Phylogenetic and codon usage analysis for replicase and capsid genes of porcine circovirus 3

Xianglong Yu1 · Kuipeng Gao1 · Molin Pi1 · Huizi Li1 · Wenxia Zhong1 · Baojian Li1 · Zhangyong Ning1,2

Received: 14 May 2021 / Accepted: 26 July 2021 / Published online: 6 August 2021
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract
Porcine circovirus type 3 (PCV3) is a highly contagious virus belonging to the family Circoviridae that causes the severe dermatitis and nephropathy syndrome. To date, PCV3 has a worldwide distribution and bring huge economic losses to swine industry. Replicase (Rep) and capsid (Cap) are two major coded proteins of PCV3. Considering the large number of new PCV3 isolates were reported in the past few years and the research for the codon usage pattern of Rep and Cap genes was still a gap, phylogenetic and codon usage analysis of these two genes was performed. Phylogenetic analyses showed that Rep genes in PCV3a were dispersed with no clear clusters while corresponding sequences in PCV3b clustered into two groups and Cap genes clustered into distinct clades according to different genotypes. Relative synonymous codon usage (RSCU) analysis revealed that the codon usage bias existed and effective number of codon (ENC) analysis showed that the bias was slight low. ENC-GC3s plot indicated that mutational pressure and other factors both played a role in PCV3 codon usage and neutrality plot analysis showed that natural selection was the main force influencing the codon usage pattern. The results presented here provided the important basic data on codon usage pattern of Rep and Cap genes, and a better understanding of the evolution and potential origin of PCV3.

Keywords PCV3 · Cap · Rep · Phylogenetic analysis · Codon usage bias

Introduction
PCV3 is an emerging circovirus which spread among a wide range of species and several syndromes were caused by this virus in pigs (Kedkovid et al. 2018). To date, PCV3 has spread worldwide with highly positive rate (Bera et al. 2020; Kwon et al. 2017; Yuzhakov et al. 2018). PCV3 is a non-enveloped, single-stranded circular DNA virus, which is the member of genus Circovirus in the family Circoviridae (Palinski et al. 2017; Phan et al. 2016). The length of genome of PCV3 is 2000 nucleotides containing three open reading frames (ORFs) (Saraiva et al. 2018). ORF1 and ORF2 are two major open reading frames that encoded replicase (Rep) and capsid (Cap) proteins in opposite directions, respectively. Replicase protein is associated with virus replication and capsid protein is the major structural protein containing immunologically epitopes associated with the viral entry and neutralization (Klaumann et al. 2018; Mankertz et al. 2004).

The research of codon usage bias is an important field, which can be of benefit to molecular biology or genetics, such as determining the origin and evolution of species and predicting the gene expression level (Butt et al. 2014; Yu et al. 2021a). Although several previous studies reported the codon usage bias of other Circovirus, such as PCV1, PCV2 (Chen et al. 2014) and duck circovirus (Xu et al. 2015), there are few reports on PCV3. Here, a comprehensive codon usage analysis using replicase and capsid genes obtained from 590 PCV3 complete sequences was performed to explore the mechanism of codon distribution and evolution of the virus.
Materials and methods

Sequence data

The complete genomic sequences of all 590 PCV3 isolates were obtained from the GenBank database (http://www.ncbi.nlm.nih.gov) until 10 February 2021. Some strains that without the collection date, country or host were excluded. Four genotypes of PCV3, including PCV3a-1, PCV3a-2, PCV3a-IM and PCV3b, were distinguished by unique epitope sequences and amino acid sites 122, 320, 323, 373 and 446 (Li et al. 2018a). The detailed information of the selected PCV3 strains was provided in Supplementary Table S1. The individual Rep and Cap genes were cut from every complete genomic sequence for next analysis.

