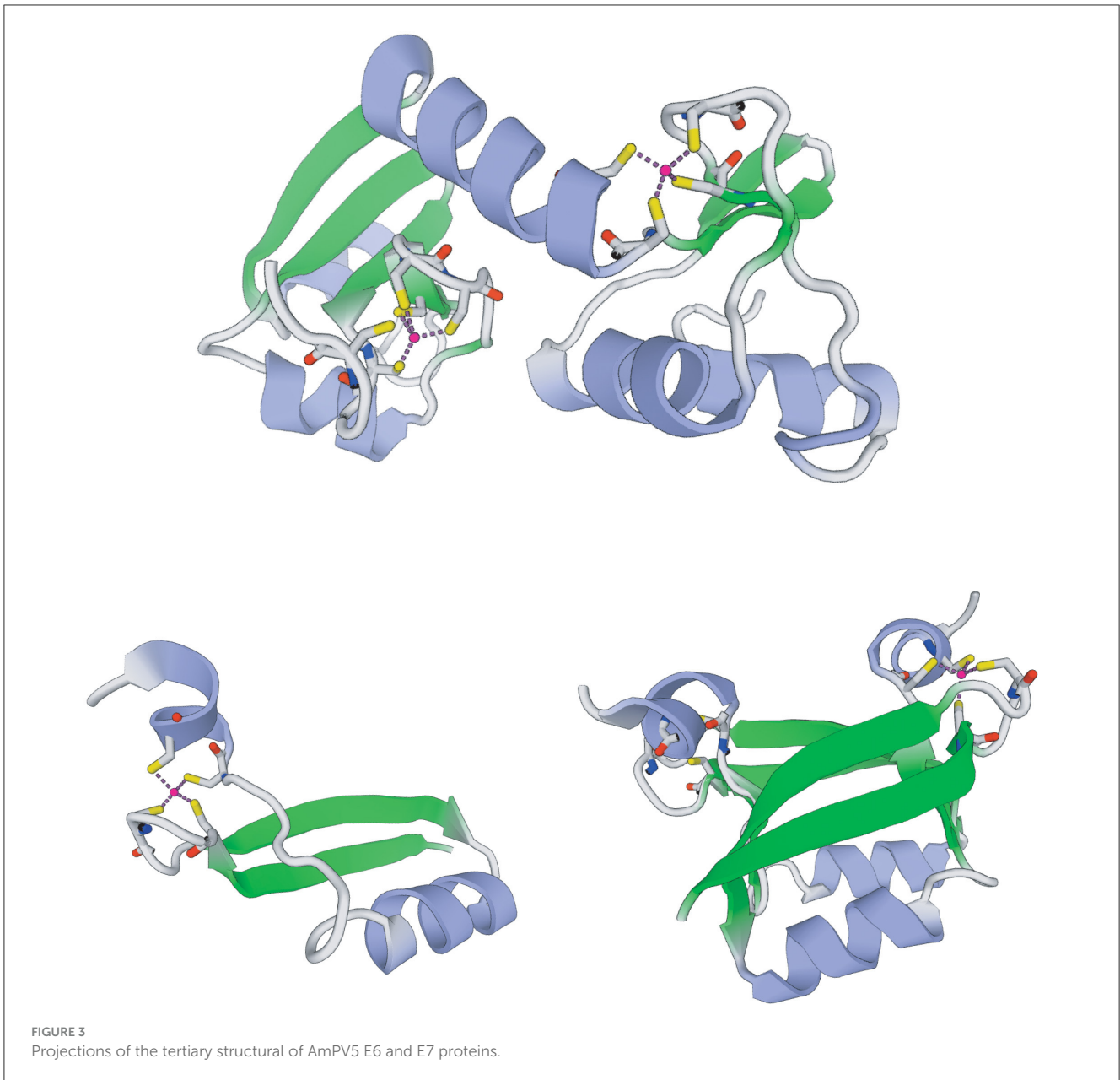


TABLE 1 Nucleotide sequence alignment of the L1 genes of all AmPVs.

