RESEARCH ARTICLE

Porcine circovirus 2 (PCV-2) genotype update and proposal of a new genotyping methodology

Giovanni Franzo[1]*, Joaquim Segalés[2,3]

1 Department of Animal Medicine, Production and Health, University of Padova, Legnaro, Italy, **2** UAB, Centre de Recerca en Sanitat Animal (CRESA, IRTA-UAB), Campus de la Universitat Autònoma de Barcelona, Bellaterra, Spain, **3** Departament de Sanitat i Anatomia Animals, Facultat de Veterinària, Universitat Autònoma de Barcelona, Bellaterra Spain

* giovanni.franzo@unipd.it
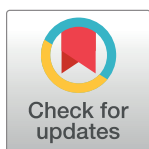
Abstract

Porcine circovirus 2 (PCV-2) is one of the most widespread viral infections of swine, causing a remarkable economic impact because of direct losses and indirect costs for its control. As other ssDNA viruses, PCV-2 is characterized by a high evolutionary rate, leading to the emergence of a plethora of variants with different biological and epidemiological features. Over time, several attempts have been made to organize PCV-2 genetic heterogeneity in recognized genotypes. This categorization has clearly simplified the epidemiological investigations, allowing to identify different spatial and temporal patterns among genotypes. Additionally, variable virulence and vaccine effectiveness have also been hypothesized. However, the rapid increase in sequencing activity, coupled with the *per se* high viral variability, has challenged the previously established nomenclature, leading to the definition of several study-specific genotypes and hindering the capability of performing comparable epidemiological studies.

Based on these premises, an updated classification scheme is herein reported. Recognizing the impossibility of defining a clear inter-cluster p-distance cut-off, the present study proposes a phylogeny-grounded genotype definition based on three criteria: maximum intra-genotype p-distance of 13% (calculated on the ORF2 gene), bootstrap support at the corresponding internal node higher than 70% and at least 15 available sequences. This scheme allowed defining 8 genotypes (PCV-2a to PCV-2h), which six of those had been previously proposed. To minimize the inconvenience of implementing a new classification, the most common names already adopted have been maintained when possible. The analysis of sequence-associated metadata highlighted a highly unbalanced sequencing activity in terms of geographical, host and temporal distribution. The PCV-2 molecular epidemiology scenario appears therefore characterized by a severe bias that could lead to spurious associations between genetic and epidemiological/biological viral features. While the suggested classification can establish a "common language" for future studies, further efforts should be paid to achieve a more homogeneous and informative representation of the PCV-2 global scenario.

Introduction

Initial sequence analyses of different *Porcine circovirus 2* (PCV-2) viruses around the world indicated a close phylogenetic relationship, with a nucleotide sequence identity higher than 93% [1]. Although genetic differences among strains/sequences of PCV-2 were considered minimal, when a nucleotide diversity cut-off of 3.5% was applied to existing sequences, they could be divided into two major phylogenetic groups. Subsequently, those genetic groups were called genotypes; after a number of nomenclature proposals for these genotypes by different research groups, a consensus (genotypes "a" and "b") was reached [2]. Subsequently, a third genotype was retrospectively reported from Denmark in the 1980s and named PCV-2c [3]. Recent data indicated the current existence of this genotype in feral pigs from Brazil [4] and in domestic swine in China [5]. A fourth genotype named PCV-2d was lately described [6], although retrospectively detected in Switzerland already in 1998 [7]. In fact, it was initially called as a mutant of PCV-2b (mPCV-2b) and linked to potential vaccination failure cases [8]. Further studies have shown that PCV-2d is widely spread all over the world nowadays, becoming more and more prevalent. More recently, a new genotype (PCV-2e) has been identified in swine samples from USA and Mexico [9,10]. In addition, a Chinese study found a year ago novel viral sequences clustering differently from the existing ones and tentatively classified them as genotype PCV-2f [11]. Currently, it is considered that PCV-2d exhibits similar virulence to PCV-2a and PCV-2b when inoculated into susceptible pigs [12], but the clinical significance of PCV-2c, PCV-2e and PCV-2f are not known. Certainly, the appearance of new genotypes in the future is likely, since PCV-2 is a single stranded DNA virus with high nucleotide substitution rate (comparable to those of RNA viruses) and the possibility to generate mutations of the genome is high [13].

Recent phylodynamic studies have indicated that PCV-2 population dynamics have been characterized by the appearance of periodic waves of different genotypes which, after an initial increase in prevalence, spread following major swine commercial routes and were then superseded by subsequent emerging genotypes [14]. PCV-2a was the most prevalent genotype in clinically affected pigs from 1996 to the early 2000s, after which PCV-2b predominated ("genotype shift"). The emergence of PCV-2b in North America and Europe was associated with the appearance of a more severe clinical disease [15–18]. PCV-2d was first reported in China [6] and a second "genotype shift" (from PCV-2b to PCV-2d) seems to be occurring globally [7,14], perhaps driven by the worldwide use of PCV-2 vaccines.

Importantly, different PCV-2 strains of the same or different genotype can be detected in the same pig [19]. Also, *in vivo* and *in vitro* studies have provided evidence of potential viral recombination. In fact, some circulating recombinant forms (named as CRFs) have followed comparable population dynamics and spreading routes to those of well-defined genotypes, suggesting that recombinant strains are able to circulate and compete with parental ones [14].

Taking into account the very recent description of novel PCV-2 genotypes and the growing difficulty for classifying consistently current and novel viral sequences based on initially proposed methods, this work aimed to review existing classification criteria and propose a further developed method for PCV-2 genotyping.

Material and methods

Sequence dataset

All available PCV-2 sequences were downloaded from GenBank (https://www.ncbi.nlm.nih.gov/genbank/) using the species-specific taxonomic identifier code (i.e. 85708) and records whose complete ORF2 gene was available were selected. All sequences were aligned at codon

level using the MAFFT method implemented in TranslatorX [20], and the obtained alignment was inspected both visually and using dedicated Python scripts. Poorly aligned sequences, those displaying premature stop codons, frameshift mutations or unknown bases, suggestive of poor sequence quality, were removed. Alignment was then performed again on the refined dataset.