Recombination and phylogenetic analysis

Potential recombination events were detected using RDP4 software by the BOOTSCAN, 3SEQ, CHIMAERA, MAX-CHI, GENECONV, SISCAN, and RDP programs (Martin et al. 2015). The type of sequences was set to circular and the P value was set to 0.05. Genomes were considered to be recombinants if a recombination event was supported by at least three of the above all algorithms (P value < 0.01) or the recombination score was above 0.6 (Liu et al. 2019; Wang et al. 2017). All Rep and Cap coding sequences were aligned by ClustalW and phylogenetic analysis was carried out by the neighbor joining (NJ) method with MEGA7 software.

Nucleotide composition analysis

The frequencies of A, U, C, G and G+C content were analyzed using Bio-edit. The frequency of nucleotide compositions at the third position (A3s, C3s, U3s and G3s) of synonymous codon was calculated by CodonW 1.4.4. GC content at third codon positions (GC3s) of synonymous codon was also computed by CodonW. All detailed information was shown in Supplementary Table S2 and S3.

Relative synonymous codon usage (RSCU)

RSCU value represents a ratio between the observed frequency of one synonymous codon and expected frequency (codons for one amino acid used equally), which is the important measure to standardize the codon usage bias. The equation of RSCU index is as follows:

\[
RSCU = \frac{X_{ij}}{\sum_{i} n_{i}}
\]

where \(X_{ij}\) is the number of the \(i\)th codon for the \(j\)th amino acid, and \(n_{i}\) represents the degenerate number of a synonymous codon. The RSCU value of 1 represents no codon bias and equal usage for the amino acid. The RSCU value > 1.0 shows positive codon usage bias while the RSCU value < 1.0 shows negative codon usage bias. Especially, if RSCU is higher than 1.6 or less than 0.6, it is considered to be over-represented or under-represented (Yu et al. 2021b). RSCU value of each codon was computed by CodonW and the RSCU values of the host (Swine) were downloaded from the codon usage database (http://www.kazusa.or.jp/codon/).

Effective number of codon (ENC) and ENC-plot analysis

The degree of codon usage bias can be quantified by the ENC analysis (Liu et al. 2010). ENC values is used to quantify and reflect the extent of preference of synonymous codons. ENC values is an absolute measure which ranges from 20 to 61 (He et al. 2019). The value of 20 indicates the codon bias is at a maximum, while the value of 61 indicates no bias. In general, if the ENC value ≤ 35, it is considered that the coding sequence has a significant codon usage bias. ENC values of Rep and Cap genes of each PCV3 sequence were calculated using codonW program.

ENC-plot analysis was used to detect the influencing factors of codon usage variation. ENC-GC3s plot was generated using GraphPad Prism 8. Expected ENC values for each GC3s was calculated using the following equation:

\[
ENC_{expected} = 2 + s + \left(\frac{29}{s^2 + (1-s)^2}\right)
\]

where \(s\) is the number of the GC3s value. If the codon usage is only constrained by the mutation pressure, points will be on or around the expected curve. Otherwise observed ENC values will lie far lower than the expected curve if several factors constrain the codon usage (Bera et al. 2017).

Principal component analysis (PCA)

PCA is a multivariate statistical method, which is performed by analyzing the relationship between variables and samples to identify major variation trends (Nasrullah et al. 2015). PCA was performed with SPSS software (Version 22) to explore the trends of codon usage pattern of Rep and Cap genes among different PCV3 strains. In detail, the RSCU values of two genes in PCV3 strains were distributed into a 59-dimensional vector corresponding to the 59 synonymous codons (excluding the codons of AUG, UGG and the three stop codons), and then they were transformed into uncorrelated variables (principal components). Among them, the first two axes are accounting for most of the component influencing the codon usage variation among genes. PCA
plots were constructed with the first two axes and the figures were drawn by Graph Pad Prism 8.0.

**General average hydropathicity (Gravy) and aromaticity (Aroma)**

Gravy and Aroma are two major indices on translation and natural selection and the values respectively show the frequencies of hydrophobic and aromatic amino acids (Xu et al. 2015). CodonW 1.4.4 program was used to calculate the two values.

