

Molecular detection and phylogenetic analysis of pigeon circovirus from racing pigeons in China during 2016-2019

Haoran Wang¹, Hui Gao¹, Zhiwen Jiang¹, Leibo Shi¹, Pengwei Zhao¹, Yanming Zhang¹, and Chengbao Wang¹

¹Northwest Agriculture and Forestry University

June 4, 2021

Abstract

The *Pigeon circovirus* (PiCV) containing a circular single stranded DNA (approximately 2 kb) genome belongs to the genus *Circovirus* of the family *Circoviridae*. PiCV infections in pigeons (*Columba livia*) have been reported worldwide. Currently, pigeon racing is becoming increasingly popular and considered to be a national sport in China, and even, the greatest competitions of racing pigeons are taking place in China. However, there are still no epidemiologic data regarding PiCV infections among racing pigeons in China. To trace the prevalence, genetic variation and evolution of PiCV in sick and healthy racing pigeons, 622 samples were collected from 11 provinces or municipalities in China from 2016 to 2019. Samples were tested by polymerase chain reaction. The results showed that the positive rate of PiCV was 19.3% (120/622) at the sample level and 59.0% (23/39) at the club level, thus suggesting that the virus was prevalent in Chinese racing pigeons. A sequence analysis revealed that the *cap* genes of the PiCV strains identified in our study displayed a high genetic diversity and shared nucleotide homologies of 71.9%–100% and amino acid homologies of 71.7%–100%. 28 and 36 unique amino acid substitutions were observed in the Cap and Rep proteins derived from our PiCV strains, respectively. Furthermore, The ATT and GTG were highly suspected to be the start codons of the Cap protein. A *cap*-gene-based phylogenetic analysis showed that the strains in this study could be further divided into seven clades (A, B, C, E, G, H, and I) and some of them were closely related to worldwide strains from different types of pigeons. A large number of recombination events (31 events) were also detected in the PiCV genomes from Chinese racing pigeons. These findings indicate that PiCV strains circulating in China exhibit a high genetic diversity.

Molecular detection and phylogenetic analysis of pigeon circovirus from racing pigeons in China during 2016-2019

Running title: PiCV in racing pigeons, China

Haoran Wang, Hui Gao, Zhiwen Jiang, Leibo Shi, Pengwei Zhao, Yanming Zhang, Chengbao Wang[§]

College of Veterinary Medicine, Northwest A&F University, Yangling, China

[§]Corresponding author

Chengbao Wang

College of Veterinary Medicine, Northwest A&F University, No. 22 Xinong Road, Yangling, Shaanxi 712100, China

Phone: +86-18049569130; Fax: +86-29-87091032

E-mail addresses: wangchengbao@nwafu.edu.cn

Summary

The *Pigeon circovirus* (PiCV) containing a circular single stranded DNA (approximately 2 kb) genome belongs to the genus *Circovirus* of the family *Circoviridae*. PiCV infections in pigeons (*Columba livia*) have been reported worldwide. Currently, pigeon racing is becoming increasingly popular and considered to be a national sport in China, and even, the greatest competitions of racing pigeons are taking place in China. However, there are still no epidemiologic data regarding PiCV infections among racing pigeons in China. To trace the prevalence, genetic variation and evolution of PiCV in sick and healthy racing pigeons, 622 samples were collected from 11 provinces or municipalities in China from 2016 to 2019. Samples were tested by polymerase chain reaction. The results showed that the positive rate of PiCV was 19.3% (120/622) at the sample level and 59.0% (23/39) at the club level, thus suggesting that the virus was prevalent in Chinese racing pigeons. A sequence analysis revealed that the *cap* genes of the PiCV strains identified in our study displayed a high genetic diversity and shared nucleotide homologies of 71.9%–100% and amino acid homologies of 71.7%–100%. 28 and 36 unique amino acid substitutions were observed in the Cap and Rep proteins derived from our PiCV strains, respectively. Furthermore, The ATT and GTG were highly suspected to be the start codons of the Cap protein. A *cap*-gene-based phylogenetic analysis showed that the strains in this study could be further divided into seven clades (A, B, C, E, G, H, and I) and some of them were closely related to worldwide strains from different types of pigeons. A large number of recombination events (31 events) were also detected in the PiCV genomes from Chinese racing pigeons. These findings indicate that PiCV strains circulating in China exhibit a high genetic diversity.

KEYWORDS

genetic diversity, molecular epidemiology, phylogenetic analysis, *Pigeon circovirus*, sequence analysis

INTRODUCTION

Members of the *Circoviridae*, within the order *Circovirus*, contain a small, circular, non-enveloped, single stranded DNA genome ranging from 1.7 to 2.5 kb in size (Stenzel et al., 2017). The *Pigeon circovirus* (PiCV) or the *Columbid circovirus* (CoCV), together with *Porcine circoviruses* (PCV) types 1, 2, 3 and 4, *Psittacine beak and feather disease virus* (BFDV), *Duck circovirus* (DuCV), *Goose circovirus* (GoCV), *Canary circovirus* (CaCV), *Raven circovirus* (RaCV), *Starling circovirus* (StCV), *Swan circovirus* (SwCV), *Finch circovirus* (FiCV), and *Gull circovirus* (GuCV), etc. belong to the genus *Circovirus* of the family *Circoviridae* (<http://www.ictvonline.org/>). PiCV is a small, circular, non-enveloped virus that contains a single-stranded DNA genome with approximately 2 kb in size. The genome of PiCV has two main open reading frames (ORFs). The ORF-V1 located on the virion sense strand encodes a replication-associated protein (Rep), and the ORF-C1 located on the complementary sense strand encodes a capsid protein (Cap) (Mankertz et al., 2000; Todd et al., 2001). The gene forming ORF-C1 of PiCV has been demonstrated to be highly genetically diverse as compared with the gene forming ORF-V1 (Stenzel et al., 2017; Zhang et al., 2015). The other ORFs including ORF-C2, ORF-C3, and ORF-C4 located on the complementary sense strand encode three viral proteins of unknown functions (Mankertz et al., 2000; Todd et al., 2008). The circovirus infection in pigeons was first identified in 1993 in the USA (Woods et al., 1993) and had been considered to be strongly associated with young pigeon disease syndrome (YPDS), including weakened racing performance, weight loss, lethargy, anorexia, respiratory distress, and diarrhea (Raue et al., 2005).

A variety of methods have been employed to detect PiCV infection in clinical specimens. The original diagnose methods electron microscopy, histology, dot blot hybridization, and in situ hybridization (Yamamoto et al., 2015) are time consuming (Smyth et al., 2001; Soike et al., 2001; Todd et al., 2002; Tsai et al., 2014). The molecular biology techniques, such as the standard PCR, Real-time PCR (Duchatel et al., 2009), NGS techniques (Wang et al., 2017), and the loop-mediated isothermal amplification method (Tsai et al., 2014), enabled more rapid and accurate detection of PiCV infections. As a result, cases of PiCV infections in pigeons were subsequently reported in various countries, including Northern Ireland (Todd et al., 2001), Germany (Mankertz et al., 2000; Raue et al., 2005), Italy (Franciosi et al., 2005), France (Abadie et al., 2001), Czech Republic (Taras et al., 2003), Belgium (Duchatel et al., 2005; Duchatel et al., 2006), Poland (Stenzel et al., 2012; Stenzel et al., 2014c), Slovenia (Krapez et al., 2012), Hungary (Cságola et al., 2012), United Arab Emirates (Ledwoń et al., 2011), Iran (Mahzounieh et al., 2014), China (Liao et al., 2015; Wang et

al., 2017; Zhang et al., 2015), Japan (Yamamoto et al., 2015), and the USA (Roy et al., 2003). Previous studies also demonstrated that PiCV had been detected in different types of pigeons, including racing, fancy, feral and meat pigeons (Stenzel et al., 2014b; Stenzel et al., 2012; Wang et al., 2017). However, all of the above-mentioned birds belong to *Columba livia* species.