Recombination and phylogenetic analyses

To reduce the computational burden and remove data redundancy, only one sequence representative of all identical ones was identified using ElimDupes (https://www.hiv.lanl.gov/content/sequence/ELIMDUPES/elimdupes.html) and maintained for further analyses.

Recombination analysis was performed on the non-redundant ORF2 sequences dataset using RDP4 v4.91 [21]. The RDP, GENECONV and MaxChi methods were selected as primary scan, while all the methods implemented in RDP4 were used for recombination detection refinement. For each method, settings were adjusted considering the database features according to the software manual recommendations. A recombination event was accepted if detected by more than two methods with a significance p-value of $p < 0.01$ with Bonferroni's correction. All sequences identified as recombinants where excluded from further analysis.

Pairwise genetic p-distance was calculated for all sequences using the *ape* version 5.2 [22] package in R version 3.3.2 [23] in order to graphically evaluate, by means of a density plot, the presence of reasonable inter-group separation cut-off.

A Neighbor-Joining phylogenetic tree based on raw genetic distances (pairwise p-distance) was reconstructed using MEGA7 [24]. The reliability of the sequence cluster was evaluated by performing 1000 bootstrap replicates.

The new proposed genotype definition was based on within-cluster genetic distance and clade reliability, defined by bootstrap support. Several arbitrary genetic distances thresholds were attempted to define a reasonable cluster number, maintaining the correspondence with the existing PCV-2 genetic diversity (genotypes defined to date). As final setting, ClusterPick [25] was used to identify sequence groups featured by a within-cluster genetic distance lower than 13% and bootstrap support of the relative ancestral node higher than 70%.

Genotype temporal, geographical and host distribution

The non-redundant database was expanded to its original size, and the accession numbers of sequences classified within each cluster were recorded. Collection country, host and date were downloaded from GenBank using homemade python scripts benefiting of the functions of the Biophyton library [26]. To facilitate data analysis and graphical representation, aggregated categories were created (e.g., countries were aggregated to continent level, similar host were merged, etc.). To assess the presence of statistically significant differences in the cluster distribution among categories, a Chi-square or Fisher test was performed. The presence of over- or under represented cluster-category pairs was assessed by evaluating the respective z-score. Statistical significance was set to p-value < 0.05.

Results

A total of 6,171 PCV-2 sequences were downloaded from GenBank (accessed 05-July-2018). However, after database refinement and potential recombinant sequence removal (about 22% of the initial sequence number), 4,586 ORF2 sequences were maintained in the analysis, corresponding to 2,610 unique sequences.

Analysis of p-distance profile distribution evidenced the absence of any clear cut-off to divide PCV-2 genetic variability in a limited, but still meaningful, genotype number (Fig 1). Therefore, the analysis of phylogenetic clusters was preferred instead.

Intra−Cluster genetic distance
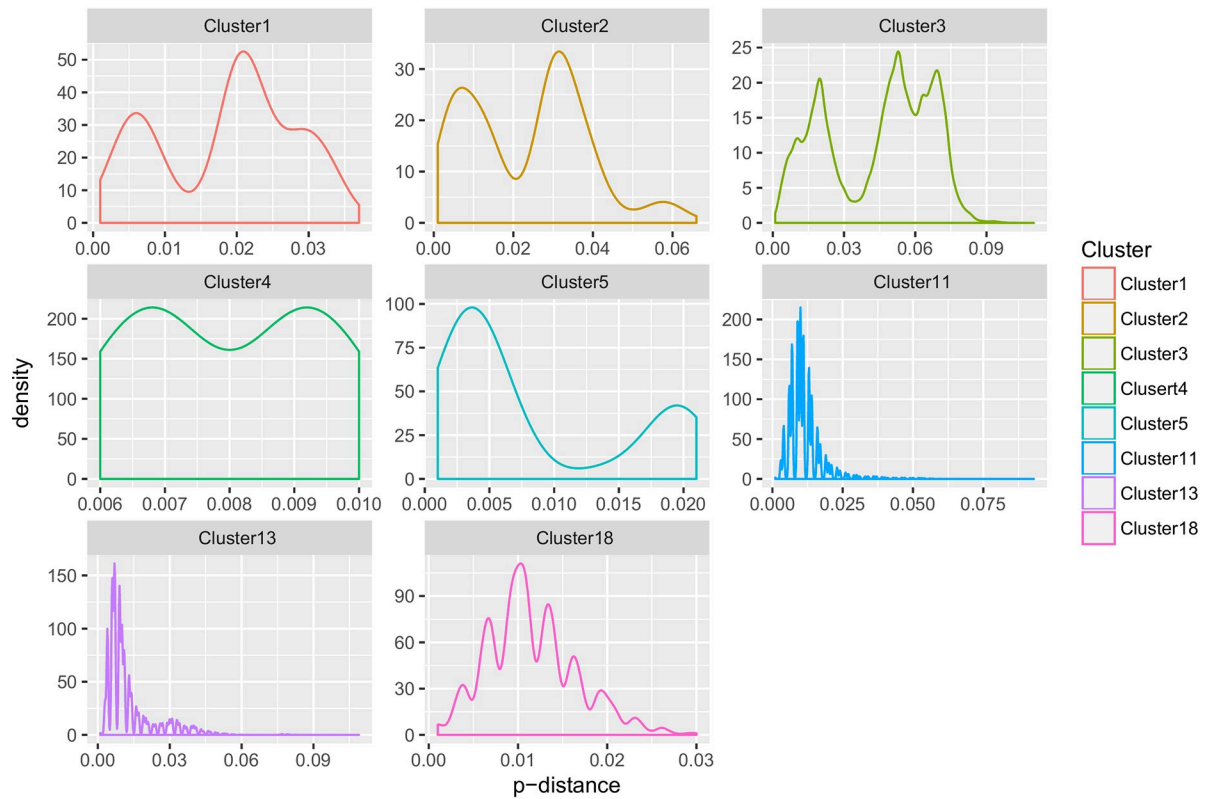
Inter−Cluster genetic distance

