Genomic Features and Evolution of the Parapoxvirus during the Past Two Decades

Xiaoting Yao [1] , **Ming Pang** [1], **Tianxing Wang** [1], **Xi Chen** [1], **Xidian Tang** [1], **Jianjun Chang** [2,3], **Dekun Chen** [1,*] **and Wentao Ma** [1,*]

[1] College of Veterinary Medicine, Northwest A&F University, Yangling 712100, China; yaoxiaoting@nwafu.edu.cn (X.Y.); pangming@nwafu.edu.cn (M.P.); 2017011072@nwafu.edu.cn (T.W.); 2017011069@nwafu.edu.cn (X.C.); tangxidian@nwafu.edu.cn (X.T.)

[2] College of Agriculture and Animal Husbandry, Qinghai University, Xining 810016, China; 2018060204@nwafu.edu.cn

[3] State Key Laboratory of Plateau Ecology and Agriculture, Qinghai University, Xining 810016, China

[*] Correspondence: cdk@nwsuaf.edu.cn (D.C.); mawentao@nwafu.edu.cn (W.M.)

Abstract: Parapoxvirus (PPV) has been identified in some mammals and poses a great threat to both the livestock production and public health. However, the prevalence and evolution of this virus are still not fully understood. Here, we performed an in silico analysis to investigate the genomic features and evolution of PPVs. We noticed that although there were significant differences of GC contents between orf virus (ORFV) and other three species of PPVs, all PPVs showed almost identical nucleotide bias, that is GC richness. The structural analysis of PPV genomes showed the divergence of different PPV species, which may be due to the specific adaptation to their natural hosts. Additionally, we estimated the phylogenetic diversity of seven different genes of PPV. According to all available sequences, our results suggested that during 2010–2018, ORFV was the dominant virus species under the selective pressure of the optimal gene patterns. Furthermore, we found the substitution rates ranged from 3.56×10^{-5} to 4.21×10^{-4} in different PPV segments, and the PPV VIR gene evolved at the highest substitution rate. In these seven protein-coding regions, purifying selection was the major evolutionary pressure, while the GIF and VIR genes suffered the greatest positive selection pressure. These results may provide useful knowledge on the virus genetic evolution from a new perspective which could help to create prevention and control strategies.

Keywords: parapoxvirus; *Poxviridae*; nucleotide composition; selection pressure; evolution

1. Introduction

Parapoxvirus (PPV) is a chordopoxvirus that causes non-systemic skin lesions in wild and domestic animals. PPV belongs to the *Poxviridae* family and contains orf virus (ORFV), pseudocowpox virus (PCPV), bovine papular stomatitis virus (BPSV), and parapoxvirus of red deer in New Zealand (PVNZ) [1–3]. The PPV genome is approximately 134–139 kb in size with an extremely high GC content, which goes up to 65% [4–6]. The genetic structures of PPVs exhibit a general pattern form of a central conserved region, with essential genes in regard to position, spacing, and orientation [5]. The terminal variable regions encode genes with important roles during viral infection such as the regulation of host immune response to infection and the determination of host range [5,6].

PPVs distributes globally, and disturbs the ruminant flocks in many continents [7–10]. This virus is known to infect mammals, such as goats, sheep, cattle, camels, reindeer, and other ruminants [7,11]. Human can be infected by PPVs, when directly exposed to skin lesions from infected mammals. Human PPVs cases are reported in a lot of regions of the world, including Europe, South America,

and South Africa [7,12–17]. Currently, many live PPVs vaccines are available to control the disease, but they may not eradicate this disease [18].

ORFV, the prototype species of PPVs can cause lesions in the lips and breasts of suckling and grazing animals with high morbidity and low mortality rates [6,19]. Ruminants can be infected with ORFV more than once, with a shorter period of time to recovery and less obvious pathological changes than the primary infection [20]. BPSV can infect cattle of all ages, however, clinical signs are usually observed in calves. Like ORFV, repeated infection of ruminants with BPSV commonly occurs, indicating that BPSV may not confer effective immune response [6]. Besides ORFV and BPSV, PCPV and PVNZ are both maintained in mammals and threaten the animal health [6,21–25].

Vaccines have been indicated to impose a powerful selection pressure on the viral evolution [26,27]. Viruses may acquire greater virulence through genetic variation and recombination, as a result of the rapid transmission in vaccinated populations. For instance, deletions in the terminal locations have been discovered for vaccinia virus (VV) [28–34], rabbitpox virus [35,36], cowpox virus [37], and monkeypox virus [38]. The transposition of duplicated sequences was also been noticed for VV [39], cowpox virus [40,41], rabbitpox virus [36], and monkeypox virus [42]. These punctuated antigenic changes led to virus escape from the host immunity that was caused by earlier infection or vaccination, thus allowing the virus to re-infect their hosts who were once immune to virus and forcing the reformulation of virus [21,26,28]. Viruses also mutate readily due to low accuracy when duplicating their genome, which may assist the viruses to evade host immune attack and evolve further. PPVs have large genome, nearly 134–139 kb in size, where many replication errors may occur. Together with the selection pressure of vaccines, PPVs are even more likely to mutate [43,44].

This work reported the population dynamics of all available PPVs regarding the evolutionary process and genetic diversity. On the basis of PPVs genomes, we performed phylogenetic analysis and genomic comparison to investigate the genomic features of viral complete sequences. From the dataset of PPVs genome sequences, the evolutionary rate was also considered by Bayesian methods. Additionally, we further explored the different genes that have played an important role in PPVs evolution by estimating the genetic diversity and the selection pressure in this study.