In China, the first PiCV infection was detected from meat pigeons in Zhejiang province in 2009, and the full genome was sequenced (Yu et al., 2009). In recent years, several studies have proved that PiCV was prevalent among meat pigeons in eastern and southern China (Wang et al., 2017; Zhang et al., 2015). Recently, the competitions of racing pigeons are becoming increasingly popular in China, and even, pigeon racing is considered to be a national sport. The Chinese Association of Racing Pigeon Breeders has over 5 million members. Chinese breeders compete in numerous races for different distances in their sections throughout the racing season, and more than 25 million racing pigeons from approximately 750 racing clubs were selected to take part in the competitions every year. However, there are no epidemiologic data on PiCV infections among racing pigeons. In order to investigate the prevalence, evolution, and genome characterization of PiCV in racing pigeons in China, an extensive epidemiological investigation and bioinformatic analysis of PiCV from racing pigeons were undertaken in this study. The purpose of this study is to provide novel epidemiologic data and genome characterization for PiCV strains identified from racing pigeons in China.

2 MATERIALS AND METHODS

2.1 Ethics statement

This study was performed according to the animal welfare guidelines of the World Organization for Animal Health, and animal sampling was strictly carried out in accordance with guidelines established by the Ethics of Animal Experiments of Northwest A&F University, Yangling, China. All protocols were approved by this committee (Permit Number: 2014BAD23B11). The field study did not include endangered or protected species. No specific permissions were required for the collection of samples because the samples were collected from public areas or non-protected areas.

2.2 Sample collection

In total, 571 serum samples were collected from healthy pigeons from 39 racing pigeon clubs between November 2016 and August 2019 in 11 provinces or municipalities of China, including Beijing, Hebei, Liaoning, Shanxi, Gansu, Qinghai, Shandong, Shaanxi, Xinjiang Uygur Autonomous Region, Ningxia Hui Autonomous Region and Inner Mongolia Autonomous Region. Moreover, 51 fatally diseased racing pigeons from 7 racing pigeon clubs in 5 provinces or municipalities were submitted to our laboratory at College of Veterinary Medicine, Northwest A&F University to determine the pathogen of the disease. The 0.2 g sections of internal organs (the heart, liver, spleen, lung, and intestine) were obtained during the postmortem examination of all mentioned dead pigeons. All samples were stored at -80°C before DNA extraction.

2.3 DNA extraction

The tissue samples were ground to powder with liquid nitrogen and diluted with three volumes of phosphate-buffered saline. The samples (1 ml) were centrifuged at 12,000xg for 10 min at 4degC and the supernatants were transferred into a 1.5 ml tube. Nucleic acids from tissues and serum samples were extracted by using an EasyPure^(r) Viral DNA/RNA Kit (TransGen Biotech, Beijing, China) according to the manufacturer's instructions. The extracted genomic DNA was stored at -20degC before use.

2.4 PCR detection of PiCV

The PiCV was first detected using a PCR method targeting a 326-base fragment of *cap* gene as described by Freick et al. (2008). The primer sequences were: PiCV-s, 5'-TTGAAAGGTTTTCAGCCTGGC-3' and PiCV-as, 5'-AGGAGACGAAGGACACGCCTC-3' (Freick et al., 2008). The full genomes of PiCV for all the positive samples were amplified by PCR using the primers: PiCV-1F, 5'-ACCCGCGACTTGGAGCCACGGAG-3' and PiCV-1R, 5'-TTGCTCCCGCATTCGCGGTCGCT-3'; PiCV-2F, 5'-GACACTAGTAAAGGGACCCAAGCCA-3' and

PiCV-2R, 5'-AAGCCTTGCAGATGCGGGGT-3', respectively. PCR was performed by using Q5 Hot Start High-Fidelity 2xMaster Mix (NEB, MA, USA). The contents of the three reactions mixture in a 50 μ l reaction volume were as follows: 0.5 μ M forward primer, 0.5 μ M reverse primer, 1 μ g genomic DNA, 25 μ l Q5 High-Fidelity 2xMaster Mix (NEB, MA, USA) and an appropriate volume of Nuclease-Free Water. The cycling parameters were 30 cycles of 98°C for 10 s, 55°C for 30 s and 72°C for 30 s, followed by a final extension at 72°C for 5 min using an automated BioRad T100 Thermal Cycler (Bio-Rad Laboratories, Inc., CA, USA). The PCR products (5 μ l) were resolved on 1% (w/v) agarose gels, and followed by staining with ethidium bromide. Finally, the bands of nucleic acid were visualized with UV illumination inside a gel documentation apparatus (Bio-Rad Laboratories, Inc., CA, USA) and saved as digital photographs.

2.5 Genome characterization and sequence analysis of the *cap*, *rep* genes and PiCV genome

The *cap* gene sequences, *rep* gene sequence and the full genome sequences of PiCV obtained in this study have been deposited in GenBank, respectively. The sequences obtained in the current study were compared with all available PiCV sequence data from different geographical locations within China (n=60) and the rest of the world (n=73) from the NCBI nucleotide database (Table S1). Multiple sequence alignments were carried out by ClustalW algorithm using MEGA 5.0 software (Tamura et al., 2011) and the homology among nucleotide and amino acid sequences was determined by using BioEdit v. 7.0.5 software (Hall, 1999). A divergence analysis of the Cap protein of PiCV was performed using WebLogoTm (<http://weblogo.threeplusone.com/>), an online software for sequence logo generator (Crooks et al., 2004). A comparison of entropy (Hx) in amino acid sequences of Cap proteins of PiCV was carried out by using GraphPad Prism 7.0 software (GraphPad Software, Inc., USA).

2.6 Phylogenetic analysis

The entire *cap* gene sequences of PiCV obtained in this study were used for a phylogenetic analysis. All available PiCV sequence data from different geographical locations within China (n=54) and the rest of the world (n=72) were retrieved from the NCBI nucleotide database as reference sequences (Table S1). Maximum likelihood phylogenetic trees were constructed with the GTR+G+I model by using the MEGA 5.0 software, with partial deletion to handle alignment gaps and 1,000 bootstrap iterations. The phylogenetic tree was annotated with the Interactive Tree Of Life (iTOL) software (<http://itol.embl.de/>), an online tool for the display and annotation of phylogenetic trees (Letunic et al., 2006).