**Neutrality plot analysis**

Neutrality plot analysis was performed to explore the effects of natural selection and mutation pressure on codon usage by plotting the GC12s against GC3s (Yu et al. 2021b). Each point represented Rep or Cap gene of an independent PCV3 strain and the regression line was plotted. If the regression curve lay near the diagonal (slope = 1), it was considered that mutation pressure was the dominant force for the codon usage bias with weak external selection pressure. Alternatively, natural selection was the main role in shaping codon usage if the slope of the regression curves tended to 0. The figure of neutrality plot was drawn using Graph Pad Prism 8.0.

**Correlation analysis**

All correlation analyses were performed by Karl Pearson’s method among the nucleotide compositions, codon compositions, Gravy, Aroma or principal axes.

**Results**

**Recombination and phylogenetic analysis**

Four recombination events of complete genomic sequences, containing MH229786, MF318448, MT183690 and MF405276, were detected by the RDP software. Take out those recombinant isolates and surplus coding sequences were carried on to the next analysis. Phylogenetic analyses showed that Rep genes in PCV3a genotypes were dispersed with no clear clusters while corresponding sequences in PCV3b clustered into two groups (Fig. S1) and Cap genes clustered into distinct clades according to different genotypes (Fig. S2).

**Nucleotide compositions of the Rep and Cap genes**

The nucleotide content and codon usage composition of coding sequences of Rep and Cap genes were calculated. For Rep gene, the mean values of nucleotide from high to low were G% (33.58 ± 0.12), U% (26.02 ± 0.14), A% (24.33 ± 0.12), C% (16.07 ± 0.15) and the mean values of the codon composition at the third position from high to low were G3s (46.62 ± 0.38), U3s (40.11 ± 0.50), A3s (22.20 ± 0.37), C3s (17.85 ± 0.45). For Cap gene, the mean values of nucleotide from high to low were A% (32.37 ± 0.23), C% (28.44 ± 0.27), U% (20.39 ± 0.20), G% (18.80 ± 0.24) and the mean values of the codon composition at the third position from high to low were C3s (43.89 ± 0.77), A3s (38.79 ± 0.92), U3s (26.12 ± 0.61), G3s (12.29 ± 0.90). Beyond that, the mean compositions of AU% were both above the GC% for two genes.

**RSCU analysis and ENC analysis**

To determine the codon usage bias of Rep and Cap genes and the effect of the host (swine), RSCU values of 59 synonymous codons were calculated and compared with the RSCU values of swine (Table 1). The results indicated that ten preferred codons were demonstrated to be common between swine and Rep or Cap gene in PCV3. Among them, three were commonly used between the host and both two genes. Nine codons of Rep gene were over-represented (mean RSCU value > 1.6) and fourteen codons were under-represented (mean RSCU value < 0.6). Meanwhile, up to twelve codons of Cap gene were over-represented and twenty-two codons were under-represented. Moreover, the number of A/U-ended preferred codons was the same as G/C-ended for both Rep and Cap genes. The ENC values of Rep gene ranged from 50.23 to 55.88, and the mean value of ENC was 54.23 ± 0.56 (mean ± SD). The ENC values of Cap gene ranged from 43.20 to 51.30, and the mean value of ENC was 44.97 ± 0.83. The detailed information of ENC values was shown in Supplementary Table S2 and S3.

**ENC-plot analysis and correlation analysis**

ENC-GC3s plot was generated to investigate the mutational pressure in shaping codon usage bias. As shown in Fig. 1, all points were lower than the standard curve no matter Rep or Cap gene. To determine the effect of mutational pressure on the codon usage, the analyses between ENC value and each nucleotide content were calculated and indicated remarkable correlations with the P values much below 0.01 (Table 2). Furthermore, the correlation analyses between nucleotide composition (A, T, G, C, and GC) and codon compositions (A3s, T3s, C3s, G3s, and GC3s) showed that most of them had a significant correlation between each other except for U with A3s content (Rep) and C with A3s content (Cap) (Table 2).
Principal component analysis (PCA)

The principal component analysis was carried out on the RSCU values to detect the variations of codon usage and the PCA plot of the first and second axes was drawn against each other according to the different countries. As shown in Fig. 2a–d, all points divided by years and continents were dispersed and no clear trend could be observed. But, points of different genotypes were clustered into two groups (3a, 3b) for Rep gene (Fig. 1e) and three groups (3a-1and 3a-IM, 3a-2, 3b) for cap gene (Fig. 1f), which were also consistent with the genotypic classification based on predicted epitopes and amino acid analysis. Subsequently, correlation analyses between the first two axes and nucleotide compositions showed that most correlations were significant (Table 3).