2. Results

2.1. G and C Nucleotides Were Highly Enriched in PPV Genomes

To analyze the basic characteristic of the virus genome, the nucleotide contents of PPV whole genomes were calculated. It was shown that the mean compositions (%) of G (32.16 ± 0.32) and C (32.12 ± 0.22) were significantly higher than A (17.95 ± 0.28) and T (17.77 ± 0.25) (Table S1, Figure 1A, *t*-test, $p < 0.01$). The mean percentages of G3 (36.92 ± 1.02) and C3 (37.12 ± 1.33) were also significantly higher than A3 (20 ± 1.02) and U3 (19.94 ± 0.89), highlighting that G- and C- ended codons are more likely to occur in the PPV whole genome sequences (Table S1, Figure 1B, *t*-test, $p < 0.01$). According to the species classification, it was important to mention that a similar trend of nucleotide composition was observed among different species, also suggesting higher G/C or G3/C3 contents in PPV genomes. However, there were significant differences in the nucleotide composition parameters trends among four PPV species, including BPSV, PCPV, ORFV, and PVNZ (Table S1, Figure 1C–D, *t*-test, $p < 0.01$). As shown in Figure 2, we further compared the nucleotide composition among four different PPV species. Strikingly, there were highly unusual patterns of nucleotide contents, and were significantly different from the nucleotide composition patterns in the various PPV species.
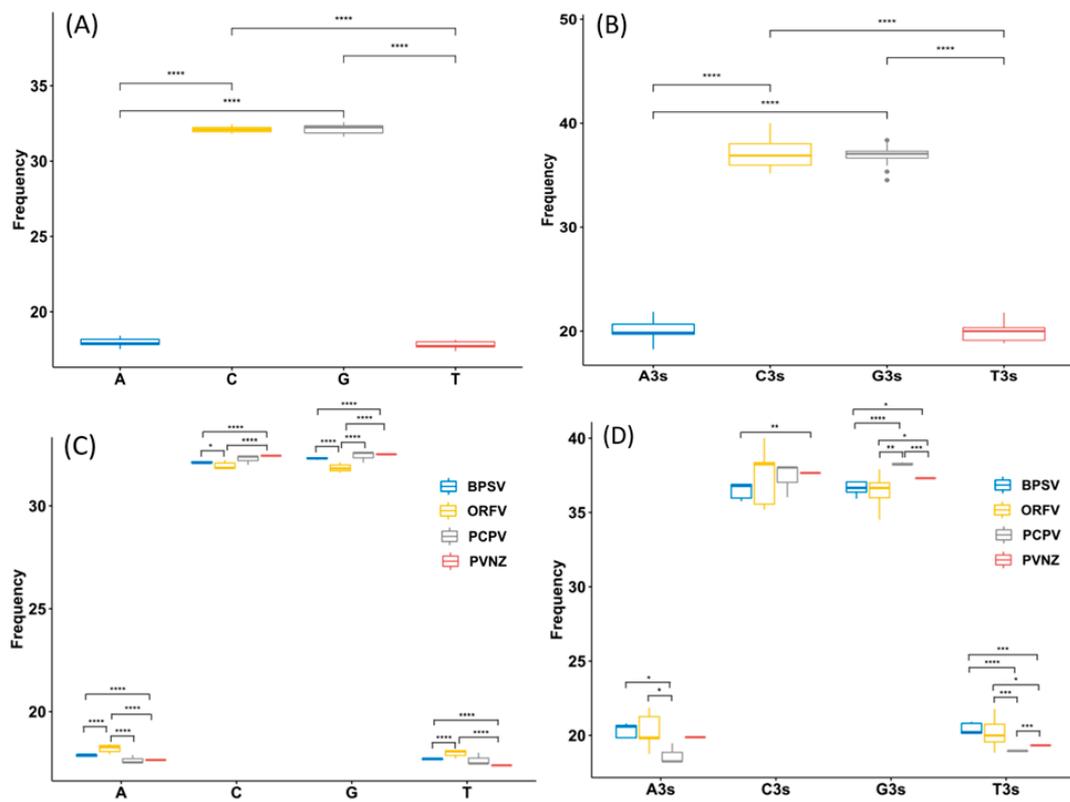
Figure 1. Nucleotide content distribution and composition. (**A**) The mean frequency for A, T, G, and C composition in 17 different parapoxvirus (PPV) sequences are shown, suggests higher GC content than AT on the whole. (**B**) The mean values of the nucleotide content frequency at the 3rd codon position. (**C**) Based on 4 various species of PPVs, analysis for the overall A, T, G, and C composition frequencies. (**D**) Based on 4 various species of PPVs, analysis for A, T, G, and C composition frequencies at the third codon position. Different asterisks represent statistical significance (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$).
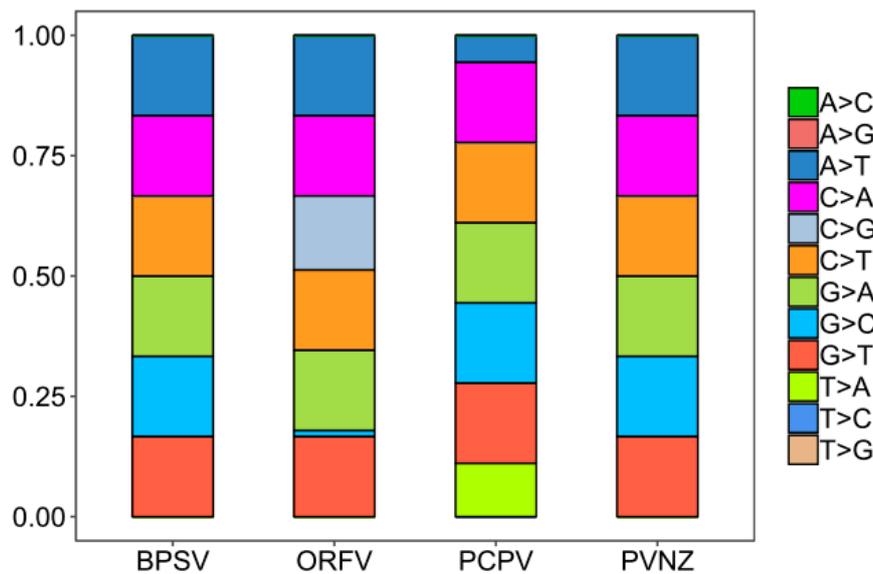
Figure 2. Nucleotide composition of PPVs for different species. The frequency of each observed nucleotide composition, reconstructed in BioEdit (v7.2.5), is shown for different species of PPVs.

2.2. Different Genome Structures Exist in Various PPVs

According to the phylogenetic relationship and genomic structure, the PPV genera belongs to *Poxviridae* and includes four species: ORFV, PCPV, BPSV, and PVNZ (Figure 3). The genome of PPV contains a large number of protein-coding regions which were encoding functional proteins, such as major envelope protein (B2L), chemokine binding protein (CBP), immunodominant envelope protein (F1L), GM-CSF/IL-2 inhibition factor (GIF), vascular endothelial growth factor (VEGF), viral interferon resistance protein (VIR), and late transcription factor (VLTF). Besides these major functional proteins, different PPVs encode some special accessory proteins, including Ankyrin protein (ANK), IL-10-like protein (vIL-10), and poxviral anaphase promoting complex regulator/ring H2 protein (PACR) (Figure 3). These results suggested that compared with other PPVs, VRGF genes existed in the left terminal region of BPSV genomes. It was also of note that PCPV genomes contained one extra VEGF gene at the left end and lacked the vIL-10 gene, while vIL-10 replaced VEGF in the left end of GQ329669.1/PCPV genome. Additionally, PPVs contained more than one ANK genes that may be involved in host range.
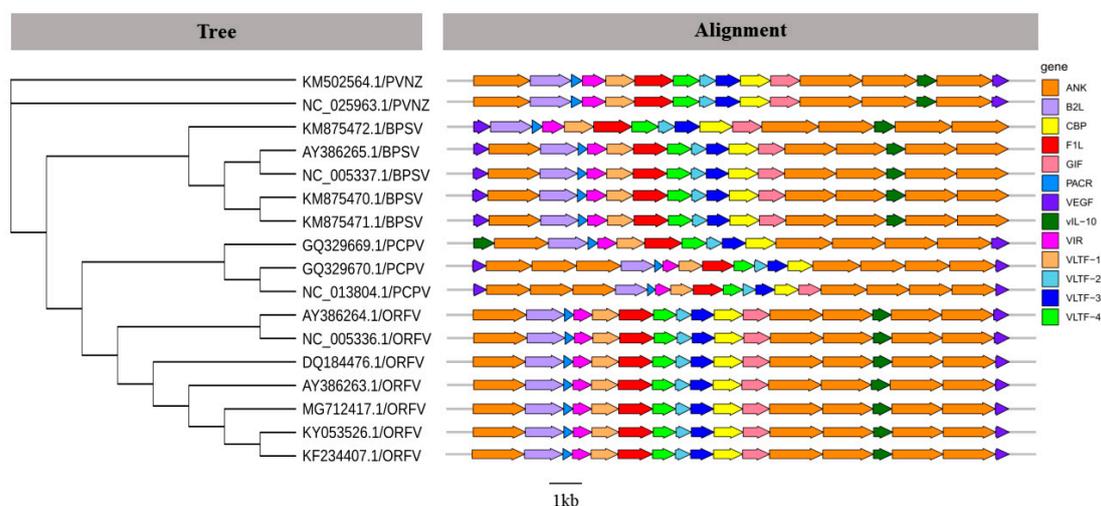
Figure 3. Phylogenetic analysis and genome comparison among different species of PPVs. According to Table S2 complete genome sequences of PPVs, Maximum Likelihood (ML) phylogenetic tree was reconstructed in phyML (v3.0) and tested with 1000 bootstraps. The predicted genes are represented by block arrows showing the direction of transcription. The genes encoding for structural proteins and replication-associated proteins of PPVs are shown in different colors.
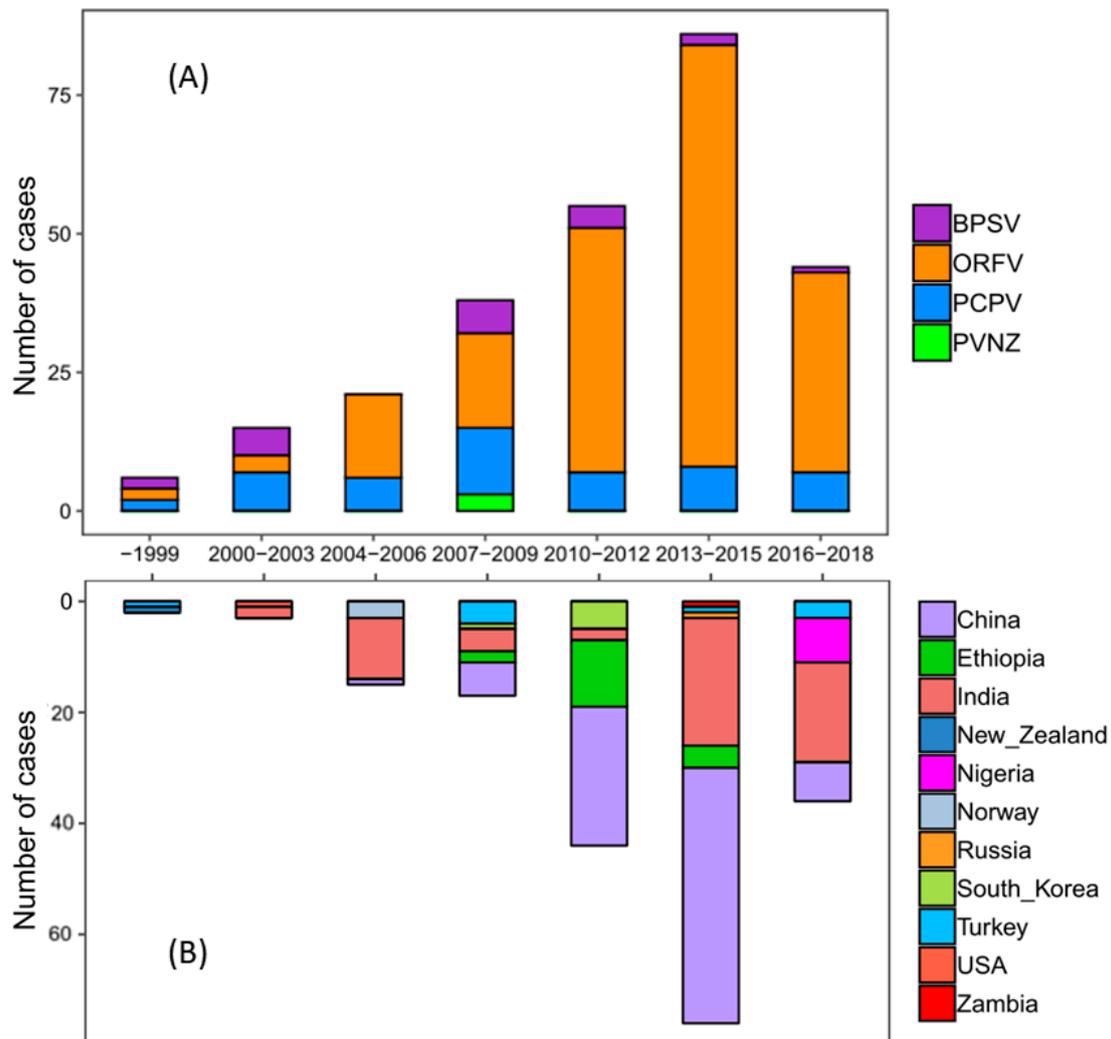
2.3. Population Dynamics of PPVs over the Past 20 Years