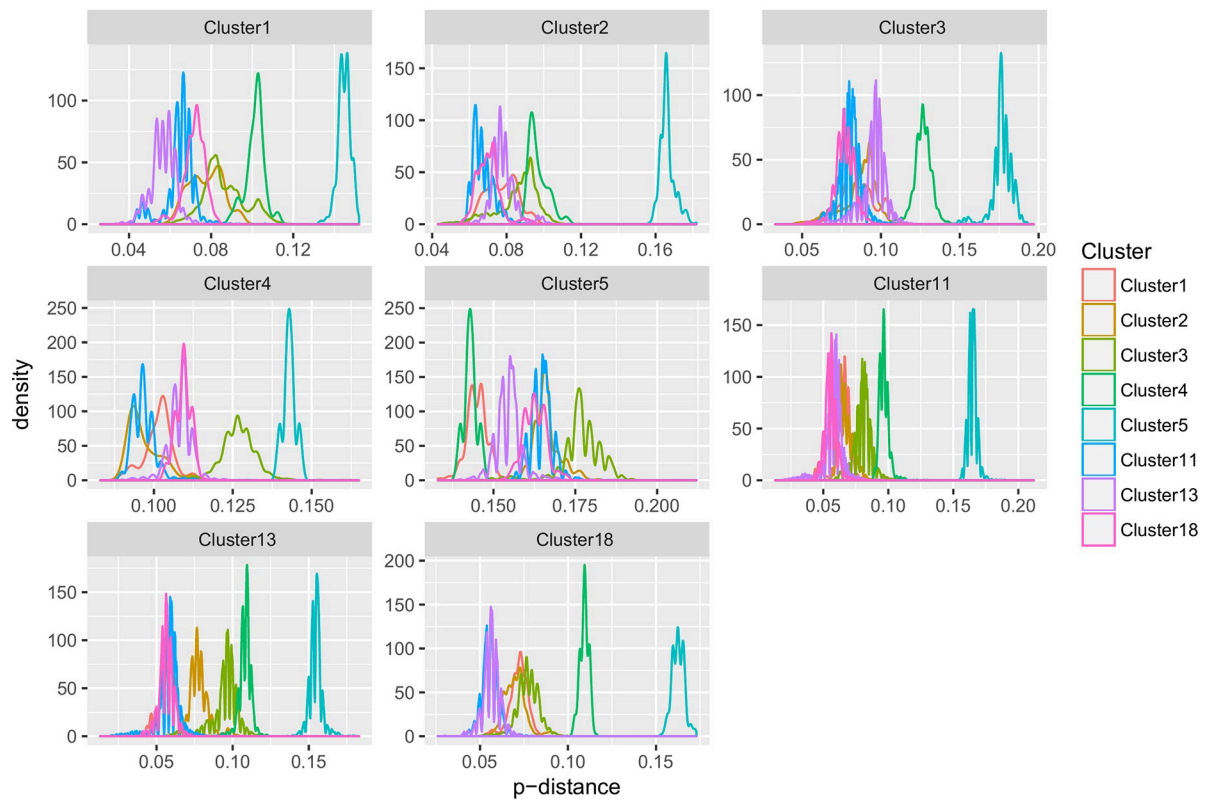**Fig 1. Intra and inter-cluster pairwise p-distances among ORF2 sequences.** Density plot reporting the pairwise p-distance distribution within and between clusters. Different clusters have been color-coded.

https://doi.org/10.1371/journal.pone.0208585.g001

Cluster feature analysis led to the definition of 18 genotypes (S1 Table). However, only 7 clusters (Cluster1, Cluster2, Cluster3, Cluster5, Cluster11, Cluster13, and Cluster18) included more than 15 sequences in the original dataset and were considered for being potentially recognized as genotypes and for further analyses. A summary of the sequence number classified in each cluster/genotype is reported in Table 1.

Some of the defined clusters mirrored previously defined genotypes (i.e. Cluster2 = PCV-2f; Cluster3 = PCV-2a; Cluster 4 = PCV-2c (maintained for historical reasons); Cluster 5 = PCV-2e; Cluster11 = PCV-2b; and Cluster13 = PCV-2d). Additionally, two more genotypes are proposed (Cluster1 = PCV-2g; Cluster18 = PCV-2h) (Fig 2).

Among all these groupings, Cluster11 (PCV-2b) included the highest sequence number, followed by Cluster13 (PCV-2d) and Cluster3 (PCV-2a). Nevertheless, a clear temporal trend was evident, being the Cluster3 the first one with a relevant number of published sequences, followed by Cluster 11 (especially from 2005 onwards) and, finally, Cluster13, which demonstrated a progressive increase in sequenced strains over time (Fig 3). Cluster3, 11 and 13 displayed an actual worldwide distribution, being detected in all continents with the exception of Africa and Oceania, for which the number of available sequences was negligible (Table 1 and Fig 4). Cluster3 was the most prevalent one in Africa and Oceania, Cluster11 in America and Europe and Cluster13 in Asia (Fig 4). A non-random distribution was observed in PCV-2 cluster geographical pattern (p<0.001).

The vast majority of available sequences were collected from domestic pig (*Sus scrofa domesticus*). In detail, among the recognized genotypes, 3,386 out of 3,555 (95.2%) sequences, for which collection host was known, originated from domestic pig, 155 (4.36%) from wild boar, 11 (0.31%) from ruminants, 2 (0.056%) from human beings and 1 from *Rattus norvegicus* (0.028%). A summary of Host-Cluster association is reported in Table 2.

A Chi-square test performed on the more represented clusters and hosts demonstrated a non-random association (p<0.001). Particularly, Cluster1 and 11 were overrepresented in wild boar, although this might be attributable to the low number of available sequences. On the other hand, *Sus scrofa domesticus* was under-represented in Cluster1 and wild boar in Cluster13 (Table 2 and S2 Fig).

To facilitate PCV-2 classification in routine settings, a dataset of reference sequences, annotated with the corresponding cluster/genotype name, has been provided (S1 Table).