The role of natural selection in the codon usage bias of Rep and Cap genes

To determine the forces of natural selection, the correlation analyses were performed for evaluating the relationship between Gravy, Aroma and A, C, G, U, GC, GC3s. As shown in Table 4, the majorities of correlations were significant whereas several correlations with Aroma of Cap gene were weak.

Neutrality plot analysis

The neutrality plot was constructed to determine the main factor shaping the codon usage pattern of Rep and Cap genes (Fig. 3). The results showed a slight negative correlation between the value of GC12s and GC3s ($r = -0.39$, $P < 0.0001$).

Table 1 Overall RSCU of the collected sequences of the PCV3 Rep and Cap genes

| Amino acid | Codon | RSCU/Rep gene | RSCU/Cap gene | RSCU/swine | Amino acid | Codon | RSCU/Rep gene | RSCU/Cap gene | RSCU/swine |
|------------|-------|---------------|---------------|------------|------------|-------|---------------|---------------|------------|
| Phe        | UUU   | 1.57          | 0.41          | 0.78       | Ser        | AGC   | 1.92          | 2.52          | 1.6238     |
|            | UUC   | 0.43          | 1.59          | 2.1232     | AGU        | 0.78  | 0.44          | 0.7713        |
| Leu        | CUA   | 0.01          | 2.00          | 0.3311     | UCA        | 0.34  | 0.04          | 0.7226        |
|            | CUC   | 0.60          | 2.46          | 1.3475     | UCC        | 0.42  | 2.14          | 1.5021        |
|            | CUG   | 1.75          | 0.95          | 2.6776     | UCG        | 1.76  | 0.01          | 0.3897        |
|            | CUU   | 0.60          | 0.58          | 0.6506     | UCU        | 0.78  | 0.85          | 0.9905        |
|            | UUA   | 0.60          | 0.00          | 0.3195     | UGU        | 0.86  | 0.50          | 1.2152        |
|            | UUG   | 2.45          | 0.01          | 0.6738     | UGU        | 1.14  | 1.50          | 0.7848        |
| Tyr        | UAC   | 0.57          | 1.03          | 1.2685     | Pro        | CCA   | 1.27          | 1.18          | 0.9423     |
|            | UAU   | 1.43          | 0.97          | 0.7315     | CCC        | 0.26  | 1.59          | 1.4563        |
| His        | CAC   | 1.49          | 2.00          | 1.2946     | CCG        | 1.22  | 0.44          | 0.5536        |
|            | CAU   | 0.51          | 0.00          | 0.7054     | CUC        | 1.25  | 0.79          | 1.0478        |
| Gln        | CAA   | 1.31          | 1.35          | 0.441      | Arg        | AGA   | 0.75          | 2.62          | 1.1155     |
|            | CAG   | 0.69          | 0.65          | 1.559      | AGG        | 2.01  | 1.04          | 1.2347        |
| Ile        | AUA   | 0.86          | 0.87          | 0.4208     | CGA        | 0.26  | 0.47          | 0.6065        |
|            | AUC   | 0.22          | 0.02          | 1.6697     | CGC        | 0.49  | 1.07          | 1.3105        |
|            | AUU   | 1.92          | 2.12          | 0.9095     | CCG        | 1.75  | 0.03          | 1.2888        |
| Asn        | AAC   | 0.80          | 1.61          | 1.2133     | CGU        | 0.75  | 0.76          | 0.444         |
|            | AAU   | 1.20          | 0.39          | 0.7867     | Thr        | ACA   | 0.31          | 1.43          | 0.9145     |
| Lys        | AAA   | 1.00          | 1.39          | 0.7603     | ACC        | 1.27  | 1.43          | 1.6803        |
|            | AAG   | 1.00          | 0.61          | 1.2397     | AGG        | 1.23  | 0.03          | 0.5725        |
| Val        | GUA   | 0.70          | 0.44          | 0.3385     | ACU        | 1.18  | 1.12          | 0.8327        |
|            | GUC   | 0.18          | 1.07          | 1.0646     | Ala        | GCA   | 0.69          | 0.57          | 0.7386     |
|            | GUG   | 1.03          | 0.01          | 2.0308     | GCC        | 0.69  | 1.25          | 1.801         |
|            | GUU   | 2.09          | 2.48          | 0.5661     | GCG        | 1.48  | 0.79          | 0.5057        |
| Asp        | GAC   | 0.33          | 1.65          | 1.1975     | GCU        | 1.14  | 1.38          | 0.9546        |
|            | GAU   | 1.67          | 0.35          | 0.8025     | Gly        | GGA   | 0.69          | 2.22          | 0.9117     |
| Glu        | GAA   | 0.64          | 1.98          | 0.7256     | GGC        | 0.37  | 1.32          | 1.4644        |
|            | GAG   | 1.36          | 0.02          | 1.2744     | GGG        | 1.82  | 0.45          | 1.0541        |
|            | GGU   | 1.12          | 0.01          | 0.5698     |            |       |               |               |