In order to explore the changes of PPV deposited sequences in the world during the past 20 years, 266 available complete sequences of PPV B2L genes were retrieved from GenBank in this study (Table S2). All B2L isolates were clustered into separate clades based on their phylogenetic tree, suggesting that ORFV was the major species, which circulated mainly in China and India (Figure 4A and Figure S1). Consistent with the pattern observed in the phylogenetic tree, the results of the case count also indicated that most of ORFVs were isolated from China and India, and their prevalence increased sharply during 2013–2015 (Figure 4B). Since 2010, ORFV has been the predominant species in circulation throughout China and India.

2.4. ORFV of Genera PPV Became Predominant Species in Recent Years

To describe the phylogenetic evolution of the prevalent PPVs, the phylogenetic trees were constructed for the seven major structural segments. Phylogenetic trees were partitioned into groups with a 20th percentile cutoff [45]. The PPV strains were not clustered exactly consistently in each tree of seven segments (Figure 5). Within each prevalent clade, a large number of ORFVs further formed a main group with few PCPV, BPSV, and PVNZ (Figure 5), demonstrating they may share a common

ancestor surviving a bottleneck event. Since 2010, the species diversity of PPVs was reduced with widespread outbreak in many countries (Figures 4 and 5), which can be reasonably considered by the high sequence similarity of PPVs isolated in the specific time span across broad regions.


Figure 4. Population dynamics of PPVs during the past 20 years. Case counts are presented in the plot. The X-axis represents the year intervals, and the Y-axis represents the number of cases. Different colors represent different species (**A**) and different collection countries of PPVs (**B**).