Table 1. Count of sequences classified in each cluster/genotype-continent pair. The respective column percentage is reported between brackets. For an easier interpretation, only major clusters corresponding to proposed genotypes are reported.

Cluster	Genotype	Africa	Asia	Europe	North America	Oceania	South America	Total
Cluster3	PCV-2a	1 (100%)	299 (12.6%)	50 (10,39%)	301 (22.69%)	8 (80%)	16 (11.67%)	675 (15,63%)
Cluster11	PCV-2b	0 (0%)	886 (37.51%)	374 (77,75%)	614 (46.30%)	2 (20%)	108 (78.83%)	1984 (45,94%)
Cluster4	PCV-2c	0 (0%)	0 (0%)	3 (0,62%)	0 (0%)	0 (0%)	1 (0,724)	4 (0,09%)
Cluster13	PCV-2d	0 (0%)	1065 (45.08%)	41 (8,52%)	373 (28.13%)	0 (0%)	12 (8.75%)	1491 (34,52%)
Cluster5	PCV-2e	0 (0%)	2 (0.08%)	0 (0%)	20 (1.50%)	0 (0%)	0 (0%)	22 (0,50%)
Cluster2	PCV-2f	0 (0%)	25 (1.05%)	1 (0,20%)	9 (0.67%)	0 (0%)	1 (0.72%)	36 (0,83%)
Cluster1	PCV-2g	0 (0%)	10 (0.42%)	12 (2,49%)	9 (0.67%)	0 (0%)	0 (0%)	31 (0,71%)
Cluster18	PCV-2h	0 (0%)	75 (3.17%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	75 (1,73%)
Total		1 (100%)	2362 (100%)	481 (100%)	1326 (100%)	11 (100%)	137 (100%)	4318 (100%)

https://doi.org/10.1371/journal.pone.0208585.t001

Fig 2. PCV-2 Neighbor-Joining phylogenetic tree. Neighbor-Joining phylogenetic tree reconstructed based on row genetic distances (i.e. pairwise p-distance) calculated on a collection of strains representative of the proposed PCV-2 genotypes. Both the cluster and genotype nomenclature are reported. Bootstrap support is displayed near the corresponding node. A tree reconstructed based on a more extensive sequence collection (n = 2610) is provided as S1 Fig.

https://doi.org/10.1371/journal.pone.0208585.g002

Discussion

Extensive genetic studies on PCV-2 have proven the high heterogeneity of this pathogen, which can be attributed to both high mutation and recombination rates [13,14]. Interestingly, PCV-2 variability has been potentially linked to different phenomena with remarkable practical implications, like differential epidemiological fitness [14], virulence [27], cross-protection [28], vaccine escape [29], etc., although definitive association has not consistently been proven in most instances because of limited data availability and poor results reproducibility. The multi-factorial nature of porcine circovirus diseases [30] can conceal the identification of causal nexus in absence of consistent and properly collected information. This scenario highlights the need for continuous, massive and updated epidemiological monitoring to investigate and understand PCV-2 epidemiological features and determinants. To this purpose, a classification scheme that simplifies and summarizes the peculiarities of each strain is mandatory for both routine diagnosis and viral characterization as well as for more complex association studies with important facts (virulence and cross-protection, mainly).

Currently, the most widely accepted classification scheme defines 4 genotypes identified based on phylogenetic analysis [31]. Since its publication, additional genetic groups have been described all around the world and claimed as new genotypes [9–11]. Nevertheless, in most instances, the newly detected strains were compared to a limited reference number, potentially magnifying some differences or missing the presence of additional inter-cluster relationships. Therefore, the present study proposes a new and updated genotype classification scheme based on a wide collection of ORF2 sequences, including all freely available non-recombinant ones. This particular region was selected since the *cap* gene is considered the more suitable phylogenetic and epidemiological marker for PCV2 and is able to reconstruct the same tree as the whole viral genome [1].