RSCU values of the preferred codon for Rep and Cap genes in PCV3 and swine are in italics
Fig. 1 ENC plots showing the relationship between the ENC values and GC3s. a Points for Rep gene classified by continent. b Points for Cap gene classified by continent. c Points for Rep gene classified by host. d Points for Cap gene classified by host. e Points for Rep gene classified by genotype. f Points for Cap gene classified by genotype. All the points were lower than the standard curve, indicating that mutational pressure and other factors both played a role in PCV3 codon usage.

Table 2 Correlation analysis of nucleotide composition and ENC

|         | A3s     | C3s     | G3s     | U3s     | GC3s    | ENC     |
|---------|---------|---------|---------|---------|---------|---------|
| PCV3 Rep gene |         |         |         |         |         |         |
| A       | 0.801** | −0.162**| −0.228**| −0.188**| −0.330**| 0.166** |
| C       | −0.229**| 0.829** | −0.222**| −0.499**| 0.616** | 0.455** |
| G       | −0.438**| −0.135**| 0.611** | −0.003  | 0.304** | 0.288** |
| U       | −0.053  | −0.630**| −0.118**| 0.704** | −0.653**| −0.367**|
| GC      | −0.522**| 0.627** | 0.252** | −0.436**| 0.726** | 0.218** |
| PCV3 Cap gene |         |         |         |         |         |         |
| A       | 0.737** | −0.195**| −0.576**| 0.066   | −0.630**| −0.099**|
| C       | 0.02    | 0.655** | −0.344**| −0.565**| 0.296** | −0.282**|
| G       | −0.575**| −0.152**| 0.836** | −0.119**| 0.540** | 0.365** |
| U       | −0.179**| −0.467**| 0.121** | 0.811** | −0.315**| 0.055** |
| GC      | −0.468**| 0.475** | 0.375** | 0.597** | 0.715** | 0.026** |

The numbers represent correlation coefficient “r” values (*P < 0.05, **P < 0.01)
for Rep gene; $r = -0.283$, $P < 0.0001$ for Cap gene). Slope of the regression line was $-0.103$ for Cap gene and only $-0.014$ for Rep gene, suggesting that natural selection was the main force while mutation pressure was a minor force for the codon usage pattern of PCV3.