2.7 Recombination analysis

In order to analyze the potential recombinant evidence, an integrated software package for the recombination detection program RDP4 (Martin et al., 2010) was used to detect potential recombinant strains, parental strains, and possible recombination breakpoints. Seven methods (RDP, GeneConv, BootScan, MaxChi, Chimaera, SiScan, and 3Seq) were implemented using the RDP4 program. Recombination events were identified by at least three of the aforementioned methods, and the *p*-value < 0.05 were considered plausible recombinant events. The sequences in the analyzed dataset that most closely resembled the parental sequences of recombinants were defined as either “minor parents” or “major parents” based on the size of the genome fragments which had contributed to the detected recombinants (with the major parent contributing the larger fragment and the minor parent the smaller).

3 RESULTS

3.1 PCR detection and prevalence analysis of PiCV infection in Chinese racing pigeons

In this study, PiCV was detected in samples from both sicked and healthy racing pigeons. The results indicated that the positive rates of PiCV were 19.3% (120/622) at the sample level. Among the samples, 14 out of 51 sick pigeons (positive rate, 27.4%) and 106 out of 571 healthy pigeons (positive rate, 18.6%) were identified to be PiCV positive, respectively. In this study, 23 out of 39 racing pigeon clubs were positive for PiCV (positive rate, 59.0%) (Table 1). Detailed information for 120 PiCV positive samples from racing pigeons in China is shown in Table 1.

3.2 Genome characterization of PiCV genomes

A total of sixty-seven full genome sequences were obtained and deposited in GenBank with accession numbers: MW181925 to MW181991 (Table S2). The assembled circular whole genome sequences were detected as eleven sizes ranging from 2030 to 2045 nt with their *cap* gene ranging from 813 to 828 nt in length and *rep* gene of 948 or 954 nt as shown in Table S2. Among these variable genome length, the most common genome sizes were 2037 (n=21) and 2042 (n=13) nucleotides, respectively (Table S2). The sequence comparison among the 67 identified PiCV strains revealed nucleotide homologies of 84.2%–100% and these sequences exhibited nucleotide homologies of 83.0%–97.8% and 82.0%–98.3% as compared with Chinese PiCV reference strains and the ones from the other countries (Table S3). In particular, four unique nucleotide substitution (CG-GCGA) was found in the intergenic region between *rep* and *cap* stop codons of two strains (SX1/SN/2017/MW181932 and YB4/SN/2018/MW181981) compared with all reference strains.

3.3 Analysis of nucleotide sequence of the *cap* genes and amino acid sequence of the Cap proteins

To explore the genetic diversity of PiCV strains, the *cap* genes of 90 PiCV strains from the 120 PiCV positive samples were sequenced. Detailed information for the 90 *cap* genes was shown in Table S2. The 90 *cap* genes were used to compare with the reference sequences from China and other countries, respectively. The results showed that the 90 *cap* genes ranged from 813 to 828 nt in length. Fourteen, thirteen, one, fifty-three, seven and two of the 90 *cap* nucleotide sequences were 813 nt, 816 nt, 819 nt, 822 nt, 825 nt, and 828 nt in length, encoding a Cap protein of 270, 271, 272, 273, 274, and 275 residues, respectively (Table S2). In addition, we found that ATT and GTG also existed in the position of the start codon site. The sequence comparison of the 90 identified *cap* genes revealed nucleotide homologies of 71.9%–100% and deduced amino acid homologies of 71.7%–100%, and the *cap* genes exhibited low sequence similarities with Chinese PiCV reference strains (73.0%–99.6% nucleotide identity, 72.3%–100% amino acid identity) and the PiCV reference strains from other countries (68.8%–98.4% nucleotide identity, 63.6%–100% amino acid identity) (Table S3).

To investigate variations in the deduced amino acid sequences, the amino acid sequences of 90 identified Cap proteins and the reference strains were aligned. The results showed that there were nine major locations of deletion (compared to the consensus sequence) among the Cap proteins including locations 7, 24, 29, 30, 35, 58, 130, 182, and 266 (Figure 1). A comparison of entropy (Hx) in amino acid sequences of Cap proteins showed that the Hx of amino acid sequences at most positions from the 90 identified Cap proteins were higher than that of the PiCV reference strains (Figure 2). In addition, some unique amino acid substitutions at 28 different positions were observed among the 90 Cap proteins as shown in Figure 3.

3.4 Analysis of nucleotide sequence of the *rep* genes and amino acid sequence of the Rep proteins

The *rep* genes for 68 out of 120 PiCV positive samples were successfully sequenced. Detailed information was shown in Table S2. The 68 identified *rep* genes were used to compare with PiCV reference sequences from China and other countries. The results showed that all of the 68 *rep* genes used the ATG start codon and were identified as two sizes: 948 and 954 nt. Forty of the 68 *rep* nucleotide sequences were 948 nt in length, encoding a Rep protein with 315 residues. Twenty-eight of the 68 *rep* nucleotide sequences were 954 nt in size, encoding a Rep protein with 317 residues (Table S2). The difference in size was due to a 2 amino acid deletion at positions 2 and 3. The sequence alignment revealed nucleotide homologies of 90.3%–100% and deduced amino acid homologies of 92.7%–100% among the 68 *rep* genes. These sequences exhibited higher similarities with PiCV reference strains from China (89.0%–99.2% nucleotide identity, 89.2%–99.6% amino acid identity) and other countries (89.5%–98.3% nucleotide identity, 90.5%–99.3% amino acid identity) as compared to the *cap* genes (Table S3). To investigate variations in the deduced amino acid sequences of *rep* gene products, the amino acid sequences of 68 identified *rep* genes and the reference strains were aligned. The results showed that some unique amino acid substitutions at 36 different positions were observed among the 68 identified PiCV strains (Data not shown).

3.5 Phylogenetic analysis of the *cap* genes

A phylogenetic analysis using 126 *cap* gene reported in GenBank, and *cap* genes of 90 PiCV strains obtained in our study was performed to investigate genetic relationship (Table S1 and S2). The phylogenetic tree indicated that a total of 216 analyzed PiCV strains could be divided into 9 clades, which were defined by the contractual letters A–I (Figure 4). The 90 *cap* gene sequences reported in the present study were clustered into 7 clades: A (14.1%), B (60.9%), C (62.2%), E (25.0%), G (51.4%), H (28.6%), and I (40.0%). As shown in Figure 4, the PiCV strains from different geographical locations were clustered into the same clades, such as QYQX1/HE/2018/MW181909, TY1/SN/2016/MW181929, DS1/GS/2018/MW181970, and QD3/SN/2018/MW181965. On the other hand, the PiCV strains from the same geographical locations were divided into different clades, such as LH1/HE/2018/MW181918 (clade B), LH2/HE/2019/MW181988 (clade A), and LH3/HE/2019/MW181989 (clade C). Furthermore, the PiCV strains collected in the different years were grouped closely in the same clade. Due to the fact that the racing pigeons were supplied by breeders from different provinces or municipalities of China, the 78 PiCV strains isolated from the pigeons in the racing clubs in Shannxi province, belonged to six separate clades except for clade D, F, and H.