It must be stressed that, as already recognized in the previous classification scheme [31], no clear genetic distance cut-off could be identified to separate PCV-2 groups, since a relevant overlap between intra- and inter-genotype p-distance is present (Fig 1). Consequently, the proposed classification is substantially based on the phylogenetic relationship among strains, with adequate bootstrap support.
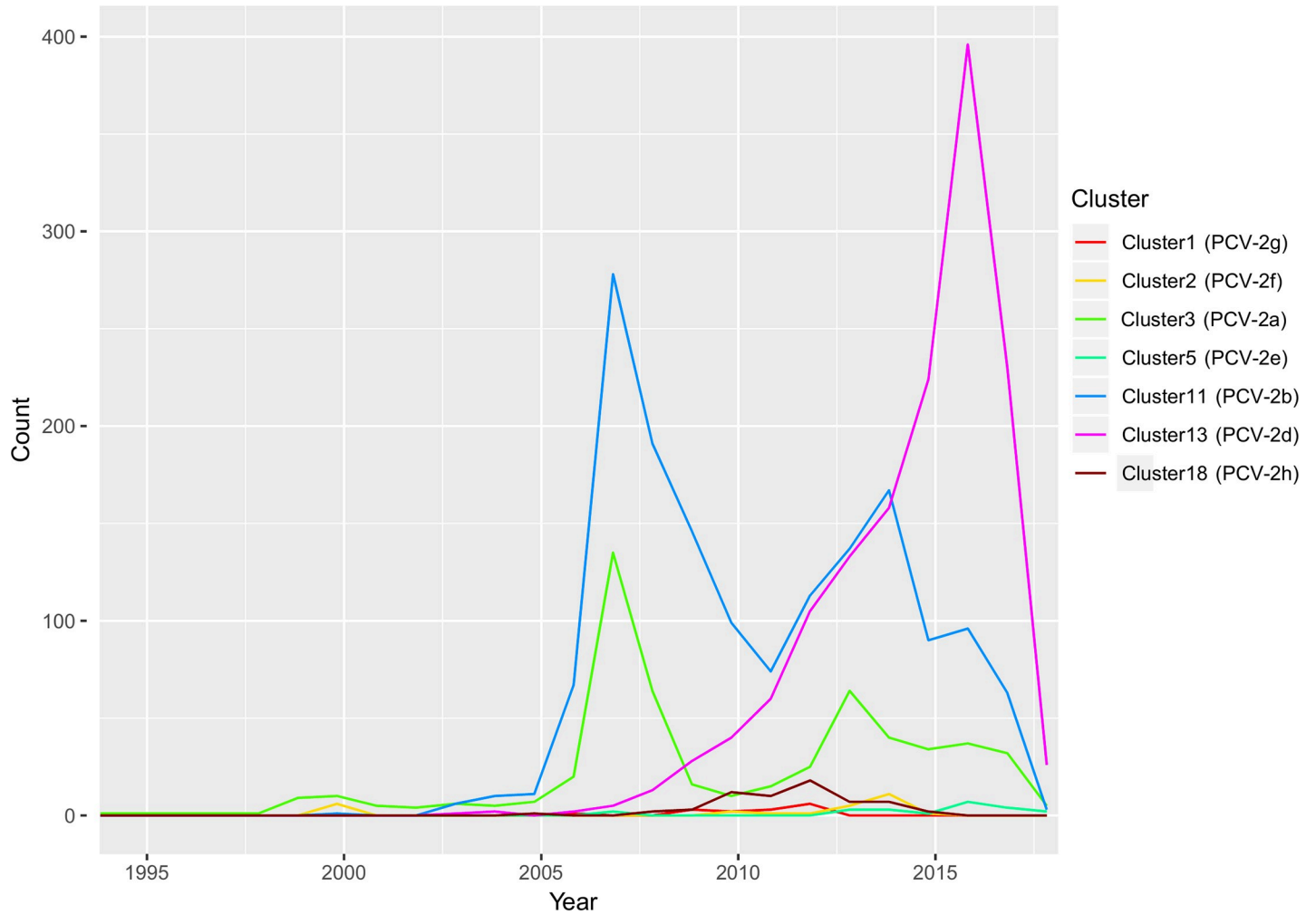
Fig 3. Count of sequences for each collection year. Different clusters have been color-coded.
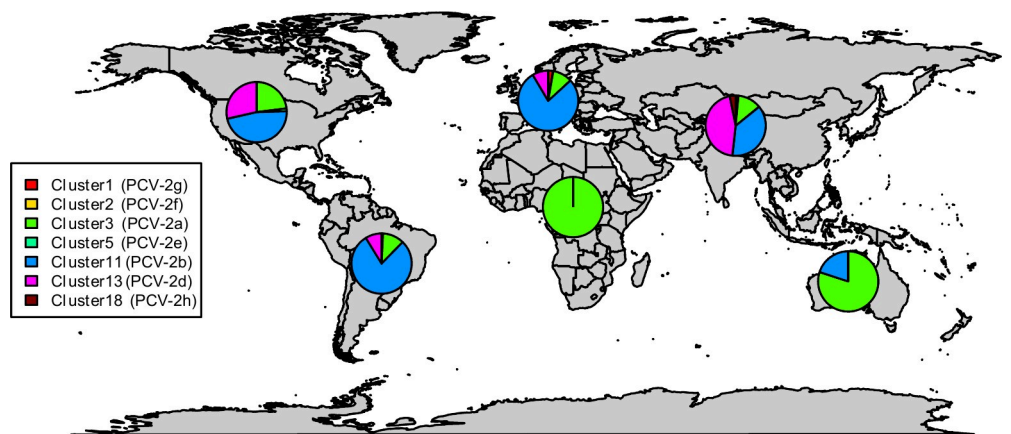
Fig 4. Geographic distribution of major PCV-2 clusters. The PCV-2 cluster distribution has been summarized using continent-specific pie-charts whose slide size is proportional to the cluster frequency (color-coded).

Table 2. Count of sequences classified in each cluster/genotype-host pair. The respective column percentage is reported between brackets. For interpretation easiness, only major clusters corresponding to proposed genotypes are reported.

.	Genotype	*Homo sapiens*	*Rattus norvegicus*	Ruminant	Domestic pig	Wild boar	Total
Cluster3	PCV-2a	1 (50%)	0 (0%)	2 (18.18%)	500 (14,74%)	17 (10.96%)	520 (14,61%)
Cluster11	PCV-2b	1 (50%)	1 (100%)	6 (54.54%)	1447 (42,68%)	114 (73.54%)	1569 (44,08%)
Cluster4	PCV-2c	0 (0%)	0 (0%)	0 (0%)	4 (0,11%)	0 (0%)	4 (0,11%)
Cluster13	PCV-2d	0 (0%)	0 (0%)	3 (27.27%)	1318 (38,88%)	11 (7.09%)	1332 (37,42%)
Cluster5	PCV-2e2e	0 (0%)	0 (0%)	0 (0%)	22 (0,64%)	0 (0%)	22 (0,62%)
Cluster2	PCV-2f	0 (0%)	0 (0%)	0 (0%)	26 (0,76%)	1 (0.64%)	27 (0,75%)
Cluster1	PCV-2g	0 (0%)	0 (0%)	0 (0%)	8 (0,23%)	12 (7.74%)	20 (0,56%)
Cluster18	PCV-2h	0 (0%)	0 (0%)	0 (0%)	65 (1,917%)	0 (0%)	65 (1,826%)
Totale		2 (100%)	1 (100%)	11 (100%)	3390 (100%)	155 (100%)	3559 (100%)

With the aim of providing a genotype definition able to depict the PCV-2 genetic heterogeneity, allowing at the same time a practical application in routine diagnostic and epidemiological studies (avoiding too many groupings and sub-groupings), it is suggested to limit the genotype definition to the following criteria: maximum within genotype p-distance of 13%, minimum cluster internal node bootstrap support of 70% and at least 15 identified sequences (A workflow of the proposed genotype analysis is depicted in S3 Fig). The last criterion was imposed to focus on more widespread genotypes and to avoid the risk of defining poor quality sequences as separate genotypes. An exception to this classification scheme is PCV-2c, which was maintained as a recognized genotype because of its historical description [3–5], although not fulfilling the third criterion yet.