In order to confirm this hypothesis, we estimated the temporal dynamic of gene sequence similarities of PPVs by collecting all available PPV sequences during 2000–2018. Pairwise identity comparison suggested a decreased diversity in all these seven segments of PPVs, starting mainly in 2010–2018 (Figure 6). At nearly the same time point, the reduced sequence diversity of almost all gene segments suggested that the virus with great fitness was selected, acquired dominance, and led to a widespread outbreak throughout the world. As we know, this is the first discovery of such genetic bottleneck in the PPV genomes (Figure 6), despite the prevalence of PPV in the world for the past 20 years.
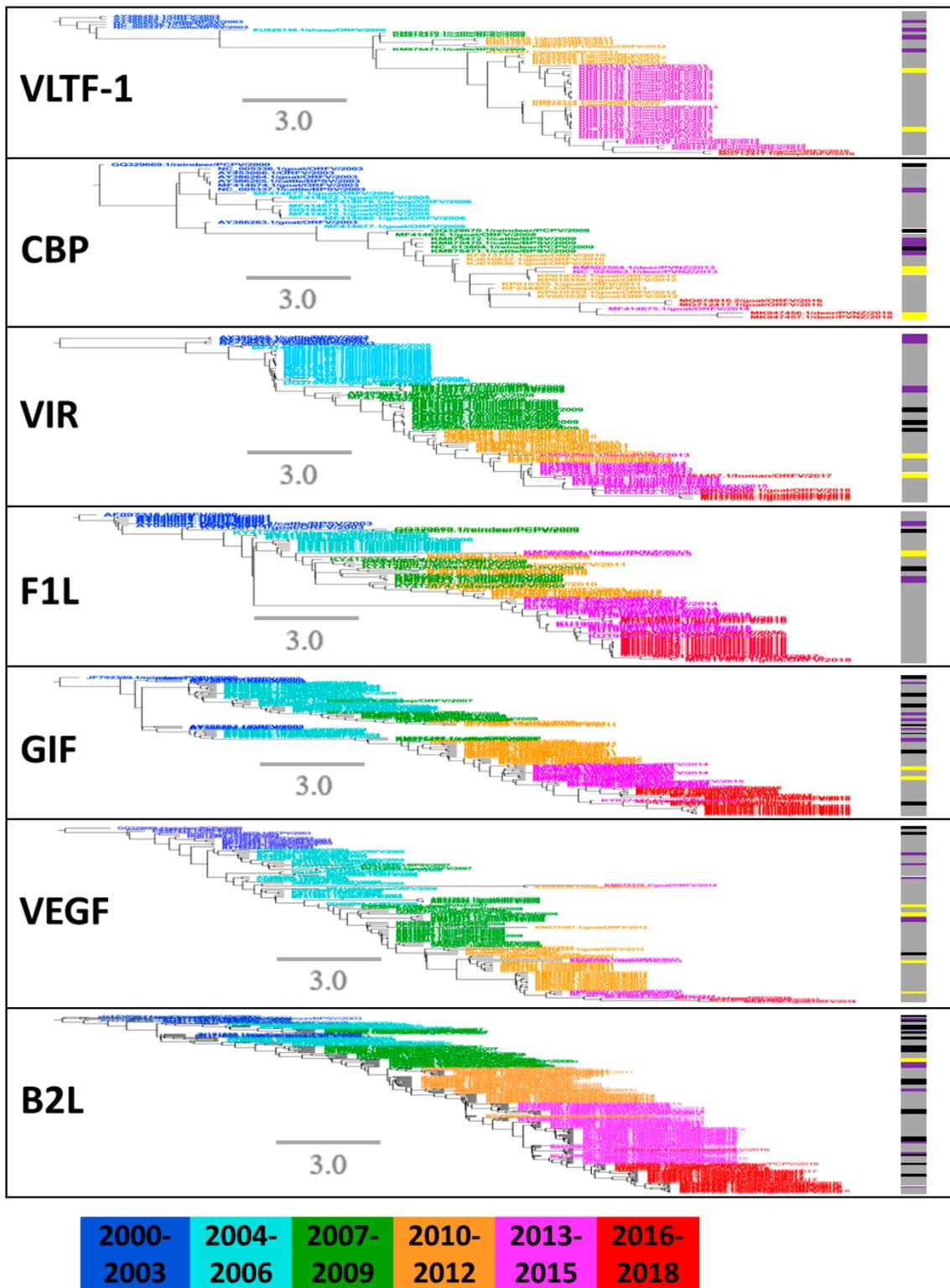
Figure 5. Phylogenetic analysis of seven structural genes with PPVs during 2000–2018. Clades with all of 2000–2018 viruses were fully shown. Color of line at right of each leaf node indicates year of isolation (see color bar). Timescale is in years. Vertical lines mark different species of PPVs: gray represents ORFV, purple represents BPSV, black represents PCPV and yellow represents PVNZ.
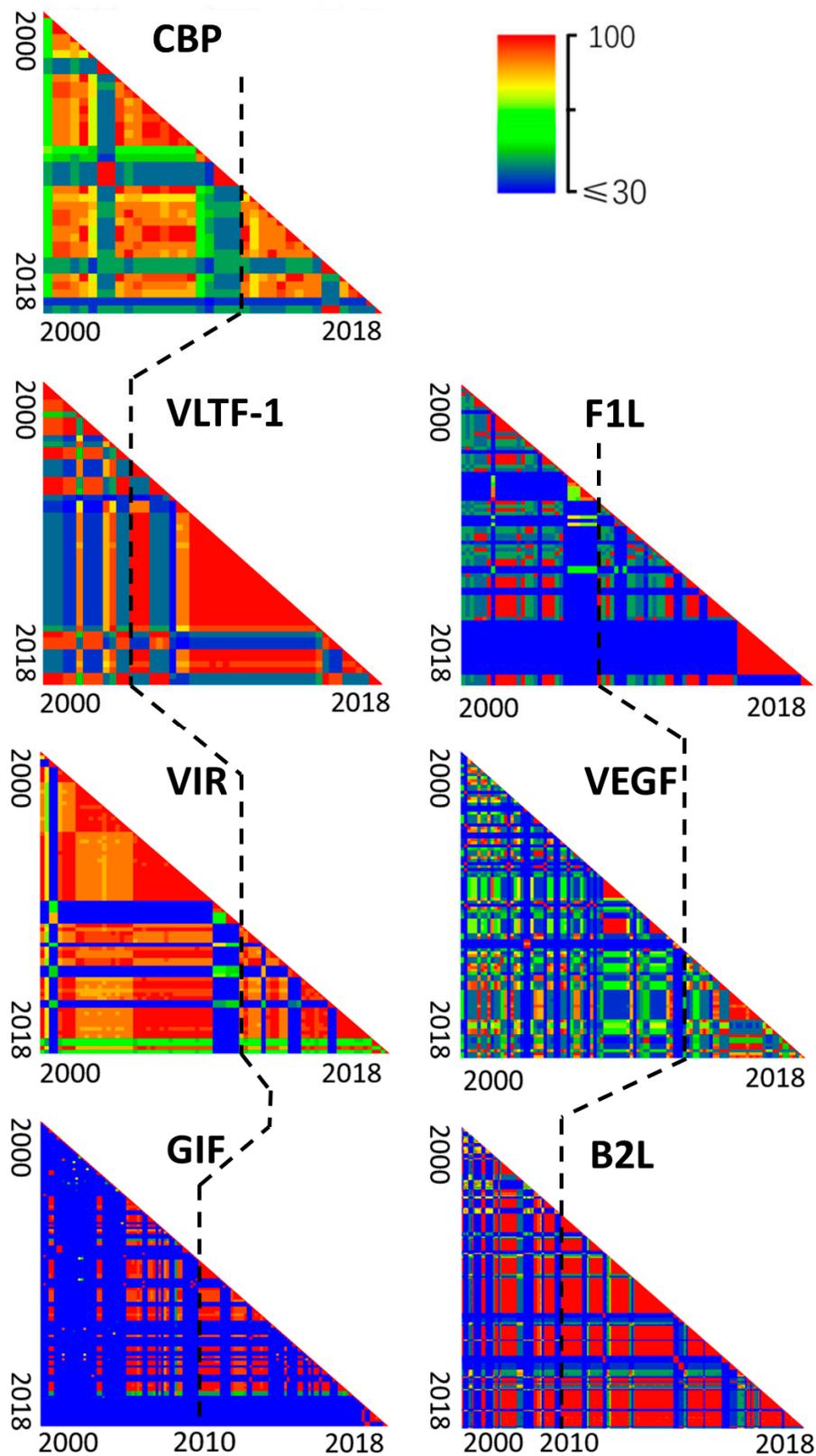
Figure 6. Decreased genetic diversity of the seven segments of PPVs during the widespread outbreaks in the world (2000–2018). Pairwise comparison of nucleotide sequences of PPVs was plotted as a heatmap. Viruses isolated from 2000 through 2018 were ordered by isolation date from left to right of the x axis and from top to bottom of the y axis. On the axes, the ticks indicated the isolation years. Black dotted line represented the year 2010. Color indicated identity levels from ≤30% (blue) to 100% (red).

2.5. Evolutionary Rates and Selection Analysis of Different Gene Sequences in PPVs

To understand the additional forces affecting PPV evolution, we analyzed the evolutionary rates and the average selective pressure of PPV strains whose isolation date were known. According to the different protein-coding gene sequences, a Bayesian coalescent approach was performed to infer the substitution rates of different segments. Based on the best-fit model, the Bayesian coalescent approach was performed to calculate the mean evolutionary rates of the seven gene segments, ranging from 3.56×10^{-5} to 4.21×10^{-4} substitutions per site per year (Table 1). The VEGF gene evolved at the slowest rate among the structural segments, with a substitution rate of 3.56×10^{-5} and the highest probability density (HPD) between 4.31×10^{-6} to 9.61×10^{-5}. The VIR gene evolved at the highest estimated mean rate of 4.21×10^{-4} (95% HPD: 1.28×10^{-4}–7.32×10^{-4}).

Table 1. Bayesian estimates of evolutionary rate of specific gene segments of *parapoxvirus*.

Gene	Evolutionary Rate (nt Substitutions per Site per Year)	95% HPD
B2L	6.34×10^{-5}	3.89×10^{-9}–1.52×10^{-4}
CBP	3.51×10^{-4}	2.39×10^{-5}–7.06×10^{-4}
F1L	2.79×10^{-4}	1.39×10^{-4}–4.31×10^{-4}
GIF	1.29×10^{-4}	3.53×10^{-5}–3.07×10^{-4}
VIR	4.21×10^{-4}	1.28×10^{-4}–7.32×10^{-4}
VEGF	3.56×10^{-5}	4.31×10^{-6}–9.61×10^{-5}
VLTF$-$1	4.45×10^{-5}	1.85×10^{-5}–7.22×10^{-5}

HPD, highest probability density.

The selective pressure for each segment was performed to suppose a constant substitution rate of all sites and provided average values across all sites. The GIF and VIR genes were found to be more variable, while the VLTF-1 gene was more conserved. On the basis of the Bayesian analysis, likelihood ratio tests were performed to compare the neutral models and the positive selection models. To estimate the selective pressure on each protein-coding segment, we applied a model (M3) for the differences among non-synonymous and synonymous rates (dN/dS) of amino acid residues, and several neutral models (M0, M1, and M7) with proportions of selected codons (M2 and M8) (Table 2). The log likelihood values showed that positive selection models (M2, M3, and M8) were more fitted in each gene regions than neutral models (M0, M1 and M7). The nested comparisons were also performed between the neutral models and the positive models, such as M0 vs M3, M1 vs M2 and M7 vs M8, and determined the better fitness of positive models, indicating that except for the VLTF-1, positive selection pressure occurred at different gene segments within the genome during the PPVs evolution process (chi-square test, $p < 0.05$) (Table 2).

Table 2. Likelihood ratio test was performed to compare positive selection models vs neutral models for different *parapoxvirus* gene segments.