3.6 Recombination analysis in PiCV genomes

For 67 identified PiCV strains, 31 recombination events were detected using RDP4 (Table S4). These recombination events indicated that the strains originally from different types of pigeons or from different geographical locations may have some evidence of recombination. For example, a segment (~1001bp) of three PiCV strains (LT3/SN/2018/MW181940, LT2/SN/2018/MW181939, and WQ6/SN/2018/MW181948) obtained in this study was likely descended from a large number of ancestral PiCV genomes that originated from Chinese meat pigeons, German racing pigeons, British racing pigeons, Polish fancy pigeons, and French meat pigeons (event 2). The possible breakpoints for recombination were determined during the recombination analysis. The results also indicated that the recombination breakpoint hot plots located within both the intergenic region between the *rep* and *cap* stop codons and near the virion strand origin of replication. Furthermore, the recombination breakpoint cold spots were detected within the central region of *cap* gene, which indicated few evidence for recombination within *cap* gene (Figure 5). Additionally, four PiCV genomes (TF1/SN/2016/MW181925, DA1/XJ/2018/MW181982, QD7/SN/2019/MW181990, and LH2/HE/2019/MW181988) had no evidence for recombination.

Discussion

In recent studies, Rotavirus A G18P[17] had been confirmed as a primary cause of YPDS-like diseases in domestic pigeons (Rubbenstroth et al., 2020). Furthermore, PiCV is also supposed to be an etiological agent of YPDS mostly affecting young pigeons worldwide (Raue et al., 2005). PiCV has been reported to be prevalent in at least fourteen countries. In China, PiCV was first detected in meat pigeons in 2009 (Yu et al., 2009). In recent years, the epidemiological survey showed that the positive rates of PiCV infection in Chinese meat pigeons were 19.67% and 75.3% in the poultry farms of eastern China in 2009 and 2015 (Wang et al., 2017; Zhang et al., 2015), implying that PiCV is widely distributed among meat pigeon populations in eastern China. However, the epidemiology and distribution of PiCV in the racing pigeons is unknown. The main objective of this work was to evaluate the genetic diversity and epidemiology of PiCV strains circulated in the racing pigeons of China. Positive samples were detected from seven provinces and the prevalence rates among these provinces were variant. Overall, our data implied that PiCV was also widely distributed in diseased racing pigeons and healthy racing pigeons in China for the first time.

Previous studies showed that the length of PiCV genome was 2031–2043 nt (Loiko et al., 2018; Sarker et al., 2019; Wang et al., 2017). However, our findings first showed that the 2030 nt (n=1), 2044 nt (n=2), and 2045 nt (n=1) complete genome were present in the PiCV positive samples. Despite many genetic diversities were found in the genome of PiCV strains, no mutation was found in the conserved nonanucleotide motif (TAGTATTAC) (Mankertz et al., 2000; Todd et al., 2001) located at the apex of a potential stem-loop which was putatively associated with the initiation of rolling circle replication (RCR) (Mankertz et al., 2000; Todd et al., 2001). Previous studies have shown many genetic diversities in the *cap* gene (Cságola et al., 2012; Stenzel et al., 2014a; Wang et al., 2017). Our findings showed that the 90 identified *cap* genes exhibited a higher diversity as compared with the reference strains. Some unique amino acid substitutions at

28 different positions were observed among the Cap proteins. In addition, the *cap* nucleotide sequence with 828 nt in length encoding a novel Cap protein of 275 amino acids was first identified in two Chinese PiCV strains, TY3/SN/2016/MW181931 (clade E) and WL4/SN/2018/MW181959 (clade E). Higher diversity of Cap protein versus the Rep protein is due to the fact that the Cap protein, as the protein shell of the virus, is exposed to the host's immune cells, resulting in a stimulation of a cascade of immune responses. An interaction with the immune system may result in adaptive mutations in the Cap protein. Some mutations can change the structure of Cap protein. This may lead to a better binding of virus with the receptors of target cells, which increases the infection ability of virus in cells. Moreover, these mutations may also protect virus from being neutralized by antibodies (Bassami et al., 1998; Bennett et al., 2006; Stenzel et al., 2017). In this study, we identified many mutations in both N- and C-terminus of the Cap protein. According to the investigations on Cap protein of BFDV, the arginine-rich N-terminus has two functions (nuclear localization and nucleic acid binding) that enable as penetration of the viral genome into the host cell nucleus through nuclear pore complex (Heath et al., 2006). Like BFDV, the N-terminus of PiCV Cap protein is also rich in arginine and has been predicted to be a nuclear localization signal and a nucleic acid binding domain by software, which means that this region is highly possible to be functionally similar to the corresponding region of the Cap protein of BFDV (Lai et al., 2014). Moreover, the results of this study showed that most of the amino acid deletion sites in Cap protein were concentrated in the N-terminus and many novel mutations also have been identified in this region. However, due to the lack of relevant experimental evidence, we cannot infer whether these mutations and deletions could affect the function of Cap protein. The Cap protein has been used as a coating antigen in ELISA due to its antigenic activity and can be recognized by PiCV specific antibodies (Daum et al., 2009; Stenzel et al., 2016). If the N-terminal region of the PiCV Cap protein, which is located within the capsid, resembles the corresponding region in BFDV, the N-terminal amino acid mutation and deletion may not affect the ability to bind to the antibody as a coating antigen (Daum et al., 2009). In addition, there are also many mutations in the C-terminus where most sites are composed of hydrophilic amino acids. It means that this region may locate on the outside of the protein and may have ability for binding with antibody. Therefore, mutations of this region may lead to changes in antigenicity. Cap protein, which contains neutralizing antibody epitopes, can be considered as a potential antigen candidate in sub-unit vaccine development (Gai et al., 2020; Stenzel et al., 2018). Moreover, the sub-unit vaccines based on PCV2 recombinant capsid proteins are successfully used in the prevention of PCV2-SD (Blanchard et al., 2003; Fort et al., 2009; Li et al., 2015; Zhu et al., 2016). It means that the development of a sub-unit vaccine may protect pigeons from infection with PiCV (Stenzel et al., 2018). Due to a high genetic diversity of Cap proteins may lead to differences in antigenicity between different strains, sub-unit vaccines may also be diversified in future. To sum up, since there is little research on the function of PiCV Cap protein, the specific meaning of these mutations cannot be accurately explained. Therefore, further experiments are needed to explain if these mutations and deletions affect the function of Cap protein. According to the obtained nucleotide sequences of *cap* gene, we found that ATT and GTG existed in the position of the start codon site through sequence alignment. In addition, several ATG was found inside *cap* genes. Since no available cell lines have been found to culture PiCV *in vitro* (Huang et al., 2021; Santos et al., 2020; Todd, 2004) and lack of commercial antibodies against PiCV Cap protein, it is hard to verify that indeed a full-length natural Cap protein is translated from the described alternative start codons, rather than a truncated version from a downstream in frame of ATG by experiments now. Therefore, the ATT and GTG were highly suspected to be the start codons of the Cap protein through sequence analysis in this study. A similar approach has been employed for identifying the alternative start codon of Cap protein in other circoviruses, such as GuCV (Todd et al., 2007) and BFDV (Bassami et al., 2001). Moreover, previous studies had confirmed that codons differing from ATG in a single position can support translation initiation in eukaryotes (Peabody, 1989; Wei, J. et al., 2013). Thus, it still needs more experiments to identify if the two non-ATG start codons can be actually used for translation initiation by Cap protein of PiCV. In all 68 identified Rep proteins, three amino acid motifs named FTLNNP (position 41-46), HLQGF (position 78-82), and YCSK (position 116-119) that putatively associated with RCR (Mankertz et al., 2000; Mankertz et al., 1998) were completely conserved. In addition, a fourth motif, which was putatively associated with dNTPase activity, namely GKS (position 197-199) (Mankertz et al., 2000; Mankertz et al., 1998), was also

completely conserved. These sequence characteristics help understand the genetic diversity of PiCV. However, the pathogenicity of different PiCV strains still need to be validated in future.