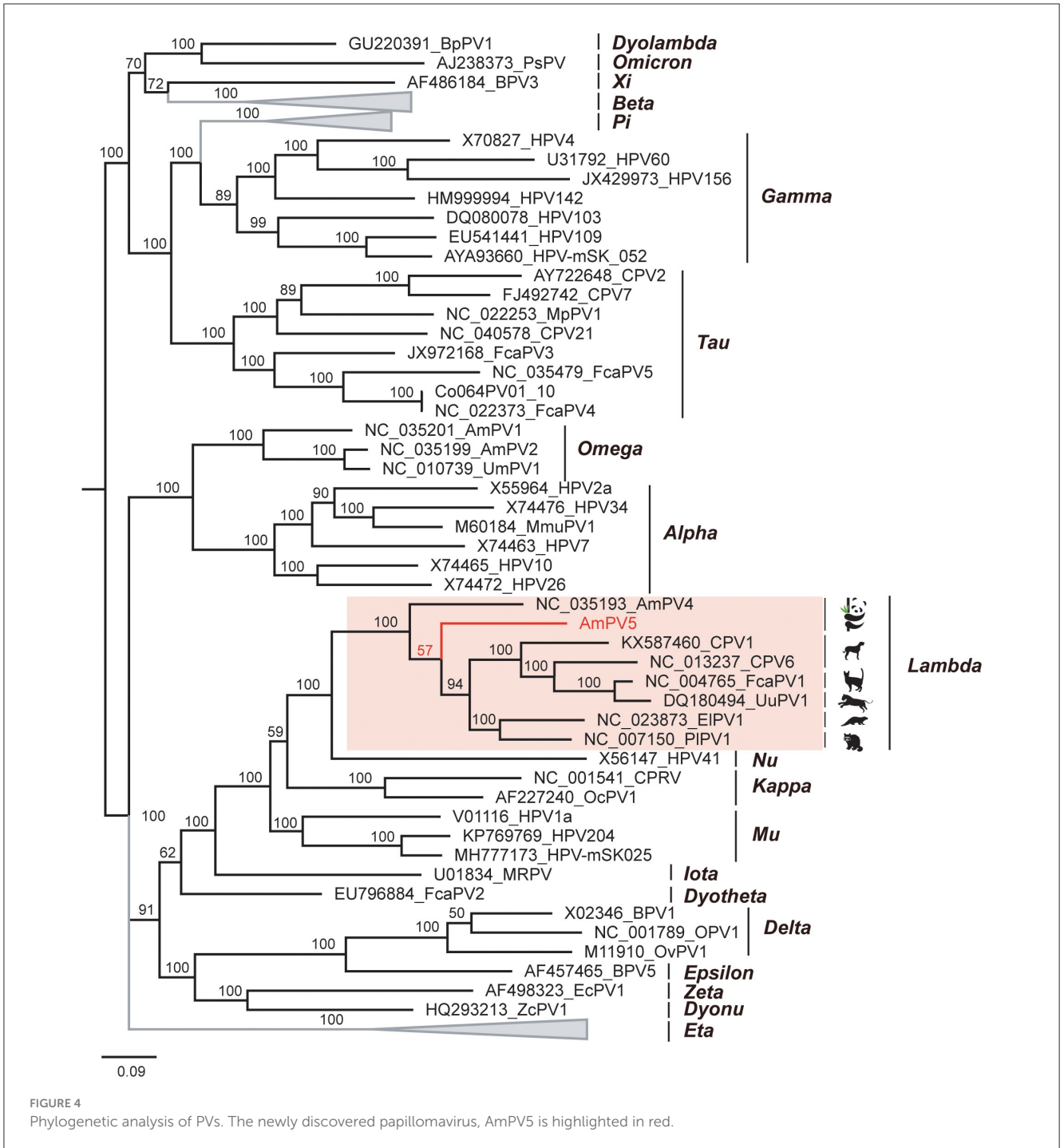
Nucleotide alignment (% identity)	AmPV5-L1	AmPV1-L1	AmPV2-L1	AmPV3-L1	AmPV4-L1
AmPV5-L1		56.35%	55.49%	52.36%	68.79%
AmPV1-L1	56.35%		68.74%	57.34%	58.33%
AmPV2-L1	55.49%	68.74%		58.10%	55.95%
AmPV3-L1	52.36%	57.34%	58.10%		54.02%
AmPV4-L1	68.79%	58.33%	55.95%	54.02%	

biochemistry, contributing to disease induction (47). Specifically, E6 has been implicated in p53 oncogene degradation by inhibiting CBP/p300 co-activator transcription, preventing apoptosis (48). E7 affects cellular transformation through conserved LXCXE motifs

and critical zinc-binding regions, contributing to cervical cancer pathogenesis (49, 50). In our study, we find that E5 is missing in AmPV5. Actually, many but not all papillomaviruses encode E5 proteins, which are short, membrane-associated proteins with

TABLE 2 L1 Amino acid sequence alignment of the L1 proteins of all AmPVs.