Gene	M0 (ω)	Likelihood Test (2ΔlnL)		
		M0 vs. M3	M1a vs. M2a	M7 vs. M8
B2L	0.496	37.600 *	7.855 *	14.975 *
CBP	0.712	714.169 **	158.452 **	465.606 **
F1L	0.810	16.101 *	7.855 *	4.673 *
GIF	1.446	54.257 **	45.127 ***	45.116 ****
VIR	1.537	27.370 *	32.219 ***	39.494 ***
VEGF	0.801	64.918 ***	17.004 *	25.106 *
VLTF-1	0.015	1.062	1.766	0.001

Positive selection models: M2a, M3, M8; neutral models: M0, M1a, M7. Positive selection pressure was defined as dN/dS >1, where dN = number of non-synonymous substitutions per site and dS = number of synonymous substitutions per site. ω, dN/dS ratio; L, log likelihood; M0, one ratio model; M1, nearly neutral model; M2, positive selection model; M3, discrete model; M7, beta model; M8, beta and ω model. * $p \leq 0.05$; ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$.

3. Discussion

PPV infections normally occur in animals and even in humans worldwide [7,12–16]. However, comprehensive studies on PPVs are still lacking. In the present study, available whole genome sequences and different structural protein-coding sequences were both utilized to investigate the structural and evolutionary genomics of PPVs. To identify the common features and differences of PPVs, it is important to analyze genomic patterns of all representative sequenced PPV genomes. To start with, the general nucleotide contents were calculated for the genomes of different PPV species have been shown in Table S1. The results suggested that the PPV coding sequences were GC-rich, especially on the third location of synonymous codons. It has been known that viruses can take on a wider range of GC compositions than other organisms [46,47]. Even within the same family, viruses have similar replication or life-cycle strategies, but may show different GC frequencies [48]. Although high GC amplification could have fixed erroneous base calls in the sequencing in an unpatterned and uneven manner, which might influence sequences reported and downstream analyses, we retrieved enough sequences and got the same conclusion with prior researches, which all indicated that the biological relevance of G/C richness and frequency [6,49,50].

These PPVs, although related, could be divided into four subgroups according to the phylogenetic tree (Figure 3), which contained ORFV, PCPV, BPSV, and PVNZ, respectively. It demonstrated a divergence among species affiliations based on complete genome sequences of PPVs. To extend the genomic description of these four subgroups of PPVs, we added the characterization of PPV genomes. PPV with large genomes (134–139 kb) encodes nearly complete replisomes, which may englobe more than one virus [4,5]. It has been suggested that VEGF plays an important role in PPV pathogenesis related to the vascularization process and epidermal lesion proliferation [6,51]. Consistent with previous studies, we identified that VEGF was located in the left terminal region of BPSV contrasting the right terminal position of ORFV, PCPV, and PVNZ VEGFs [6,52], which indicated that the divergence of BPSV VEGF may have distinct functions or binding specificities compared with other VEGFs. Notably, PCPV was divergent from other PPVs with a deficiency of vIL-10 or left terminal vIL-10. As a pleiotrophic cytokine, vIL-10 can generate immunosimulatory/immunosuppressive effects on cells. Additionally, vIL-10 is an important anti-inflammatory cytokine to inhibit non-specific immunity and plays the particular role in macrophage and Th1 effector [53,54]. Overall, these divergent genomic features may reflect the specific adaptation to their natural hosts [6].

In addition to the genomic characteristics of PPVs, the phylogenetic diversity of the segmented genome of PPV may be a crucial contributor to understanding the temporal and spatial patterns of viral outbreaks [55,56]. As shown in this study, the proportion of ORFV strains sequenced and available indicates a sharp increase in the past 20 years. Our results suggested that PPVs evolved under mutation pressure and natural selection over 20 years of prevalence in animals, and the predominant species, ORFV, was selected recently. Since 2010, ORFV has rapidly increased in many countries, especially in China and India, and is now the predominant PPV species in farmed livestock [57–62]. It has been indicated that under selective pressure, ORFV eventually became the dominant species during 2010–2018. The reduced sequence diversity of almost all gene segments at nearly the same time point indicates that the virus with greatest fitness was selected and resulted in a widespread outbreak throughout China and India.

Furthermore, according to the supposition of rate constancy, the differences in viral isolation dates can provide us with knowledge about the substitution rate of molecular evolution. The previous study has suggested that the magnitude of evolutionary change may accumulate since the isolation date [63]. A Bayesian coalescent approach suggested the mean substitution rates were between 3.56×10^{-5} to 4.21×10^{-4} in different gene segments. Besides, the highest substitution rate of the PPV VIR gene reflected its variability in viral synthesis and replication process than other gene segments, which may be caused by the long-lasting interaction with the immune response.

When we analyzed the selection profiles of PPV protein-encoding genes, the overall rates of dN/dS were mostly less than 1, demonstrating that purifying selection was the major force to drive the

evolution of PPV genes. However, there were clear differences in this selection profiles of various genes. Among these protein coding genes, GIF and VIR were under greater positive selection than other gene segments, which were likely interacting with proteins of the immune system [18,64,65]. The variability of genes was constrained by the positive selection during adaptation, which is important for virus replication [66]. Our results suggested that VLTF-1 may suffer from greater purifying selection in the evolutionary process of PPVs. Additionally, the angiogenic factor, VEGF which has a huge influence on embryonic development or tumor neovascularization by binding to its receptor tyrosine kinases is reasonably conserved during the viral evolution in host populations. These results suggested that PPV genes are evolving at the rapid rate under selection pressure, which may improve viral fitness in their infected hosts. Therefore, it is likely that the PPV genomes have been reshaped by the animal immune system, although genetic mutation or recombination contribute to the natural evolution of PPVs.

4. Materials and Methods

4.1. PPVs Sequences

PPVs complete genome sequences and nucleotide sequences of the B2L, CBP, F1L, GIF, VIR, VEGF, and VLTF-1 genes were downloaded from GenBank (www.ncbi.nlm.nih.gov/genbank/). To explore the emergent time of the earliest PPV isolate, the sequences with an unknown collection date were excluded. The host, collection date, and sampling location of each PPV isolates was determined either from GenBank, or were obtained from the related publications. This led to a dataset of 17 complete genome sequences and 776 gene sequences, including 266 B2L genes, 38 CBP genes, 83 F1L genes, 133 GIF genes, 81 VIR genes, 122 VEGF genes, and 51 VLTF-1 genes, respectively (Table S2). Partial sequences from the aforementioned genes were obtained from the complete genomes available at GenBank. The resulting sequences were aligned using MUSCLE v3.7 (http://www.drive5.com/muscle/) [67].

4.2. Gene Structure and Nucleotide Composition

The gene structure of PPV was visualized by comparing their genomic nucleotide sequences using the R program "gggenes" project. Nucleotide compositional analysis of the PPVs complete genome sequences was measured using BioEdit v7.2.5 (https://bioedit.software.informer.com/7.2/) [68], including the frequencies of A, T, G, C, AT, and GC. The nucleotide occurrence frequencies of the third codon position (T3, G3, C3, and A3) of synonymous codons were measured by the CodonW v1.3 (http://sourceforge.net/projects/codonw).

4.3. Compositional Change of PPV Isolates over Time

During 1974–2018, 266 B2L genes of PPV strains were downloaded from GenBank and analyzed with the phylogenetic analysis by maximum likelihood program (PAML v4.9, http://abacus.gene.ucl.ac.uk/software/paml.html), with 1000 bootstrap replicates. An online tool, the Interactive Tree Of Life v2, was performed to design the tree [69]. The resulting isolates were composed of 4 species in PPVs population, including ORFV, PCPV, BPSV, and PVNZ. Isolates were assigned based on the PPVs species and the initial collection country. The variation number of PPV was calculated using statistical methods and the plot was drawn with the project ggplot2 in R program [70].

4.4. Evolution Substitution Rates

For all PPV genes, the molecular evolution rates were estimated by the MCMC program BEAST v2.4.8 (http://beast.community). This program gives a maximum-likelihood estimate of the evolution substitution rate, based on a model that considered a constant evolution rate (molecular clock). We also estimated the confidence intervals for all the parameters.

To obtain the best-fitting models for each viral gene, jModelTest v2.1.7 was performed by Akaike Information Criterion (AIC) (https://www.softpedia.com/get/Science-CAD/jModelTest.shtml) [71], which was estimated for each model tested according to the approach of Newton and Raftery [72].