**Discussion**

Porcine circovirus type 3 (PCV3), first recognized in USA in recent years, is an important pathogen for swine with...
severe dermatitis and nephropathy syndrome (Palinski et al. 2017). Previous phylogenetic analyses showed the PCV3 had genetic diversity (Chen et al. 2019; Saraiva et al. 2018; Tan et al. 2020), the corresponding analysis of this study also proved this. Although the phylogenetic data of PCV3 could help us to deeply understand this virus, a systematic analysis for the codon usage pattern is still needed. A previous study for host adaptability partly reflected the codon usage bias of PCV3 complete genomes with 52 strains collected up to 2017 (Li et al. 2018b). Considering many new PCV3 isolates were reported in the past four years and the codon usage pattern of Rep and Cap genes was still a gap, a comprehensive analysis was performed to explore the mechanism of codon distribution and procedure of evolution of PCV3.

In this study, several analyses were used to standardize the codon usage bias. Nucleotide composition demonstrated the codon usage bias existed in PCV3 genomes. RSCU analysis showed that more than half of all codons were over-represented or under-represented, which further revealed that the codon usage bias appeared and the virus had a relatively stable nucleotide composition of synonymous codons. Moreover, about half of preferred codons were common between host and PCV3 Rep or Cap genes, highlighting a coincident and antagonistic codon usage patterns compared with host, which was also consistent with the results of the previous study (Li et al. 2018b). For ENC analysis, all ENC values were higher than 35 which revealed a slight low codon usage bias for two genes. Significantly, the ENC mean value of Rep was higher than that of Cap, demonstrating that codon usage bias for Rep was more stable, which also reflected the phenomenon that mutations mainly existed in Cap. PCA is a method which can identify major variation trends. As shown in Fig. 2, points of different years or continents were dispersed on the plot but points divided by genotype were clustered relatively, reflecting the high genetic stability of PCV3 genomes.

Mutation pressure and natural selection are thought to be the two most important factors for codon usage bias. ENC-GC3s plot showed that all of the points were below

Table 3 Summary of correlation between the first two axes and nucleotide constraints

|                  | PCV3 Rep gene | PCV3 Cap gene |
|------------------|---------------|---------------|
|                  | Axis 1        | Axis 2        | Axis 1        | Axis 2        |
| A                | −0.154**      | −0.030        | −0.251**      | 0.345**       |
| C                | 0.072         | 0.075         | −0.372**      | −0.415**      |
| G                | 0.237**       | 0.093*        | 0.483**       | −0.133**      |
| U                | −0.159**      | −0.143**      | 0.258**       | 0.312**       |
| A3s              | −0.295**      | 0.018         | −0.588**      | 0.182**       |
| C3s              | 0.061         | 0.170**       | 0.232**       | −0.309**      |
| G3s              | −0.245**      | −0.078        | 0.592**       | −0.240**      |
| U3s              | 0.359**       | −0.154**      | −0.003        | 0.481**       |
| GC               | 0.239**       | 0.109**       | 0.026         | −0.483**      |
| GC3s             | −0.103**      | 0.120**       | 0.541**       | −0.443**      |
| ENC              | 0.133**       | −0.333**      | 0.079         | −0.025        |

The numbers represent correlation coefficient “r” values (*P < 0.05, **P < 0.01)

Table 4 Correlation analyses among AROMO, GRAVY, the first two axes, ENC, GC3s and GC

|                  | A       | C       | G       | U       | GC      | GC3s    |
|------------------|---------|---------|---------|---------|---------|---------|
| PCV3 Rep gene    |         |         |         |         |         |         |
| Gravy            | −0.292**| 0.049   | 0.128** | 0.083*  | 0.134** | 0.000   |
| Aromo            | 0.084*  | −0.235**| −0.04   | 0.215** | −0.215**| −0.108**|
| PCV3 Cap gene    |         |         |         |         |         |         |
| Gravy            | −0.179**| −0.383**| 0.171** | 0.508** | −0.196**| 0.182** |
| Aromo            | −0.078  | −0.023  | −0.022  | −0.092* | 0.007   | 0.038   |

The numbers represent correlation coefficient “r” values (*P < 0.05, **P < 0.01)