In the present study, the 90 PiCV strains were divided into seven clades (A, B, C, E, G, H, and I) based on a phylogenetic analysis of *cap* gene. Additionally, we found that PiCV strains isolated from the same club belonged to different clades and shared a low sequence identity. These data suggested that the infection and evolution of PiCV in Chinese racing pigeons might have different evolutionary origins. This may result from the fact that the import of racing pigeons from all around the world to China is very significant and a lack of oversight in the international pigeon trade (racing pigeons mainly) (Ashton, 1984; Stenzel et al., 2012). Interestingly, two isolates, QYQX1/HE/2018/MW181909 (clade A) and QYQX2/HE/2018/MW181910 (clade C), with lower identity (75.5% nucleotide identity and 76.8% amino acid identity) in the *cap* gene were detected from the same pigeon, suggesting a horizontal transmission occurred among the racing pigeons in the same racing clubs (Gerdes, 1993; Woods et al., 1993). The complexity of epidemic PiCV strains in China may cause difficulties to protect pigeon from infection by vaccines in future.

Viral recombination had been proved to play a significant role in the evolution of many ssDNA viruses (Lefevre et al., 2009). The extensive recombination events had been reported in PiCV genomes and other circoviruses, such as PCV (Kleymann et al., 2020; Wei, C. et al., 2019) and BFDV (Julian et al., 2013; Varsani et al., 2011). Similar to the previous reports, 31 recombination events were detected in the 67 identified PiCV strains in this study. Thus, the recombination seemed to be a key mechanism for PiCV evolution (Csagola et al., 2012; Loiko et al., 2018; Sarker et al., 2019; Stenzel et al., 2014c).

In conclusion, our study demonstrated that PiCV infection in racing pigeons was widespread in northern China and revealed the characteristics of the PiCV genome. Furthermore, the identified PiCV strains displayed a high genetic diversity. Our data also demonstrated that PiCV in Chinese racing pigeons had an extensive recombination for the first time. Thus, these data could increase our understanding of the epidemiology and genetic variation of PiCV circulated in northern China and evolutionary relationships among different strains.

ACKNOWLEDGMENTS

This study was funded by grants from Xi'an IRIS Livestock Technology Co. Ltd. (K4030218170), China Postdoctoral Science Foundation (2017M610659, 2018T111113), the State Key Laboratory of Veterinary Etiological Biology, Lanzhou Veterinary Research Institute, Chinese Academy of Agricultural Sciences (SKLVEB2016KFKT014) and the Fundamental Research Funds for the Central Universities (2452019053). The authors thank Ass. Prof. Tomasz Stenzel from Department of Bird Diseases, Faculty of Veterinary Medicine in Olsztyn for proofreading the manuscript.

CONFLICT OF INTERESTS STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

AUTHOR CONTRIBUTIONS

Chengbao Wang and Haoran Wang drafted the manuscript and communicated with the coauthors to coordinate the document editing. Haoran Wang, Hui Gao, Zhiwen Jiang, Leibo Shi and Pengwei Zhao were responsible for sample collection, detection and sequence acquisition. Haoran Wang was responsible for data analysis and interpretation. Chengbao Wang and Yanming Zhang reviewed the manuscript and finalized it. All authors read and approved the final manuscript.

REFERENCES

- Abadie, J., Nguyen, F., Groizeleau, C., Amenna, N., Fernandez, B., Guereaud, C., . . . Wyers, M. (2001). Pigeon circovirus infection: Pathological observations and suggested pathogenesis. *Avian Pathology*, 30, 149-158. doi:10.1080/03079450124811
- Ashton, W. L. (1984). The risks and problems connected with the import and export of captive birds. *British Veterinary Journal*, 140, 317-327. doi:10.1016/0007-1935(84)90121-0
- Bassami, M. R., Berryman, D., Wilcox, G. E., & Raidal, S. R. (1998). Psittacine beak and feather disease virus nucleotide sequence analysis and its relationship to porcine circovirus, plant circoviruses, and chicken anaemia virus. *Virology*, 249, 453-459. doi:org/10.1006/viro.1998.9324
- Bassami, M. R., Ypelaar, I., Berryman, D., Wilcox, G., & Raidal, S. (2001). Genetic Diversity of Beak and Feather Disease Virus Detected in Psittacine Species in Australia. *Virology*, 279, 392-400. doi:10.1006/viro.2000.0847
- Bennett, S., Holmes, E., Chirivella, M., Rodriguez, D., Beltran, M., Vorndam, V., . . . McMillan, W. (2006). Molecular evolution of dengue 2 virus in Puerto Rico: Positive selection in the viral envelope accompanies clade reintroduction. *The Journal of General Virology*, 87, 885-893. doi:10.1099/vir.0.81309-0
- Blanchard, Mahe, D., Cariolet, R., Keranflec'h, A., Baudouard, M. A., Cordoli, P., . . . A.Jestin. (2003). Protection of swine against post-weaning multisystemic wasting syndrome (PMWS) by porcine circovirus type 2 (PCV2) proteins. *Vaccine*, 21, 4565-4575. doi:10.1016/S0264-410X(03)00503-6
- Crooks, G. E., Hon, G., Chandonia, J. M., & Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Research*, 14, 1188-1190. doi:10.1101/gr.849004
- Csagola, A., Lőrincz, M., Tombácz, K., Wladár, Z., Kovács, E., & Tuboly, T. (2012). Genetic diversity of pigeon circovirus in Hungary. *Virus Genes*, 44, 75-79. doi:10.1007/s11262-011-0669-6
- Daum, I., Finsterbusch, T., Härtle, S., Göbel, T., Mankertz, A., Korbel, R., & Grund, C. (2009). Cloning and expression of a truncated pigeon circovirus capsid protein suitable for antibody detection in infected pigeons. *Avian Pathology*, 38, 135-141. doi:10.1080/03079450902737797
- Duchatel, J. P., Todd, D., Curry, A., Smyth, J. A., Bustin, J. C., & Vindevogel, H. (2005). New data on the transmission of pigeon circovirus. *The Veterinary Record*, 157, 413-415. doi:10.1136/vr.157.14.413
- Duchatel, J. P., Todd, D., Smyth, J. A., Bustin, J. C., & Vindevogel, H. (2006). Observations on detection, excretion and transmission of pigeon circovirus in adult, young and embryonic pigeons. *Avian Pathology*, 35, 30-34. doi:10.1080/03079450500465692
- Duchatel, J. P., Todd, D., Willeman, C., & Losson, B. (2009). Quantification of pigeon circovirus in serum, blood, semen and different tissues of naturally infected pigeons using a real-time polymerase chain reaction. *Avian Pathology*, 38, 143-148. doi:10.1080/03079450902737805
- Fort, M., Marina, S., Perez-Martin, E., Nofrarias, M., Mateu, E., & Segalés, J. (2009). One dose of a porcine circovirus 2 (PCV2) sub-unit vaccine administered to 3-week-old conventional piglets elicits cell-mediated immunity and significantly reduces PCV2 viremia in an experimental model. *Vaccine*, 27, 4031-4037. doi:10.1016/j.vaccine.2009.04.028
- Franciosini, M. P., Fringuelli, E., Tarhuni, O., Guelfi, G., Todd, D., Casagrande Proietti, P., . . . Asdrubali, G. (2005). Development of a polymerase chain reaction-based in vivo method in the diagnosis of subclinical pigeon circovirus infection. *Avian Diseases*, 49, 340-343. doi:10.1637/7334-012705R.1
- Freick, M., Müller, H., & Raue, R. (2008). Rapid detection of pigeon herpesvirus, fowl adenovirus and pigeon circovirus in young racing pigeons by multiplex PCR. *Journal of Virological Methods*, 148, 226-231. doi:10.1016/j.jviromet.2007.11.003
- Gai, W., Zheng, W., Zhao, Z., Wong, G., Sun, P., Yan, L., . . . Zheng, X. (2020). Assembly of pigeon circovirus-like particles using baculovirus expression system. *Microbial Pathogenesis*, 139, 103905.