The suggested classification is based on a NJ tree, reconstructed using raw genetic distances (i.e. pairwise p-distance). Certainly, other approaches like maximum likelihood, with more complex substitution models and rate heterogeneity among sites can better represent the phylogenetic relationship among strains. However, these methods imply a heuristic search in the tree likelihood space and estimation of all model parameters, being both database-dependent. In other words, somewhat different estimates (and thus different tree topology and branch lengths) can be obtained depending on the sequences included in the dataset. Since the present study aim was to establish a robust and long lasting approach for genotype definition, a simpler approach, less susceptible to future (and desirable) database updates, was selected. Overall, albeit all the previously proposed genotypes have been confirmed with the new methodology, other divergent genetic groups were discovered, demonstrating a higher variability than that previously expected for PCV-2.

Although the increase in the number of genotypes may be potentially surprising, these results could have been predicted for several reasons. The progressive increase in sequence number, obtained from previously non-sampled countries, could have revealed a hitherto undetected genetic variability. Additionally, the inclusion of thousands of sequences in the same analysis (instead of a limited set of them) understandably increases the capability of depicting the overall PCV-2 variability picture and identifying previously neglected genetic groups. Remarkably, most of the sequences included in newly defined genotypes were marked as "direct submission" in GenBank. Thus, although a relevant subset of those strains was sequenced years ago, most of them were not properly analyzed. The relevance of periodic and systematic analysis of freely available molecular data is therefore highly recommended to recognize the emerging genetic variability promptly.

The evolution of viruses with a high mutation rate genome is not expectable to follow a limited number of directions and a tendency to explore several paths in the fitness landscape can

be predicted [32–34]. Accordingly, only 7 out of 18 recognized clusters/genotypes included more than 15 sequences and only 3 were consistently detected for the whole considered time period and in substantially all continents. Therefore, a relevant proportion of the detected heterogeneity could be attributed to epidemiological dead ends, which underwent extinguished or circulated in extremely reduced host populations. The well-known history of PCV-2c (here named Cluster4) represents a good example in this sense [3,4].

Some of the clusters (i.e., Cluster1 and Cluster18) included a subset of sequences proposed as potential recombinant strains by some (but not all) previous studies [35,36]. The inclusion of hundreds of additional sequences could have affected the recombination detection in several ways. The sequencing of new strains, representative of "intermediate steps" in the PCV-2 evolution could have allowed explaining the phylogenetic relationship among strains without involving the presence of recombination events. However, a potential statistical power reduction in recombination detection cannot be excluded, both because of a higher complexity of the analyzed scenario and a more stringent multiple test correction.

Despite the efforts, it must be recognized that the presence in the final database of some, low fitness recombinant clusters cannot be completely excluded. In fact, the current methods are still perfectible, and recombination detection represents a partially unsolved challenge in viral phylogenesis. Moreover, the limited distance between parental strains and/or the progressive evolution after recombination occurrence, modifying the original recombinant sequence, could clearly complicate the detection of ancestral recombination events. Consequently, the confident identification of actual recombinant strains is probably far to be achieved. Nevertheless, the few potentially involved sequences and their limited practical role in the overall PCV-2 epidemiological scenario lessen the relevance of this formal shortcoming. Finally, apart from the theoretical knowledge on PCV-2 evolution, the presence of a highly divergent PCV2 cluster can be considered interesting *per se*, independently if originating through evolution or recombination.

The present study results confirm the well-known pattern of PCV-2 genotype natural history, characterized by an initial highest detection frequency of PCV-2a (Cluster3), which was outclassed by PCV-2b (Cluster11) and afterwards by PCV-2d (Cluster13). Differently from those major genotypes, the other ones showed a more restricted distribution. Unfortunately, the causes of the differential evolutionary fitness remain elusive and should deserve more focused investigations since they could contribute to the understanding of PCV-2 success (and potentially virulence) determinants.

Although the available sequence number approximately mirrors the overall PCV-2 genotypes population dynamics [14], it can hardly be considered and adequate proxy, being evident a relevant sampling bias due to the different sequencing intensity over time. This issue is particularly clear when combined with the concomitant location-based bias. For example, Asian available sequence number constantly rose over time. Thus, Asia became the most represented continent in the last 10 years, followed by North America, where a higher number of strains collected after 2011 was sequenced.

On the contrary, the amount of sequences provided by European countries is extremely limited, especially in recent years. Interestingly, PCV-2d (Cluster13) was the most commonly reported genotype in Asia, followed by North America, while its presence in Europe appeared low. Based on these evidences, it is hard to understand if the differences observed in the geographical distribution are actually due to a certain spatial clustering or can be mostly explained by sampling bias. The herein described scenario impedes any meaningful comparison among locations and, therefore, among the different applied control strategies, causing a lack of information that could be of pivotal importance for swine farming.

Similar conclusions can be drawn for the host distribution. In this case, most of the non *Sus scrofa domesticus* derived sequences originated from a limited number of countries. Therefore,

even if an apparent over-representation of certain clusters was detected for wild boar, to confidentially establish if this association is due to actual biological reasons or to the sampling structure is almost impossible.

In conclusion, obtained results emphasize the limitations of the current PCV-2 molecular epidemiology knowledge and warrant the implementation of more effective monitoring measures. A unified classification scheme, proposed in the present study, surely represents a fundamental tool, establishing a "common language" among different research groups and diagnostic laboratories. At the same time, a huge effort of will must be set in place to guarantee a more organic and structured sampling activity and to encourage the sharing of properly annotated sequences in freely available databases.