Here, we used general time reversible model, a strict molecular clock, and a coalescent model with constant available population size, as performed in the BEAST package. Bayesian Markov chain Monte Carlo (MCMC) analysis was implemented with 10 million steps and was sampled every 10,000 steps with 10% burn-in. Tracer v1.6 was used to analyze the results of running BEAST. The effective sample size was greater than 200 for each parameter assessment in the MCMC analysis. Statistical uncertainty in the dataset was estimated in the 95% HPD values.

4.5. Genetic Diversity and Phylogenetic Analysis

The genetic diversity (π) of PPV was evaluated as average pairwise difference between viral sequences, using the Tamura–Nei nucleotide substitution models performed in MEGA 5 [73]. BioEdit was implemented to estimate the sequence identity for different gene segments [68], and the heatmap was plotted by R program "pheatmap"project. For each PPV gene, bayesian phylogenetic tree was inferred by Bayesian Evolutionary Analysis Sampling Trees Software (BEAST) [74]. For all analyses we performed a coalescent constant population model and the general-time reversible model of nucleotide substitution [75]. Based on the correlation between collection date and evolutionary distance for each gene, a strict molecular clock was selected, as evaluated by TempEst [76]. Here, MCMC analysis was run for 10 million steps, 10% of which was removed as burn-in and parameters sampled every 10,000 steps. The effective sample size for all the traces was greater than 200.

4.6. Selection Pressures in PPVs

Selection pressure was highlighted as the ratio between the average number of non-synonymous nucleotide substitutions (dN) and synonymous (dS) nucleotide substitutions per site (dN/dS) and was implemented in the CODEML program within the PAML software package. Besides, paired models of dN/dS were estimated among nucleotide sequences, containing M0 versus M3, M2a versus M1a, and M8 versus M7. We applied the likelihood ratio test to compare the neutral and positive selection models, and chi-square test to indicate the divergence between the compared models. We also calculated the dN and dS values respectively.

5. Conclusions

In conclusion, this study indicated that PPVs share several common features within different virus species. However, we noticed that GC contents were significantly different between ORFV and other PPV species. The genetic characteristics were also not similar among all PPVs, which may have resulted from the specific adaptation to their natural hosts. Furthermore, the evolutionary features of PPVs suggested that ORFV has been the predominant PPV species and increased rapidly in many countries since 2010. PPV isolates have been evolving at a sharp increased rate under the natural selection pressure. Combining genetic mutation of PPVs, positive selection may reshape the PPV genomes during the evolutionary process. The genetic analysis of PPVs could extend our knowledge of the mechanisms that promote virus evolution.

Author Contributions: D.C., J.C. and W.M. conceived and designed the experiments; X.Y. and M.P. performed all experiments. X.Y., T.W. and X.C. collected and analyzed the data. X.Y., M.P. and X.T. drafted the manuscript. All authors have read and agreed to the published version of the manuscript.

References

1. Becher, P.; König, M.; Müller, G.; Siebert, U.; Thiel, H.-J. Characterization of sealpox virus, a separate member of the parapoxviruses. *Arch. Virol.* **2002**, *147*, 1133–1140. [CrossRef] [PubMed]

2. Robinson, A.; Mercer, A.A. Parapoxvirus of Red Deer: Evidence for Its Inclusion as a New Member in the Genus Parapoxvirus. *Virology* **1995**, *208*, 812–815. [CrossRef] [PubMed]

3. Brown, D. The Encyclopedia of Virology plus. *Trends Biochem. Sci.* **1997**, *22*, 34–35. [CrossRef]

4. Wittek, R.; Kuenzle, C.C.; Wyler, R. High C + G content in parapoxvirus DNA. *J. Gen. Virol.* **1979**, *43*, 231–234. [CrossRef] [PubMed]

5. Buttner, M.; Rziha, H.-J. Parapoxviruses: From the Lesion to the Viral Genome. *J. Vet Med. Ser. B* **2002**, *49*, 7–16. [CrossRef] [PubMed]

6. Delhon, G.; Tulman, E.R.; Afonso, C.L.; Lu, Z.; De La Concha-Bermejillo, A.; Lehmkuhl, H.D.; Piccone, M.E.; Kutish, G.F.; Rock, D.L. Genomes of the Parapoxviruses Orf Virus and Bovine Papular Stomatitis Virus. *J. Virol.* **2004**, *78*, 168–177. [CrossRef] [PubMed]

7. Chakhunashvili, G.; Carlson, B.F.; Power, L.; Khmaladze, E.; Tsaguria, D.; Gavashelidze, M.; Zakhashvili, K.; Imnadze, P.; Boulton, M.L.; Johnson, L.R. Parapoxvirus Infections in the Country of Georgia: A Case Series. *Am. J. Trop. Med. Hyg.* **2018**, *98*, 1870–1875. [CrossRef] [PubMed]

8. Müller, G.; Gröters, S.; Siebert, U.; Rosenberger, T.; Driver, J.; König, M.; Becher, P.; Hetzel, U.; Baumgärtner, W. Parapoxvirus infection in harbor seals (Phoca vitulina) from the German North Sea. *Vet. Pathol.* **2003**, *40*, 445–454. [CrossRef] [PubMed]

9. Roess, A.A.; Mccollum, A.M.; Gruszynski, K.; Zhao, H.; Davidson, W.; Lafon, N.; Engelmeyer, T.; Moyer, B.; Godfrey, C.; Kilpatrick, H.; et al. Surveillance of Parapoxvirus Among Ruminants in Virginia and Connecticut. *Zoonoses Public Health* **2013**, *60*, 543–548. [CrossRef] [PubMed]

10. Velazquez-Salinas, L.; Ramirez-Medina, E.; Bracht, A.J.; Hole, K.; Brito, B.P.; Gladue, D.; Carrillo, C. Phylodynamics of parapoxvirus genus in Mexico (2007–2011). *Infect. Genet. Evol.* **2018**, *65*, 12–14. [CrossRef]

11. Hosamani, M.; Scagliarini, A.; Bhanuprakash, V.; McInnes, C.J.; Singh, R.K. Orf: An update on current research and future perspectives. *Expert Rev. Anti-Infect. Ther.* **2009**, *7*, 879–893. [CrossRef] [PubMed]

12. De Sant'Ana, F.J.; Rabelo, R.E.; Vulcani, V.A.; Cargnelutti, J.F.; Flores, E.F. Bovine papular stomatitis affecting dairy cows and milkers in midwestern Brazil. *J. Vet. Diagn. Investig.* **2012**, *24*, 442–445. [CrossRef]

13. Nougairede, A.; Fossati, C.; Salez, N.; Cohen-Bacrie, S.; Ninove, L.; Michel, F.; Aboukais, S.; Büttner, M.; Zandotti, C.; De Lamballerie, X.; et al. Sheep-to-Human Transmission of Orf Virus during Eid al-Adha Religious Practices, France. *Emerg. Infect. Dis.* **2013**, *19*, 102–105. [CrossRef] [PubMed]

14. Kitchen, M.; Muller, H.; Zobl, A.; Windisch, A.; Romani, N.; Huemer, H. Orf Virus Infection in a Hunter in Western Austria, Presumably Transmitted by Game. *Acta Dermato-Venereol.* **2014**, *94*, 212–214. [CrossRef]

15. Osadebe, L.U.; Manthiram, K.; McCollum, A.M.; Li, Y.; Emerson, G.L.; Gallardo-Romero, N.F.; Doty, J.B.; Wilkins, K.; Zhao, H.; Drew, C.P.; et al. Novel poxvirus infection in 2 patients from the United States. *Clin. Infect. Dis.* **2015**, *60*, 195–202. [CrossRef]

16. Midilli, K.; Erkılıç, A.; Kuşkucu, M.; Analay, H.; Erkiliç, S.; Benzonana, N.; Yildirim, M.S.; Mülayim, K.; Acar, H.; Ergonul, O. Nosocomial outbreak of disseminated orf infection in a burn unit, Gaziantep, Turkey, October to December 2012. *Eurosurveillance* **2013**, *18*, 20425. [CrossRef]

17. Lederman, E.R.; Green, G.M.; DeGroot, H.E.; Dahl, P.; Goldman, E.; Greer, P.W.; Li, Y.; Zhao, H.; Paddock, C.D.; Damon, I.K. Progressive Orf Virus Infection in a Patient with Lymphoma: Successful Treatment Using Imiquimod. *Clin. Infect. Dis.* **2007**, *44*, e100–e103. [CrossRef] [PubMed]