Protein alignment (% identity)	AmPV5-L1	AmPV1-L1	AmPV2-L1	AmPV3-L1	AmPV4-L1
AmPV5-L1		49.12%	50.68%	51.39%	66.80%
AmPV1-L1	49.12%		73.37%	58.30%	51.28%
AmPV2-L1	50.68%	73.37%		57.00%	53.15%
AmPV3-L1	51.39%	58.30%	57.00%		53.09%
AmPV4-L1	66.80%	51.28%	53.15%	53.09%	



transforming activity. In non-mammalian PVs or certain low-risk mammalian PVs, the E5 gene is frequently absent, suggesting that its functions may not be essential in all host-virus interactions (51). Notably, all currently identified PVs infecting giant pandas lack the E5 gene, which may represent a characteristic feature of PVs specific to this host species. This consistent absence may reflect evolutionary adaptations to the host environment or particular aspects of the viral life cycle in giant pandas. Further investigation into these key proteins and their binding regions in new papillomavirus species is essential to elucidate papillomavirus-host interactions.

In addition to papillomaviruses (PVs), several other viruses have been identified in giant pandas, some of which may pose potential risks to other species, including humans. For instance, canine distemper virus (CDV) has caused outbreaks in giant panda populations, raising concerns about its potential for cross-species transmission (52). Similarly, feline panleukopenia parvovirus (FPV), known to infect a range of carnivorous mammals, and certain enteric viruses, such as rotaviruses and coronaviruses, have been detected in giant pandas (53). While there are no documented cases of direct virus transmission from giant pandas to humans, the presence of zoonotic pathogens highlights the importance of biosecurity in conservation settings (54). These findings underscore the need for ongoing surveillance of viral infections in giant pandas to protect both the species and public health.

Overall, this study significantly expands our understanding of papillomavirus diversity in giant pandas, enriching the viral ontogenetic database related to these animals and their surroundings. Although no direct evidence was found linking AmPV5 to disease in giant pandas, the structural features of E6 and E7 suggest the virus may affect the host cell cycle or immune response, highlighting the need for further research. Future studies could include longitudinal studies to monitor the prevalence and transmission dynamics of AmPV5 across different panda populations, regions, and age groups, thereby assessing potential health risks. Additionally, functional validation of E6 and E7 proteins through cell culture or animal models could help determine their impact on host cellular pathways and pathogenic potential. These efforts will provide a comprehensive understanding of AmPV5's role in panda health and conservation, supporting proactive measures against emerging diseases.

## 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 in the article/supplementary material.

## Ethics statement

The animal study was approved by Committee of the Chengdu Research Base of Giant Panda Breeding protocol no. 2019013 and Ethics Committee of Jiangsu University nos. 2015011 and 2017021. The study was conducted in accordance with the local legislation and institutional requirements.

## Author contributions

YZ: Writing – original draft, Writing – review & editing. MZ: Writing – review & editing. WZ: Methodology, Resources, Writing – review & editing.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was supported by National Key Research and Development Programs of China (Nos. 2023YFD1801301 and 2022YFC2603801) and the National Natural Science Foundation of China (No. 82341106).

## 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.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

1. Sommers P, Chatterjee A, Varsani A, Trubl G. Integrating viral metagenomics into an ecological framework. *Ann Rev Virol.* (2021) 8:133–58. doi: 10.1146/annurev-virology-010421-053015
2. Truchado DA, Williams RAJ, Benitez L. Natural history of avian papillomaviruses. *Virus Res.* (2018) 252:58–67. doi: 10.1016/j.virusres.2018.05.014
3. Tse H, Tsang ALK, Tsoi H-E, Leung ASP, Ho C-C, Lau SPK, et al. Identification of a novel bat papillomavirus by metagenomics' edited

by T. Davis. *PLoS ONE.* (2012) 7:e43986. doi: 10.1371/journal.pone.0043986

4. Cubie HA. Diseases associated with human papillomavirus infection. *Virology.* (2013) 445:21–34. doi: 10.1016/j.virol.2013.06.007