Supporting information

S1 Fig. Complete dataset based phylogenetic tree. Neighbor-Joining phylogenetic tree reconstructed based on a complete collection of strains representative of the proposed PCV-2 clusters (color-coded). The bootstrap support has been displayed as a color-coded (from white (low) to black (high)) circle drawn at the corresponding node.
(PDF)

S2 Fig. Mosaic plot depicting the relationship between Cluster and Host. The area of each cell is proportional to the count size. Cells have been color-coded and lines dotted based on standardized residuals (a standardized residual greater than 2 or lower than -2 is indicative of statistical significance). For graphical reasons, only the more relevant clusters and host categories have been plotted.
(PDF)

S3 Fig. Workflow of genotype classification. A schematic workflow of the criteria to be used for strain classification and new genotype definition is provided.
(PDF)

S1 Table. List of sequences belonging to each cluster and relative metadata. List of ORF2 sequences classified according to the respective cluster and genotype (when present). Additional metadata including collection date, host and continent are reported.
(XLSX)

Author Contributions

Conceptualization: Giovanni Franzo, Joaquim Segalés.

Formal analysis: Giovanni Franzo.

Methodology: Giovanni Franzo.

Supervision: Joaquim Segalés.

Writing – original draft: Giovanni Franzo, Joaquim Segalés.

Writing – review & editing: Giovanni Franzo, Joaquim Segalés.

References

1. Olvera A, Cortey M, Segalés J. Molecular evolution of porcine circovirus type 2 genomes: Phylogeny and clonality. Virology. 2007; 357: 175–185. https://doi.org/10.1016/j.virol.2006.07.047 PMID: 16963096

2. Segalés J, Olvera A, Grau-Roma L, Charreyre C, Nauwynck H, Larsen L, et al. PCV-2 genotype definition and nomenclature. Vet Rec. 2008; 162: 867–868. https://doi.org/10.1136/vr.162.26.867 PMID: 18587066

3. Dupont K, Nielsen EO, Bækbo P, Larsen LE. Genomic analysis of PCV2 isolates from Danish archives and a current PMWS case-control study supports a shift in genotypes with time. Vet Microbiol. 2008; 128: 56–64. https://doi.org/10.1016/j.vetmic.2007.09.016 PMID: 17996404

4. Franzo G, Cortey M, Castro AMMG de, Piovezan U, Szabo MPJ, Drigo M, et al. Genetic characterisation of Porcine circovirus type 2 (PCV2) strains from feral pigs in the Brazilian Pantanal: An opportunity to reconstruct the history of PCV2 evolution. Vet Microbiol. 2015; 178: 158–162. https://doi.org/10.1016/j.vetmic.2015.05.003 PMID: 25975522

5. Liu X, Wang FX, Zhu HW, Sun N, Wu H. Phylogenetic analysis of porcine circovirus type 2 (PCV2) isolates from China with high homology to PCV2c. Arch Virol. 2016; 161: 1591–1599. https://doi.org/10.1007/s00705-016-2823-x PMID: 27016927

6. Guo LJ, Lu YH, Wei YW, Huang LP, Liu CM. Porcine circovirus type 2 (PCV2): Genetic variation and newly emerging genotypes in China. Virol J. 2010; 7: 273. https://doi.org/10.1186/1743-422X-7-273 PMID: 20955622

7. Xiao CT, Halbur PG, Opriessnig T. Global molecular genetic analysis of porcine circovirus type 2 (PCV2) sequences confirms the presence of four main PCV2 genotypes and reveals a rapid increase of PCV2d. J Gen Virol. 2015; 96: 1830–1841. https://doi.org/10.1099/vir.0.000100 PMID: 25711965

8. Opriessnig T, Xiao CT, Gerber PF, Halbur PG. Emergence of a novel mutant PCV2b variant associated with clinical PCVAD in two vaccinated pig farms in the U.S. concurrently infected with PPV2. Vet Microbiol. 2013; 163: 177–183. https://doi.org/10.1016/j.vetmic.2012.12.019 PMID: 23305615

9. Harmon KM, Gauger PC, Zhang J, Piñeyro PE, Dunn DD, Chriswell AJ. Whole-Genome Sequences of Novel Porcine Circovirus Type 2 Viruses Detected in Swine from Mexico and the United States. Genome Announc. 2015; 3: e01315–15. https://doi.org/10.1128/genomeA.01315-15 PMID: 26679573

10. Davies B, Wang X, Dvorak CMT, Marthaler D, Murtaugh MP. Diagnostic phylogenetics reveals a new Porcine circovirus 2 cluster. Virus Res. 2016; 217: 32–37. https://doi.org/10.1016/j.virusres.2016.02.010 PMID: 26948261

11. Bao F, Mi S, Luo Q, Guo H, Tu C, Zhu G, et al. Retrospective study of porcine circovirus type 2 infection reveals a novel genotype PCV2f. Transbound Emerg Dis. 2018; 65: 432–440. https://doi.org/10.1111/tbed.12721 PMID: 28963755

12. Opriessnig T, Xiao CT, Gerber PF, Halbur PG, Matzinger SR, Meng XJ. Mutant USA strain of porcine circovirus type 2 (mPCV2) exhibits similar virulence to the classical PCV2a and PCV2b strains in caesarean-derived, colostrum-deprived pigs. J Gen Virol. 2014; 95: 2495–2503. https://doi.org/10.1099/vir.0.066423-0 PMID: 25053562

13. Firth C, Charleston MA, Duffy S, Shapiro B, Holmes EC. Insights into the Evolutionary History of an Emerging Livestock Pathogen: Porcine Circovirus 2. J Virol. 2009; 83: 12813–12821. https://doi.org/10.1128/JVI.01719-09 PMID: 19812157

14. Franzo G, Cortey M, Segalés J, Hughes J, Drigo M. Phylodynamic analysis of porcine circovirus type 2 reveals global waves of emerging genotypes and the circulation of recombinant forms. Mol Phylogenet Evol. 2016; 100: 269–280. https://doi.org/10.1016/j.ympev.2016.04.028 PMID: 27114187

15. Carman S, McEwen B, DeLay J, van Dreumel T, Lusis P, Cai H, et al. Porcine circovirus-2 associated disease in swine in Ontario (2004 to 2005). Can Vet J = La Rev Vet Can. 2006; 47: 761–2. Available: http://www.ncbi.nlm.nih.gov/pubmed/16933552