Fig. 3 Neutrality analysis performed by plotting the GC12s against GC3s. a The Neutrality analysis for Rep gene. b The Neutrality analysis for Cap gene. The regression line was represented by the black straight line. The regression equation was also showed on the plot.
the standard curve, indicating that mutational pressure and other factors both played a role in codon usage of PCV3. Particularly, points divided by continent were dispersed on the plot but points divided by hosts were relatively concentrated except swine, which implied that PCV3 originated from swine. The neutrality plot analysis is a widely used method to explore the effects of natural selection and mutation pressure. The results revealed that the relative constraint (natural selection) was 89.7% for Cap gene and 98.6% for Rep, suggesting that natural selection was the main force for the codon usage pattern of the two genes and the contribution of mutation pressure for Cap gene was greater than that for Rep.

In conclusion, a low level of codon usage bias was detected in Rep and Cap genes of PCV3, and natural selection was the primary force impacting the codon usage. The results provided the important basic data for better understanding the evolution and potential origin of PCV3.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12259-021-09816-0.

Authors' contributions Xianglong Yu and Kuipeng Gao: Data curation, Methodology, Investigation, Writing—original draft, Writing—review & editing. Molin Pi, Huizi Li, Wenxia Zhong and Baojian Li: Methodology, Investigation. Zhangyong Ning: Conceptualization, Supervision, Writing—original draft, Writing—review & editing.

Funding This work was supported by the Project of Swine Innovation Team in Guangdong Modern Agricultural Research System (2020KJ126) and Natural Science Foundation of Guangdong Province (2019A1515011735).

Data availability All data included in this study are available on request to the corresponding author.

Declarations

Consent for publication All authors gave their consent for research publication.

Conflict of interest The authors declare that they have no conflict of interest.

References

Bera BC, Virmani N, Kumar N, Anand T, Pavulraj S, Rash A, Elton D, Rash N, Bhatia S, Sood R, Singh RK, Tripathi BN (2017) Genetic and codon usage bias analyses of polymerase genes of equine influenza virus and its relation to evolution. BMC Genomics 18(1):652

Bera BC, Choudhary M, Anand T, Virmani N, Sundaram K, Choudhary B, Tripathi BN (2020) Detection and genetic characterization of porcine circovirus 3 (PCV3) in pigs in India. Transbound Emerg Dis 67(3):1062–1067

Butt AM, Nasrullah I, Tong Y (2014) Genome-wide analysis of codon usage and influencing factors in chikungunya viruses. PLoS ONE 9(3):e90905

Chen Y, Sun J, Tong X, Xu J, Deng H, Jiang Z, Jiang C, Duan J, Li J, Zhou P, Wang C (2014) First analysis of synonymous codon usage in porcine circovirus. Arch Virol 159(8):2145–2151

Chen Y, Xu Q, Chen H, Luo X, Wu Q, Tan C, Pan Q, Chen JL (2019) Evolution and genetic diversity of porcine circovirus 3 in China. Viruses 11(9):786

He Z, Gan H, Liang X (2019) Analysis of synonymous codon usage bias in potato virus M and its adaption to hosts. Viruses 11(8):752

Kedkovid R, Woonwong Y, Arunorat J, Sirisereeaw C, Sangpratum N, Lumyai M, Kesdangsakonwat S, Teankum K, Jittimanees S, Thanawongnuwech R (2018) Porcine circovirus type 3 (PCV3) infection in grower pigs from a Thai farm suffering from porcine respiratory disease complex (PRDC). Vet Microbiol 215:71–76

Klaumann F, Correa-Fiz F, Franko G, Sibilia M, Nunez JI, Segales J (2018) Current knowledge on porcine circovirus 3 (PCV3): a novel virus with a yet unknown impact on the swine industry. Front Vet Sci 5:315

Kwon T, Yoo SJ, Park CK, Lyoo YS (2017) Prevalence of novel porcine circovirus 3 in Korean pig populations. Vet Microbiol 207:178–180