18. Tedla, M.G.; Berhan, N.; Molla, W.; Temesgen, W.; Alemu, S. Molecular identification and investigations of contagious ecthyma (Orf virus) in small ruminants, North west Ethiopia. *BMC Vet. Res.* **2018**, *14*, 13. [CrossRef]

19. Mazur, C.; Machado, R.D. Detection of contagious pustular dermatitis virus of goats in a severe outbreak. *Vet. Rec.* **1989**, *125*, 419–420. [CrossRef] [PubMed]

20. McKeever, D.; Jenkinson, D.M.; Hutchison, G.; Reid, H. Studies of the pathogenesis of orf virus infection in sheep. *J. Comp. Pathol.* **1988**, *99*, 317–328. [CrossRef]

21. De Sant'Ana, F.J.D.; Leal, F.A.; Rabelo, R.E.; Vulcani, V.A.; Moreira, C.A., Jr.; Cargnelutti, J.F.; Flores, E.F. Coinfection by Vaccinia virus and an Orf virus-like parapoxvirus in an outbreak of vesicular disease in dairy cows in midwestern Brazil. *J. Vet. Diagn. Investig.* **2013**, *25*, 267–272. [CrossRef] [PubMed]

22. Schmidt, C.; Cargnelutti, J.; Brum, M.C.; Traesel, C.K.; Weiblen, R.; Flores, E.F. Partial sequence analysis of B2L gene of Brazilian orf viruses from sheep and goats. *Vet. Microbiol.* **2013**, *162*, 245–253. [CrossRef]

23. Abrahão, J.S.; Borges, I.A.; Mazur, C.; Lobato, Z.I.P.; Ferreira, P.C.P.; Bonjardim, C.A.; Trindade, G.S.; Kroon, E.G. Looking back: A genetic retrospective study of BrazilianOrf virusisolates. *Vet. Rec.* **2012**, *171*, 476. [CrossRef]

24. Cargnelutti, J.F.; Flores, M.M.; Teixeira, F.R.M.; Weiblen, R.; Flores, E.F. An outbreak of pseudocowpox in fattening calves in southern Brazil. *J. Vet. Diagn. Investig.* **2012**, *24*, 437–441. [CrossRef] [PubMed]

25. Oliveira, C.H.; Assis, F.L.; Neto, J.D.B.; Oliveira, C.M.; Lopes, C.T.D.A.; Bomjardim, H.D.A.; Vinhote, W.M.; E Silva, A.G.M.; Abrahão, J.S.; Kroon, E.G. Multifocal Cutaneous Orf Virus Infection in Goats in the Amazon Region, Brazil. *Vector-Borne Zoonotic Dis.* **2012**, *12*, 336–340. [CrossRef] [PubMed]

26. Petrova, V.N.; Russell, C.A. The evolution of seasonal influenza viruses. *Nat. Rev. Microbiol.* **2018**, *16*, 47–60. [CrossRef]

27. Zhao, Y.; Zhang, H.; Zhao, J.; Guo-Zhong, Z.; Jin, J.-H.; Zhang, G.-Z. Evolution of infectious bronchitis virus in China over the past two decades. *J. Gen. Virol.* **2016**, *97*, 1566–1574. [CrossRef]

28. Drillien, R.; Koehren, F.; Kirn, A. Host range deletion mutant of vaccinia virus defective in human cells. *Virology* **1981**, *111*, 488–499. [CrossRef]

29. Moss, B.; Winters, E.; A Cooper, J. Deletion of a 9,000-base-pair segment of the vaccinia virus genome that encodes nonessential polypeptides. *J. Virol.* **1981**, *40*, 387–395. [CrossRef]

30. Panicali, D.; Davis, S.W.; Mercer, S.R.; Paoletti, E. Two major DNA variants present in serially propagated stocks of the WR strain of vaccinia virus. *J. Virol.* **1981**, *37*, 1000–1010. [CrossRef]

31. Paez, E.; Dallo, S.; Esteban, M. Generation of a dominant 8-MDa deletion at the left terminus of vaccinia virus DNA. *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 3365–3369. [CrossRef] [PubMed]

32. Dallo, S.; Esteban, M. Isolation and characterization of attenuated mutants of vaccinia virus. *Virology* **1987**, *159*, 408–422. [CrossRef]

33. Meyer, H.; Sutter, G.; Mayr, A. Mapping of deletions in the genome of the highly attenuated vaccinia virus MVA and their influence on virulence. *J. Gen. Virol.* **1991**, *72 Pt 5*, 1031–1038. [CrossRef]

34. Cottone, R.; Büttner, M.; Bauer, B.; Henkel, M.; Hettich, E.; Rziha, H.-J. Analysis of genomic rearrangement and subsequent gene deletion of the attenuated Orf virus strain D1701. *Virus Res.* **1998**, *56*, 53–67. [CrossRef]

35. Lake, J.R.; Cooper, P.D. Deletions of the Terminal Sequences in the Genomes of the White Pock (u) and Host-restricted (p) Mutants of Rabbitpox Virus. *J. Gen. Virol.* **1980**, *48*, 135–147. [CrossRef] [PubMed]

36. Moyer, R.W.; Rothe, C.T. The white pock mutants of rabbit poxvirus. I. Spontaneous host range mutants contain deletions. *Virology* **1980**, *102*, 119–132. [CrossRef]

37. Archard, L.C.; Mackett, M. Restriction Endonuclease Analysis of Red Cowpox Virus and its White Pock Variant. *J. Gen. Virol.* **1979**, *45*, 51–63. [CrossRef] [PubMed]

38. Esposito, J.J.; Cabradilla, C.D.; Nakano, J.H.; Obijeski, J.F. Intragenomic sequence transposition in monkeypox virus. *Virology* **1981**, *109*, 231–243. [CrossRef]

39. Kotwal, G.J.; Moss, B. Analysis of a large cluster of nonessential genes deleted from a vaccinia virus terminal transposition mutant. *Virology* **1988**, *167*, 524–537. [CrossRef]

40. Archard, L.C.; Mackett, M.; Barnes, D.E.; Dumbell, K.R. The Genome Structure of Cowpox Virus White Pock Variants. *J. Gen. Virol.* **1984**, *65*, 875–886. [CrossRef]

41. Pickup, D.J.; Ink, B.S.; Parsons, B.L.; Hu, W.; Joklik, W.K. Spontaneous deletions and duplications of sequences in the genome of cowpox virus. *Proc. Natl. Acad. Sci. USA* **1984**, *81*, 6817–6821. [CrossRef] [PubMed]

42. Dumbell, K.R.; Archard, L.C. Comparison of white pock (h) mutants of monkeypox virus with parental monkeypox and with variola-like viruses isolated from animals. *Nat. Cell Biol.* **1980**, *286*, 29–32. [CrossRef]

43. Karki, M.; Kumar, A.; Arya, S.; Ramakrishnan, M.A.; Venkatesan, G. Poxviral E3L ortholog (Viral Interferon resistance gene) of orf viruses of sheep and goats indicates species-specific clustering with heterogeneity among parapoxviruses. *Cytokine* **2019**, *120*, 15–21. [CrossRef] [PubMed]

44. Rziha, H.-J.; Büttner, M.; Müller, M.; Salomon, F.; Reguzova, A.; Laible, D.; Amann, R. Genomic Characterization of Orf Virus Strain D1701-V (Parapoxvirus) and Development of Novel Sites for Multiple Transgene Expression. *Viruses* **2019**, *11*, 127. [CrossRef]

45. Prosperi, M.C.; Ciccozzi, M.; Fanti, I.; Saladini, F.; Pecorari, M.; Borghi, V.; Di Giambenedetto, S.; Bruzzone, B.; Capetti, A.; Vivarelli, A.; et al. A novel methodology for large-scale phylogeny partition. *Nat. Commun.* **2011**, *2*, 321. [CrossRef] [PubMed]

46. Bronson, E.C.; Anderson, J.N. Nucleotide composition as a driving force in the evolution of retroviruses. *J. Mol. Evol.* **1994**, *38*, 506–532. [CrossRef]

47. Shackelton, L.A.; Parrish, C.R.; Holmes, E.C. Evolutionary basis of codon usage and nucleotide composition bias in vertebrate DNA viruses. *J. Mol. Evol.* **2006**, *62*, 551–563. [CrossRef] [PubMed]