doi:10.1016/j.micpath.2019.103905

Gerdes, G. H. (1993). Two very small viruses—a presumptive identification. *Journal of the South African Veterinary Association*, 64, 2.

Hall, T. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucleic Acids Symposium Series*, 41, 95-98. doi:10.1021/bk-1999-0734.ch008

Heath, L., Williamson, A.-L., & Rybicki, E. (2006). The Capsid Protein of Beak and Feather Disease Virus Binds to the Viral DNA and Is Responsible for Transporting the Replication-Associated Protein into the Nucleus. *Journal of Virology*, 80, 7219-7225. doi:10.1128/JVI.02559-05

Huang, H.-Y., Silva, B. B. I., Tsai, S.-P., Tsai, C.-Y., Tyan, Y.-C., Lin, T.-C., . . . Chuang, K.-P. (2021). Immunogenicity and Protective Activity of Pigeon Circovirus Recombinant Capsid Protein Virus-Like Particles (PiCV rCap-VLPs) in Pigeons (*Columba livia*) Experimentally Infected with PiCV. *Vaccines*, 9, 98. doi:10.3390/vaccines9020098

Julian, L., Piasecki, T., Chrzastek, K., Walters, M., Muhire, B., Harkins, G. W., . . . Varsani, A. (2013). Extensive recombination detected among beak and feather disease virus isolates from breeding facilities in Poland. *The Journal of General Virology*, 94, 1086-1095. doi:10.1099/vir.0.050179-0

Kleymann, A., Soto, E., Illanes, O., Malik, Y. S., Fuentealba, C., & Ghosh, S. (2020). High rates of detection and complete genomic analysis of porcine circovirus 2 (PCV2) in the Lesser Antilles island of St. Kitts: Identification of PCV2b-PCV2d recombinants. *Transboundary and Emerging Diseases*, 67, 2282-2289. doi:10.1111/tbed.13583

Krapez, U., Slavec, B., Steyer, A. F., Pintaric, S., Dobeic, M., Rojs, O. Z., & Dovc, A. (2012). Prevalence of pigeon circovirus infections in feral pigeons in Ljubljana, Slovenia. *Avian Diseases*, 56, 432-435. doi:10.1637/9929-091211-Case.1

Lai, G., Lin, Y., Tsai, Y., Lien, Y., Lin, M., Chen, H., . . . Lee, M. (2014). High yield production of pigeon circovirus capsid protein in the E. coli by evaluating the key parameters needed for protein expression. *BMC Veterinary Research*, 10, 115-115. doi:10.1186/1746-6148-10-115

Ledwoń, A., Bailey, T., O'Donovan, D., Mckeown, S., Lloyd, C., Wieckowski, T., . . . Wernery, U. (2011). Prevalence of circovirus and adenovirus in pigeons in Dubai. *Medycyna Weterynaryjna*, 67, 752-756. doi:10.1016/j.prevetmed.2011.04.010

Lefevre, P., Lett, J. M., Varsani, A., & Martin, D. P. (2009). Widely conserved recombination patterns among single-stranded DNA viruses. *Journal of Virology*, 83, 2697-2707. doi:10.1128/JVI.02152-08

Letunic, I., & Bork, P. (2006). Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics*, 23, 127-128. doi:10.1093/bioinformatics/btl529

Li, P., Qiao, X., Zheng, Q., & Hou, J. (2015). Immunogenicity and immunoprotection of porcine circovirus type 2 (PCV2) Cap protein displayed by *Lactococcus lactis*. *Vaccine*, 34, 696-702. doi:10.1016/j.vaccine.2015.09.007

Liao, P., Wang, K., Tsai, S., Liu, H., Huang, B., & Chuang, K. (2015). Recurrent positive selection and heterogeneous codon usage bias events leading to coexistence of divergent pigeon circoviruses. *The Journal of General Virology*, 96, 2262-2273. doi:10.1099/vir.0.000163

Loiko, M. R., Junqueira, D. M., Varela, A., Tochetto, C., Scheffer, C. M., Lima, D. A., . . . Roehe, P. M. (2018). Columbidae circoviruses detected in free ranging pigeons from Southern Brazil: insights on PiCV. *Archives of Virology*, 163, 3083-3090. doi:10.1007/s00705-018-3990-8

Mahzounieh, M., Heidari Khoei, H., Ghasemi Shamsabadi, M., & Dastjerdi, A. (2014). Detection and phylogenetic characterization of Columbidae circoviruses in Chaharmahal va Bakhtiari province, Iran. *Avian Pathology*, 43, 524-528. doi:10.1080/03079457.2014.966648