16. Cortey M, Pileri E, Sibila M, Pujols J, Balasch M, Plana J, et al. Genotypic shift of porcine circovirus type 2 from PCV-2a to PCV-2b in Spain from 1985 to 2008. Vet J. 2011; 187: 363–368. https://doi.org/10.1016/j.tvjl.2009.12.023 PMID: 20106683

17. Timmusk S, Wallgren P, Brunborg IM, Wikström FH, Allan G, Meehan B, et al. Phylogenetic analysis of porcine circovirus type 2 (PCV2) pre- and post-epizootic postweaning multisystemic wasting syndrome (PMWS). Virus Genes. 2008; 36: 509–520. https://doi.org/10.1007/s11262-008-0217-1 PMID: 18343985

18. Wiederkehr DD, Sydler T, Buergi E, Haessig M, Zimmermann D, Pospischil A, et al. A new emerging genotype subgroup within PCV-2b dominates the PMWS epizooty in Switzerland. Vet Microbiol. 2009; 136: 27–35. https://doi.org/10.1016/j.vetmic.2008.10.028 PMID: 19081207

19. Correa-Fiz F, Franzo G, Llorens A, Segalés J, Kekarainen T. "Porcine circovirus 2 (PCV-2) genetic variability under natural infection scenario reveals a complex network of viral quasispecies." Sci Rep. 2018: 8(1), p.15469. https://doi.org/10.1038/s41598-018-33849-2 PMID: 30341330

20. Abascal F, Zardoya R, Telford MJ. TranslatorX: Multiple alignment of nucleotide sequences guided by amino acid translations. Nucleic Acids Res. 2010; 38: W7–W13. https://doi.org/10.1093/nar/gkq291 PMID: 20435676

21. Martin DP, Murrell B, Golden M, Khoosal A, Muhire B. RDP4: Detection and analysis of recombination patterns in virus genomes. Virus Evol. 2015; 1: 1–5. https://doi.org/10.1093/ve/vev001

22. Paradis E, Bolker B, Claude J, Cuong HS, Desper R, Durand B, et al. ape: Analyses of Phylogenetics and Evolution. Bioinformatics. 2014; 20: 289–290. Available: http://cran.r-project.org/web/packages/ape/index.html

23. Team RC. R A Lang Environ Stat Comput Found Stat Comput Vienna, Austria2013. ISBN 3-900051-07-0; 2014;

24. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol. 2016; 33: 1870–4. https://doi.org/10.1093/molbev/msw054 PMID: 27004904

25. Ragonnet-Cronin M, Hodcroft E, Hué S, Fearnhill E, Delpech V, Brown AJL, et al. Automated analysis of phylogenetic clusters. BMC Bioinformatics. 2013; 14: 317. https://doi.org/10.1186/1471-2105-14-317 PMID: 24191891

26. Cock PJA, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, et al. Biopython: Freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics. 2009; 25: 1422–1423. https://doi.org/10.1093/bioinformatics/btp163 PMID: 19304878

27. Guo L, Fu Y, Wang Y, Lu Y, Wei Y, Tang Q, et al. A porcine circovirus type 2 (PCV2) mutant with 234 amino acids in Capsid protein showed more virulence in vivo, compared with classical PCV2a/b strain. PLoS One. 2012; 7: e41463. https://doi.org/10.1371/journal.pone.0041463 PMID: 22829951

28. Xiao C-T, Halbur PG, Opriessnig T. Complete Genome Sequence of a Novel Porcine Circovirus Type 2b Variant Present in Cases of Vaccine Failures in the United States. J Virol. 2012; 86: 12469–12469. https://doi.org/10.1128/JVI.02345-12 PMID: 23087125

29. Franzo G, Tucciarone CM, Cecchinato M, Drigo M. Porcine circovirus type 2 (PCV2) evolution before and after the vaccination introduction: A large scale epidemiological study. Sci Rep. 2016; 6: 39458. https://doi.org/10.1038/srep39458 PMID: 27991573

30. Grau-Roma L, Stockmarr A, Kristensen CS, Enøe C, López-Soria S, Nofrarías M, et al. Infectious risk factors for individual postweaning multisystemic wasting syndrome (PMWS) development in pigs from affected farms in Spain and Denmark. Res Vet Sci. 2012; 93: 1231–1240. https://doi.org/10.1016/j.rvsc.2012.07.001 PMID: 22884005

31. Franzo G, Cortey M, Olvera A, Novosel D, De Castro AMMG, Biagini P, et al. Revisiting the taxonomical classification of Porcine Circovirus type 2 (PCV2): Still a real challenge. Virol J. 2015; 12: 131. https://doi.org/10.1186/s12985-015-0361-x PMID: 26311322

32. Manrubia SC. Modelling viral evolution and adaptation: Challenges and rewards. Curr Opin Virol. 2012; 2: 531–537. https://doi.org/10.1016/j.coviro.2012.06.006 PMID: 22796325

33. Holmes EC. The comparative genomics of viral emergence. Proc Natl Acad Sci. 2010; 107: 1742–1746. https://doi.org/10.1073/pnas.0906193106

34. Holmes EC. The Evolutionary Genetics of Emerging Viruses. Annu Rev Ecol Evol Syst. Annual Reviews; 2009; 40: 353–372. https://doi.org/10.1146/annurev.ecolsys.110308.120248

35. Cadar D, Cságola A, Lorincz M, Tombácz K, Spînu M, Tuboly T. Detection of natural inter- and intra-genotype recombination events revealed by cap gene analysis and decreasing prevalence of PCV2 in wild boars. Infect Genet Evol. 2012; 12: 420–427. https://doi.org/10.1016/j.meegid.2012.01.014 PMID: 22293465

36. Anoopraj R, Rajkhowa TK, Cherian S, Arya RS, Tomar N, Gupta A, et al. Genetic characterisation and phylogenetic analysis of PCV2 isolates from India: Indications for emergence of natural inter-genotypic recombinants. Infect Genet Evol. 2015; 31: 25–32. https://doi.org/10.1016/j.meegid.2015.01.006 PMID: 25592870