Li G, He W, Zhu H, Bi Y, Wang R, Xing G, Zhang C, Zhou J, Yuen KY, Gao GF, Su S (2018a) Origin, genetic diversity, and evolutionary dynamics of novel porcine circovirus 3. Adv Sci (Weinheim) 5(9):1800275

Li G, Wang H, Wang S, Xing G, Zhang C, Zhang W, Liu J, Zhang J, Su S, Zhou J (2018b) Insights into the genetic and host adaptability of emerging porcine circovirus 3. Virulence 9(1):1301–1313

Liu X, Wu Č, Chen AY (2010) Codon usage bias and recombination events for neuraminidase and hemagglutinin genes in Chinese isolates of influenza A virus subtype H9N2. Arch Virol 155(5):685–693

Liu J, Zha Y, Li H, Sun Y, Wang F, Lu R, Ning Z (2019) Novel recombinant seneca valley virus isolated from slaughtered pigs in Guangdong Province. Virol Sin 34(6):722–724

Mankertz A, Caliskan R, Hattermann K, Hillenbrand B, Kurzendoerfer P, Mueller B, Schmitt C, Steinfeldt T, Finsterbusch T (2004) Molecular biology of Porcine circovirus: analyses of gene expression and viral replication. Vet Microbiol 98(2):81–88

Martin DP, Murrell B, Golden M, Khoosal A, Muhire B (2015) RDP4: Detection and analysis of recombination patterns in virus genomes. Virus Evol 1(1):evw003

Nasrullah I, Butt AM, Tahir S, Idrees M, Tong Y (2015) Genomic analysis of codon usage shows influence of mutation pressure, natural selection, and host features on Marburg virus evolution. BMC Evol Biol 15:174

Palinski R, Pinyero P, Shang P, Yuan F, Guo R, Fang Y, Byers E, Hause BM (2017) A novel porcine circovirus distantly related to known circoviruses is associated with porcine dermatitis and nephropathy syndrome and reproductive failure. J Virol 91(1):e01879-e01916

Phan TG, Giannitti F, Rossow S, Marthaler D, Knutson TP, Li L, Deng X, Resende T, Vannucci F, Delwart E (2016) Detection of a novel porcine circovirus PCV3 in pigs with cardiac and multi-systemic inflammation. Virol J 13(1):184

Saraiya GL, Vidigal PMF, Fietto JLR, Bressan GC, Silva Junior A, de Almeida MR (2018) Evolutionary analysis of Porcine circovirus 3 (PCV3) indicates an ancient origin for its current strains and a worldwide dispersion. Virus Genes 54(3):376–384

Tan CY, Opaskornkul K, Thanawongnuwech R, Arshad SS, Hassan L, Ooi PT (2020) First molecular detection and complete sequence analysis of porcine circovirus type 3 (PCV3) in Peninsular Malaysia. PLoS ONE 15(7):e0235832

Wang J, Ling J, Wang Z, Huang Y, Zhu J, Zhu G (2017) Molecular characterization of a novel Muscovy duck parvovirus isolate:
evidence of recombination between classical MDPV and goose parvovirus strains. BMC Vet Res 13(1):327
Xu Y, Jia R, Zhang LuY, Wang M, Zhu D, Chen S, Liu M, Yin Z, Cheng A (2015) Analysis of synonymous codon usage pattern in duck circovirus. Gene 557(2):138–145
Yu X, Liu J, Li H, Liu B, Zhao B, Ning Z (2021a) Comprehensive analysis of synonymous codon usage bias for complete genomes and E2 gene of atypical porcine pestivirus. Biochem Genet 59(3):799–812
Yu X, Liu J, Li H, Liu B, Zhao B, Ning Z (2021b) Comprehensive analysis of synonymous codon usage patterns and influencing factors of porcine epidemic diarrhea virus. Arch Virol 166(1):157–165
Yuzhakov AG, Raev SA, Alekseev KP, Grebennikova TV, Verkhovsky OA, Zaberezhny AD, Aliper TI (2018) First detection and full genome sequence of porcine circovirus type 3 in Russia. Virus Genes 54(4):608–611

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