- Mankertz, A., Hattermann, K., Ehlers, B., & Soike, D. (2000). Cloning and sequencing of columbid circovirus (CoCV), a new circovirus from pigeons. *Archives of Virology*, 145, 2469-2479. doi:10.1007/s007050070002
- Mankertz, A., Mankertz, J., Wolf, K., & Buhk, H.-J. (1998). Identification of a protein essential for replication of porcine circovirus. *The Journal of General Virology*, 79 (Pt 2), 381-384. doi:10.1099/0022-1317-79-2-381
- Martin, D. P., Lemey, P., Lott, M., Moulton, V., Posada, D., & Lefevre, P. (2010). RDP3: a flexible and fast computer program for analyzing recombination. *Bioinformatics*, 26, 2462-2463. doi:10.1093/bioinformatics/btq467
- Peabody, D. S. (1989). Translation Initiation at Non-Aug Triplets in Mammalian Cells. *Journal of Biological Chemistry*, 264, 5031-5035. doi:10.1016/S0021-9258(18)83694-8
- Raue, R., Schmidt, V., Freick, M., Reinhardt, B., Johne, R., Kamphausen, L., . . . Krautwald-Junghanns, M.-E. (2005). A disease complex associated with pigeon circovirus infection, young pigeon disease syndrome. *Avian Pathology*, 34, 418-425. doi:10.1080/03079450500267825
- Roy, P., Dhillon, A. S., Lauerman, L., & Shivaprasad, H. L. (2003). Detection of Pigeon Circovirus by Polymerase Chain Reaction. *Avian Diseases*, 47, 218-222. doi:10.1637/0005-2086(2003)047[0218:DOPCBP]2.0.CO;2
- Rubbenstroth, D., Ulrich, R., Wylezich, C., Rautenschlein, S., Beer, M., & Mohr, L. (2020). First experimental proof of Rotavirus A (RVA) genotype G18P[17] inducing the clinical presentation of 'young pigeon disease syndrome' (YPDS) in domestic pigeons (*Columba livia*). *Transboundary and Emerging Diseases*, 67, 1507-1516. doi:10.1111/tbed.13485
- Santos, H. M., Chen, C. C., Tsai, C.-Y., Hsish, Y. C., Chung, F. C., Tyan, Y.-C., . . . Chuang, K. P. (2020). Influence of pigeon interferon alpha (PiIFN- α) on pigeon circovirus (PiCV) replication and cytokine expression in *Columba livia*. *Veterinary Microbiology*, 242, 108591. doi:10.1016/j.vetmic.2020.108591
- Sarker, S., Das, S., Ghorashi, S. A., Forwood, J. K., & Raidal, S. R. (2019). Pigeon circoviruses from feral pigeons in Australia demonstrate extensive recombination and genetic admixture with other circoviruses. *Avian Pathology*, 48, 512-520. doi:10.1080/03079457.2019.1629391
- Smyth, J. A., Weston, J., Moffett, D. A., & Todd, D. (2001). Detection of Circovirus Infection in Pigeons by in Situ Hybridization Using Cloned DNA Probes. *Journal of Veterinary Diagnostic Investigation*, 13, 475-482. doi:10.1177/104063870101300604
- Soike, D., Hattermann, K., Albrecht, K., Segalés, J., Domingo, M., Schmitt, C., & Mankertz, A. (2001). A diagnostic study on columbid circovirus infection. *Avian Pathology*, 30, 605-611. doi:10.1080/03079450120092099
- Stenzel, T., Dziewulska, D., Tykałowski, B., Śmiałek, M., Kowalczyk, J., & Koncicki, A. (2018). Immunogenicity of Pigeon Circovirus Recombinant Capsid Protein in Pigeons. *Viruses*, 10, 596. doi:10.3390/v10110596
- Stenzel, T., & Koncicki, A. (2017). The epidemiology, molecular characterization and clinical pathology of circovirus infections in pigeons – current knowledge. *The Veterinary Quarterly*, 37, 166-174. doi:10.1080/01652176.2017.1325972
- Stenzel, T., & Pestka, D. (2014a). Occurrence and genetic diversity of pigeon circovirus strains in Poland. *Acta Veterinaria Hungarica*, 62, 274-283. doi:10.1556/AVet.2014.004
- Stenzel, T., Pestka, D., & Choszcz, D. (2014b). The prevalence and genetic characterization of Chlamydia psittaci from domestic and feral pigeons in Poland and the correlation between infection rate and incidence of pigeon circovirus. *Poultry Science*, 93, 3009–3016. doi:10.3382/ps.2014-04219
- Stenzel, T., Pestka, D., Tykałowski, B., Śmiałek, M., & Koncicki, A. (2012). Epidemiological investigation of selected pigeon viral infections in Poland. *The Veterinary Record*, 171, 562. doi:10.1136/vr.100932

- Stenzel, T., Piasecki, T., Chrzastek, K., Julian, L., Muhire, B. M., Golden, M., . . . Varsani, A. (2014c). Pigeon circoviruses display patterns of recombination, genomic secondary structure and selection similar to those of Beak and feather disease viruses. *The Journal of General Virology*, 95, 1338-1351. doi:10.1099/vir.0.063917-0
- Stenzel, T., Wozniakowski, G., Dziewulska, D., Choszcz, D., Tykałowski, B., Smialek, M., & Koncicki, A. (2016). Application of pigeon circovirus recombinant capsid protein for detecting anti-PiCV antibodies in the sera of asymptomatic domestic pigeons and the potential use of a combination of serological and molecular tests for controlling circovirus infections in pigeon breeding flocks. *Poultry Science*, 96, 303-308. doi:10.3382/ps/pew266
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., & Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, 28, 2731-2739. doi:10.1093/molbev/msr121
- Taras, L., Kubicek, O., Juranova, R., & Jurajda, V. (2003). The First Demonstration of Pigeon Circovirus Infection in the Czech Republic Based on Histology and Nested PCR. *Acta Veterinaria Brno*, 72, 577-582. doi:10.1016/S0378-1135(03)00117-2
- Todd, D. (2004). Avian circovirus diseases: lessons for the study of PMWS. *Veterinary Microbiology*, 98, 169-174. doi:10.1016/j.vetmic.2003.10.010
- Todd, D., Duchatel, J. P., Weston, J. H., Ball, N. W., Borghmans, B. J., Moffett, D. A., & Smyth, J. A. (2002). Evaluation of polymerase chain reaction and dot blot hybridisation tests in the diagnosis of pigeon circovirus infections. *Veterinary Microbiology*, 89, 1-16. doi:10.1016/S0378-1135(02)00154-2
- Todd, D., Fringuelli, E., Scott, A. N., Borghmans, B. J., Duchatel, J. P., Shivaprasad, H. L., . . . Smyth, J. A. (2008). Sequence comparison of pigeon circoviruses. *Research in Veterinary Science*, 84, 311-319. doi:10.1016/j.rvsc.2007.03.007
- Todd, D., Scott, A. N. J., Fringuelli, E., Shivaprasad, H. L., Gavier-Widen, D., & Smyth, J. A. (2007). Molecular characterization of novel circoviruses from finch and gull. *Avian Pathology*, 36, 75-81. doi:10.1080/03079450601113654
- Todd, D., Weston, J. H., Soike, D., & Smyth, J. A. (2001). Genome Sequence Determinations and Analyses of Novel Circoviruses from Goose and Pigeon. *Virology*, 286, 354-362. doi:10.1006/viro.2001.0985
- Tsai, S. S., Chang, Y. L., Huang, Y. L., Liu, H. J., Ke, G. M., Chiou, C. J., . . . Chuang, K. P. (2014). Development of a loop-mediated isothermal amplification method for rapid detection of pigeon circovirus. *Archives of Virology*, 159, 921-926. doi:10.1007/s00705-013-1906-1
- Van Borm, S., Rosseel, T., Steensels, M., van den Berg, T., & Lambrecht, B. (2013). What's in a strain? Viral metagenomics identifies genetic variation and contaminating circoviruses in laboratory isolates of pigeon paramyxovirus type 1. *Virus Research*, 171, 186-193. doi:10.1016/j.virusres.2012.11.017
- Varsani, A., Regnard, G. L., Bragg, R., Hitzeroth, I. I., & Rybicki, E. P. (2011). Global genetic diversity and geographical and host-species distribution of beak and feather disease virus isolates. *The Journal of General Virology*, 92, 752-767. doi:10.1099/vir.0.028126-0
- Wang, K., Zhuang, Q., Qiu, Y., Wang, T., & Chen, J. (2017). Genome sequence characterization of pigeon circoviruses in China. *Virus Research*, 233, 1-7. doi:10.1016/j.virusres.2017.03.007
- Wei, C., Lin, Z., Dai, A., Chen, H., Ma, Y., Li, N., . . . Liu, J. (2019). Emergence of a novel recombinant porcine circovirus type 2 in China: PCV2c and PCV2d recombinant. *Transboundary and Emerging Diseases*, 66, 2496-2506. doi:10.1111/tbed.13307
- Wei, J., Zhang, Y., Ivanov, I. P., & Sachs, M. S. (2013). The Stringency of Start Codon Selection in the Filamentous Fungus *Neurospora crassa*. *The Journal of Biological Chemistry*, 288, 9549-9562.