48. Schachtel, G.A.; Bucher, P.; Mocarski, E.S.; Blaisdell, B.E.; Karlin, S. Evidence for selective evolution in codon usage in conserved amino acid segments of human alphaherpesvirus proteins. *J. Mol. Evol.* **1991**, *33*, 483–494. [CrossRef] [PubMed]

49. Chen, H.; Li, W.; Kuang, Z.; Chen, D.; Liao, X.; Li, M.; Luo, S.; Hao, W. The whole genomic analysis of orf virus strain HN3/12 isolated from Henan province, central China. *BMC Vet. Res.* **2017**, *13*, 260. [CrossRef] [PubMed]

50. Andreani, J.; Fongue, J.; Khalil, J.Y.B.; David, L.; Mougari, S.; Le Bideau, M.; Abrahão, J.; Berbis, P.; La Scola, B. Human Infection with Orf Virus and Description of Its Whole Genome, France, 2017. *Emerg. Infect. Dis.* **2019**, *25*, 2197–2204. [CrossRef] [PubMed]

51. Savory, L.J.; Stacker, S.A.; Fleming, S.B.; Niven, B.E.; Mercer, A.A. Viral vascular endothelial growth factor plays a critical role in orf virus infection. *J. Virol.* **2000**, *74*, 10699–10706. [CrossRef]

52. Rziha, H.-J.; Bauer, B.; Adam, K.-H.; Röttgen, M.; Cottone, R.; Henkel, M.; Dehio, C.; Büttner, M. Relatedness and heterogeneity at the near-terminal end of the genome of a parapoxvirus bovis 1 strain (B177) compared with parapoxvirus ovis (Orf virus). *J. Gen. Virol.* **2003**, *84*, 1111–1116. [CrossRef] [PubMed]

53. Wise, L.M.; Stuart, G.S.; Real, N.C.; Fleming, S.B.; Mercer, A.A. Orf virus IL-10 accelerates wound healing while limiting inflammation and scarring. *Wound Repair Regen.* **2014**, *22*, 356–367. [CrossRef]

54. Fleming, S.B.; Haig, D.M.; Nettleton, P.; Reid, H.W.; McCaughan, C.A.; Wise, L.M.; Mercer, A.A. Sequence and functional analysis of a homolog of interleukin-10 encoded by the parapoxvirus orf virus. *Virus Genes* **2000**, *21*, 85–95. [CrossRef]

55. Perrin, A.; Larsonneur, E.; Nicholson, A.C.; Edwards, D.J.; Gundlach, K.M.; Whitney, A.M.; Gulvik, C.A.; Bell, M.E.; Rendueles, O.; Cury, J.; et al. Evolutionary dynamics and genomic features of the Elizabethkingia anophelis 2015 to 2016 Wisconsin outbreak strain. *Nat. Commun.* **2017**, *8*, 15483. [CrossRef] [PubMed]

56. Pu, J.; Wang, S.; Yin, Y.; Zhang, G.; Carter, R.A.; Wang, J.; Xu, G.; Sun, H.; Wang, M.; Wen, C.; et al. Evolution of the H9N2 influenza genotype that facilitated the genesis of the novel H7N9 virus. *Proc. Natl. Acad. Sci. USA* **2014**, *112*, 548–553. [CrossRef]

57. Ahanger, S.A.; Parveen, R.; Nazki, S.; Dar, Z.; Dar, T.; Dar, K.H.; Dar, A.; Rai, N.; Dar, P.A. Detection and phylogenetic analysis of Orf virus in Kashmir Himalayas. *Virus Dis.* **2018**, *29*, 405–410. [CrossRef] [PubMed]

58. Bala, J.A.; Balakrishnan, K.N.; Abdullah, A.A.; Mohamed, R.; Haron, A.W.; Jesse, F.F.A.; Noordin, M.M.; Mohd-Azmi, M.L. The re-emerging of orf virus infection: A call for surveillance, vaccination and effective control measures. *Microb. Pathog.* **2018**, *120*, 55–63. [CrossRef]

59. Bora, M.; Bora, D.P.; Barman, N.N.; Borah, B.; Bora, P.L.; Talukdar, A.; Tamuly, S. Isolation and molecular characterization of Orf virus from natural outbreaks in goats of Assam. *Virus Dis.* **2015**, *26*, 82–88. [CrossRef]

60. Chan, K.W.; Lin, J.W.; Lee, S.H.; Liao, C.J.; Tsai, M.C.; Hsu, W.L.; Wong, M.L.; Shih, H.C. Identification and phylogenetic analysis of orf virus from goats in Taiwan. *Virus Genes* **2007**, *35*, 705–712. [CrossRef]

61. Flores, C.; Gonzalez Altamiranda, E.A.; Verna, A.E.; Peralta, A.V.; Madariaga, C.; Odeón, A.C.; Canton, G. Orf virus in human, confirmation in case report from Chile. *Rev. Chilena Infectol.* **2017**, *34*, 607–609. [CrossRef] [PubMed]

62. Şevik, M. Association of two clusters of Orf virus isolates in outbreaks of infection in goat in the Central Anatolian region of Turkey. *Virus Dis.* **2017**, *28*, 345–348. [CrossRef]

63. Soylemez, O.; Kondrashov, F.A. Estimating the Rate of Irreversibility in Protein Evolution. *Genome Biol. Evol.* **2012**, *4*, 1213–1222. [CrossRef]

64. Haig, D.; McInnes, C.J.; Deane, D.; Reid, H.; Mercer, A.A. The immune and inflammatory response to orf virus. *Comp. Immunol. Microbiol. Infect. Dis.* **1997**, *20*, 197–204. [CrossRef]

65. Paillot, R. A systematic review of the immune-modulators Parapoxvirus ovis and Propionibacterium acnes for the prevention of respiratory disease and other infections in the horse. *Vet. Immunol. Immunopathol.* **2013**, *153*, 1–9. [CrossRef] [PubMed]

66. Soderholm, J.; Ahlén, G.; Kaul, A.; Frelin, L.; Alheim, M.; Barnfield, C.; Liljeström, P.; Weiland, O.; Milich, D.R.; Bartenschlager, R.; et al. Relation between viral fitness and immune escape within the hepatitis C virus protease. *Gut* **2006**, *55*, 266–274. [CrossRef] [PubMed]

67. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [CrossRef]

68. Hall, T.A. Bioedit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **1999**, *41*, 95–98.

69. Letunic, I.; Bork, P. Interactive Tree Of Life v2: Online annotation and display of phylogenetic trees made easy. *Nucleic Acids Res.* **2011**, *39*, W475–W478. [CrossRef]

70. Ito, K.; Murphy, D. Application of ggplot2 to Pharmacometric Graphics. *Cpt Pharmacomet. Syst. Pharmacol.* **2013**, *2*, e79. [CrossRef]

71. Darriba, D.; Taboada, G.L.; Doallo, R.; Posada, D. jModelTest 2: More models, new heuristics and parallel computing. *Nat. Methods* **2012**, *9*, 772. [CrossRef]

72. Newton, M.A.; Raftery, A.E. Approximate bayesian inference with the weighted likelihood bootstrap. *J. R. Stat. Soc. Ser. B (Methodol.)* **1994**, *56*, 3–48. [CrossRef]

73. Tamura, K.; Peterson, N.; Stecher, G.; Nei, M.; Kumar, S. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol. Biol. Evol.* **2011**, *28*, 2731–2739. [CrossRef]

74. Bouckaert, R.; Heled, J.; Kühnert, D.; Vaughan, T.; Wu, C.-H.; Xie, D.; Suchard, M.A.; Rambaut, A.; Drummond, A.J. BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Comput. Biol.* **2014**, *10*, e1003537. [CrossRef] [PubMed]

75. Shapiro, B.; Rambaut, A.; Drummond, A.J. Choosing Appropriate Substitution Models for the Phylogenetic Analysis of Protein-Coding Sequences. *Mol. Biol. Evol.* **2005**, *23*, 7–9. [CrossRef] [PubMed]

76. Rambaut, A.; Lam, T.T.; Carvalho, L.M.; Pybus, O.G. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol.* **2016**, *2*, vew007. [CrossRef] [PubMed]