5. Johannsen E, Lambert PF. Epigenetics of human papillomaviruses. *Virology.* (2013) 445:205–12. doi: 10.1016/j.virol.2013.07.016



6. Sobhy NM, Singh V, El Damaty HM, Mor SK, Youssef CRB, Goyal CM. Molecular characterization of a novel *Camelus dromedarius* papillomavirus. *Comp Immunol Microbiol Infect Dis*. (2020) 73:101561. doi: 10.1016/j.cimid.2020.101561
7. Jensen JE, Becker GL, Brooks Jackson J, Rysavy MB. Human papillomavirus and associated cancers: a review. *Viruses*. (2024) 16:680. doi: 10.3390/v16050680
8. Bzhalava D, Guan P, Franceschi S, Dillner J, Clifford G. A systematic review of the prevalence of mucosal and cutaneous human papillomavirus types. *Virology*. (2013) 445:224–31. doi: 10.1016/j.virol.2013.07.015
9. Ikechukwu CK, Qin K, Zhang H, Pan J, Zhang W. Novel equid papillomavirus from domestic donkey. *Equine Vet J*. (2024) 56:171–7. doi: 10.1111/evj.13957
10. Yamashita-Kawanishi N, Sawanobori R, Matsumiya K, Uema A, Chambers JK, Uchida K, et al. Detection of *Felis catus* papillomavirus type 3 and 4 DNA from squamous cell carcinoma cases of cats in Japan. *J Vet Med Sci*. (2018) 80:1236–40. doi: 10.1292/jvms.18-0089
11. Rattay S, Hufbauer M, Hagen C, Putschli B, Coch C, Akgül B, et al. Human beta papillomavirus type 8 E1 and E2 proteins suppress the activation of the RIG-I-like receptor MDA5. *Viruses*. (2022) 14:1361. doi: 10.3390/v14071361
12. Van Doorslaer K. Evolution of the papillomaviridae. *Virology*. (2013) 445:11–20. doi: 10.1016/j.virol.2013.05.012
13. De Villiers E-M, Fauquet C, Broker TR, Bernard H-I, Zur Hausen H. Classification of papillomaviruses. *Virology*. (2004) 324:17–27. doi: 10.1016/j.virol.2004.03.033
14. Dunowska M, Munday JS, Laurie RE, Hills SFK. Genomic Characterisation of *Felis catus* papillomavirus 4, a novel papillomavirus detected in the oral cavity of a domestic cat. *Virus Genes*. (2014) 48:111–9. doi: 10.1007/s11262-013-1002-3
15. Schwartz S. Papillomavirus transcripts and posttranscriptional regulation. *Virology*. (2013) 445:187–96. doi: 10.1016/j.virol.2013.04.034
16. Buck CB, Day PM, Trus BL. The papillomavirus major capsid protein L1. *Virology*. (2013) 445:169–74. doi: 10.1016/j.virol.2013.05.038
17. Bernard H-U. Regulatory elements in the viral genome. *Virology*. (2013) 445:197–204. doi: 10.1016/j.virol.2013.04.035
18. Munday JS, Hardcastle MR, Sim M. Detection of a putative novel papillomavirus type within a large exophytic papilloma on the fetlock of a horse. *Pathogens*. (2020) 9:816. doi: 10.3390/pathogens9100816
19. Liu S, Li Y, Zhang D, Su X, Yue C, Ayala JE, et al. Mortality analysis of captive red panda cubs within Chengdu, China. *BMC Vet Res*. (2022) 18:68. doi: 10.1186/s12917-022-03170-2
20. Zhang W, Yang S, Shan T, Hou R, Liu Z, Li W, et al. Virome comparisons in wild-diseased and healthy captive giant pandas. *Microbiome*. (2017) 5:90. doi: 10.1186/s40168-017-0308-0
21. Zhou L, Lu X, Zhao C, Zhang Y, Ning S, Zhang W. Characterization of a novel picornavirus prevalent in experimental rabbits (*Oryctolagus cuniculus*). *Heliyon*. (2023) 9:e15702. doi: 10.1016/j.heliyon.2023.e15702
22. Bao S, Wang H, Li W, Wu H, Lu C, Yong L, et al. viral metagenomics of the gut virome of diarrheal children with rotavirus A infection. *Gut Microbes*. (2023) 15:2234653. doi: 10.1080/19490976.2023.2234653
23. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. (2012) 28:1647–9. doi: 10.1093/bioinformatics/bts199
24. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods*. (2015) 12:59–60. doi: 10.1038/nmeth.3176
25. Bagci C, Patz S, Huson DH. DIAMOND+MEGAN: fast and easy taxonomic and functional analysis of short and long microbiome sequences. *Curr Protoc*. (2021) 1:e59. doi: 10.1002/cpz1.59
26. Zhao M, Yue C, Yang Z, Li Y, Zhang D, Zhang J, et al. Viral metagenomics unveiled extensive communications of viruses within giant pandas and their associated organisms in the same ecosystem. *Sci Total Environ*. (2022) 820:153317. doi: 10.1016/j.scitotenv.2022.153317
27. Wang J, Chitsaz F, Derbyshire MK, Gonzales NR, Gwadz M, Lu S, et al. The conserved domain database in 2023. *Nucleic Acids Res*. (2023) 51:D384–8. doi: 10.1093/nar/gkac1096
28. Aziz T, Sarwar A, Naveed M, Shahzad M, Shabbir MA, Dabool AS, et al. Bio-molecular analysis of selected food derived *Lactiplantibacillus* strains for CLA production reveals possibly a complex mechanism. *Food Res Int*. (2022) 154:111031. doi: 10.1016/j.foodres.2022.111031
29. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res*. (2018) 46:W296–303. doi: 10.1093/nar/gky427
30. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms' edited by F. U Battistuzzi. *Mol Biol Evol*. (2018) 35:1547–9. doi: 10.1093/molbev/msy096
31. Lu X, Zhu R, Dai Z. Characterization of a novel papillomavirus identified from a whale (*Delphinapterus leucas*) pharyngeal metagenomic library. *Virol J*. (2023) 20:48. doi: 10.1186/s12985-023-02009-y
32. Huelsenbeck JP, Ronquist F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*. (2001) 17:754–5. doi: 10.1093/bioinformatics/17.8.754
33. Tamura K, Stecher G, Kumar S. MEGA11: molecular evolutionary genetics analysis version 11' edited by F. U Battistuzzi. *Mol Biol Evol*. (2021) 38:3022–7. doi: 10.1093/molbev/msab120
34. Bergvall M, Melendy T, Archambault J. The E1 proteins. *Virology*. (2013) 445:35–56. doi: 10.1016/j.virol.2013.07.020
35. Parkinson NJ, Ward A, Malbon AJ, Reardon RJM, Kelly PG. Bovine papillomavirus gene expression and inflammatory pathway activation vary between equine sarcoid tumour subtypes. *Vet Immunol Immunopathol*. (2024) 277:110838. doi: 10.1016/j.vetimm.2024.110838
36. Leroy L, Barbosa JARG, Prat-Gay GD, Polikarpov I, Pinheiro CB. The structure of the extended E2 DNA-binding domain of the bovine papillomavirus-1. *Prot Struct Funct Bioinf*. (2020) 88:106–12. doi: 10.1002/prot.25773
37. Da Silva-Júnior AHP, De Oliveira Silva RC, Diniz Gurgel APA, Barros-Júnior MR, Gomes Nascimento KC, et al. Identification and functional implications of the E5 oncogene polymorphisms of human papillomavirus type 16. *Trop Med Infect Dis*. (2024) 9:140. doi: 10.3390/tropicalmed9070140
38. Liu X, Clements A, Zhao K, Marmorstein R. Structure of the human papillomavirus E7 oncoprotein and its mechanism for inactivation of the retinoblastoma tumor suppressor. *J Biol Chem*. (2006) 281:578–86. doi: 10.1074/jbc.M508452000
39. Zhao X, Dai C, Qian S, Tang Q, Li L, Hao Y, et al. Viral Diversity and epidemiology in critically endangered yangtze finless porpoises (*Neophocaena asiaeorientalis asiaeorientalis*) edited by B. He. *Microbiol Spect*. (2023) 11:e00810-23. doi: 10.1128/spectrum.00810-23