doi:10.1074/jbc.M112.447177

Woods, L. W., Latimer, K. S., Barr, B. C., Niagro, F. D., Campagnoli, R. P., Nordhausen, R. W., & Castro, A. E. (1993). Circovirus-Like Infection in a Pigeon. *Journal of Veterinary Diagnostic Investigation*, 5, 609-612. doi:10.1177/104063879300500417

Yamamoto, E., Ito, H., Kitamoto, E., Morinishi, K., Yano, A., Miyoshi, S., & Ito, T. (2015). Complete genome sequence of pigeon circovirus detected in racing pigeons in western Japan. *Virus Genes*, 51, 140–143. doi:10.1007/s11262-015-1211-z

Yu, X., Zhu, C., Zheng, X., Mu, A., & Yu, H. (2009). [Cloning and analysis of the complete genomes of pigeon circovirus from Zhejiang Province]. *Bing du xue bao = Chinese journal of virology*, 25, 355-361.

Zhang, Z., Dai, W., Wang, S., & Dai, D. (2015). Epidemiology and genetic characteristics of pigeon circovirus (PiCV) in eastern China. *Archives of Virology*, 160, 199-206. doi:10.1007/s00705-014-2255-4

Zhu, S., Zhang, C., Wang, J., Wei, L., Quan, R., Yang, J., . . . Liu, J. (2016). Immunity Elicited by an Experimental Vaccine Based on Recombinant Flagellin-Porcine Circovirus Type 2 Cap Fusion Protein in Piglets. *PloS One*, 11, e0147432. doi:10.1371/journal.pone.0147432

Table 1 The results of PiCV detection in Chinese racing pigeons.

Classification	Sick	Healthy	PiCV Positive number
	Pos/Tot (%)	Pos/Tot (%)	Pos/Tot (%)
Pigeon Club	5/7 (71.4%)	18/32 (56.3%)	23/39 (59.0%)
Geographic distributions			
Beijing	0/0 (0.0%)	1/19 (5.3%)	1/19 (5.3%)
Shaanxi	1/13 (7.7%)	101/466 (21.7%)	102/479 (21.3%)
Hebei	9/23 (39.1%)	1/25 (4.0%)	10/48 (20.8%)
Liaoning	0/0 (0.0%)	0/18 (0.0%)	0/18 (0.0%)
Inner Mongolia	0/0 (0.0%)	0/6 (0.0%)	0/6 (0.0%)
Shanxi	0/0 (0.0%)	0/11 (0.0%)	0/11 (0.0%)
Ningxia	0/0 (0.0%)	0/8 (0.0%)	0/8 (0.0%)
Gansu	2/10 (20%)	0/12 (0.0%)	2/22 (9.1%)
Qinghai	0/0 (0.0%)	3/6 (50.0%)	3/6 (50%)
Xinjiang	1/2 (50%)	0/0 (0.0%)	1/2 (50%)
Shandong	1/3 (33.3%)	0/0 (0.0%)	1/3 (33.3%)
Total	14/51 (27.4%)	106/571 (18.6%)	120/622 (19.3%)

Figure legends:

Figure 1 Size and amino acid sequence differences of 90 Cap proteins obtained in this study. Bars indicating the Cap proteins are filled according to the phylogenetic clade they occupy. The upper bar represents the consensus of the known PiCV Cap proteins; Amino acids are numbered from 1 to 277. Gaps indicate the location of deleted (compared to the consensus sequence) amino acids. The number on the right is the length of Cap proteins.

Figure 2 Comparison of the amino acid sequence conservation at a particular position between the 90 identified PiCV strains and the PiCV reference strains available in GenBank. A: The 90 identified PiCV strains from the racing pigeons in China. B: The 54 Chinese PiCV reference strains available in GenBank. C: The 72 PiCV reference strains available in GenBank from other countries. Note. The GenBank accession no. of the 54 Chinese PiCV reference strains, 72 PiCV reference strains from other countries, and the 90 identified PiCV strains was shown in Table S1 and S2.

Figure 3 Divergence analysis of Cap proteins of the 90 identified PiCV strains. A: The 90 identified PiCV strains from the racing pigeons in China. B: The 54 Chinese PiCV reference strains available in GenBank. C: The 72 PiCV reference strains available in GenBank from other countries. The unique amino acid mutations found in this study are indicated in red. Note. All PiCV reference strains were the same to the Figure 2.

Figure 4 Phylogenetic analysis of PiCV strains based on 90 *cap* gene sequences identified in this study and 126 *cap* gene sequences available in GeneBank. A maximum likelihood phylogenetic tree was automatically constructed with 1000 bootstrap replicates and the remaining default parameters in the MEGA 5.0 software, and the Interactive Tree Of Life (iTOL) software was used for the display and annotation of the phylogenetic tree. Labels at branch tips refer to the strain name and GenBank accession number. Red taxa highlight the 90 PiCV genome sequences isolated from the racing pigeons geographically located in China.

Figure 5 Recombination breakpoint distribution plots for 67 PiCV full genome sequences identified in this study and 113 PiCV full genome sequences available in GenBank. The red and blue areas of the plots indicate recombination breakpoint hot-spots and cold-spots, respectively. The dark and light grey areas represent the 95% and 99% confidence intervals of the expected degrees of breakpoint clustering under random recombination, respectively.

Hosted file

Figure_1.docx available at <https://authorea.com/users/374251/articles/524810-molecular-detection-and-phylogenetic-analysis-of-pigeon-circovirus-from-racing-pigeons-in-china-during-2016-2019>

Hosted file

Figure_2.docx available at <https://authorea.com/users/374251/articles/524810-molecular-detection-and-phylogenetic-analysis-of-pigeon-circovirus-from-racing-pigeons-in-china-during-2016-2019>

Hosted file

Figure_3.docx available at <https://authorea.com/users/374251/articles/524810-molecular-detection-and-phylogenetic-analysis-of-pigeon-circovirus-from-racing-pigeons-in-china-during-2016-2019>

Hosted file

Figure_4.docx available at <https://authorea.com/users/374251/articles/524810-molecular-detection-and-phylogenetic-analysis-of-pigeon-circovirus-from-racing-pigeons-in-china-during-2016-2019>

Hosted file

Figure_5.docx available at <https://authorea.com/users/374251/articles/524810-molecular-detection-and-phylogenetic-analysis-of-pigeon-circovirus-from-racing-pigeons-in-china-during-2016-2019>

Hosted file

Table 1.docx available at <https://authorea.com/users/374251/articles/524810-molecular-detection-and-phylogenetic-analysis-of-pigeon-circovirus-from-racing-pigeons-in-china-during-2016-2019>