40. Khattri M, Amako Y, Gibbs JR, Collura JL, Arora R, Harold A, et al. Methyltransferase-independent function of enhancer of zeste homologue 2 maintains tumorigenicity induced by human oncogenic papillomavirus and polyomavirus. *Tumour Virus Res*. (2023) 16:200264. doi: 10.1016/j.tvr.2023.200264
41. Roman A, Munger K. The papillomavirus E7 proteins. *Virology*. (2013) 445:138–68. doi: 10.1016/j.virol.2013.04.013
42. De Freitas L, Natália MGD, Gomes D, Pavan AR, Sousa Â, Dos Santos JL, et al. Histone deacetylase inhibitors as therapeutic interventions on cervical cancer induced by human papillomavirus. *Front Cell Dev Biol*. (2021) 8:592868. doi: 10.3389/fcell.2020.592868
43. Roux S, Hallam SJ, Woyke T, Sullivan MB. Viral dark matter and virus-host interactions resolved from publicly available microbial genomes. *Elife*. (2015) 4:e08490. doi: 10.7554/eLife.08490
44. Gysens L, Vanmechelen B, Maes P, Martens A, Haspeslagh M. Complete genomic characterization of bovine papillomavirus type 1 and 2 strains infers ongoing cross-species transmission between cattle and horses. *Vet J*. (2023) 298–9:106011. doi: 10.1016/j.tvjl.2023.106011
45. Maarouf M, Rai KR, Goraya MU, Chen J-L. Immune ecosystem of virus-infected host tissues. *Int J Mol Sci*. (2018) 19:1379. doi: 10.3390/ijms19051379
46. De Falco F, Cutarelli A, Cuccaro B, Catoi C, De Carlo E, Roperto S. Evidence of a novel cross-species transmission by ovine papillomaviruses. *Transbound Emerg Dis*. (2022) 69:3850–7. doi: 10.1111/tbed.14756
47. Braspenning J, Marchini A, Albarani V, Levy L, Ciccolini F, Cremonesi C, et al. The CXXC ZN binding motifs of the human papillomavirus type 16 E7 oncoprotein are not required for its *in vitro* transforming activity in rodent cells. *Oncogene*. (1998) 16:1085–9. doi: 10.1038/sj.onc.1201617
48. Martinez-Zapien D, Ruiz FX, Poirson J, Mitschler A, Ramirez J, et al. Structure of the E6/E6AP/P53 complex required for HPV-mediated degradation of P53. *Nature*. (2016) 529:541–5. doi: 10.1038/nature16481
49. Ledl A, Schmidt D, Müller S. Viral oncoproteins E1A and E7 and cellular LxCxE proteins repress SUMO modification of the retinoblastoma tumor suppressor. *Oncogene*. (2005) 24:3810–8. doi: 10.1038/sj.onc.1208539
50. Putta S, Alvarez L, Lütke S, Sehr P, Müller GA, Fernandez SM, et al. Structural basis for tunable affinity and specificity of LxCxE-dependent protein interactions with the retinoblastoma protein family. *Structure*. (2022) 30:1340–53.e3. doi: 10.1016/j.str.2022.05.019
51. Fehrmann E, Klumpp DJ, Laimins LA. Human papillomavirus type 31 E5 protein supports cell cycle progression and activates late viral functions upon epithelial differentiation. *J Virol*. (2003) 77:2819–31. doi: 10.1128/JVI.77.5.2819-2831.2003
52. Feng N, Yu Y, Wang T, Wilker P, Wang J, Li Y, et al. Fatal canine distemper virus infection of giant pandas in China. *Sci Rep*. (2016) 6:27518. doi: 10.1038/srep27518
53. Zhao S, Hu H, Lan J, Yang Z, Peng Q, Yan L, et al. Characterization of a fatal feline panleukopenia virus derived from giant panda with broad cell tropism and zoonotic potential. *Front Immunol*. (2023) 14:1237630. doi: 10.3389/fimmu.2023.1237630
54. Peng Q, Yang Z, Wu L, Yu P, Li Q, Lan J, et al. Evaluation of the inactivation efficacy of four disinfectants for feline parvovirus derived from giant panda. *Microorganisms*. (2023) 11:1844. doi: 10.3390/microorganisms11071844