Available online at www.sciencedirect.com

ScienceDirect

Virology 357 (2007) 175–185

VIROLOGY

www.elsevier.com/locate/yviro

Molecular evolution of porcine circovirus type 2 genomes: Phylogeny and clonality

A. Olvera [a,1], M. Cortey [b,1], J. Segalés [a,c,*]

[a] *Centre de Recerca en Sanitat Animal, Universitat Autonoma de Barcelona, Edifici CReSA, Campus Bellaterra, Cerdanyola del Valles (Barcelona), 08193, Spain*

[b] *Bofill i Codina 14, Calella de P, (Girona), 17210, Spain*

[c] *Departament de Sanitat Animal, Facultat de Veterinaria, Universitat Autonoma de Barcelona, Campus Bellaterra, Cerdanyola del Valles (Barcelona), 08193, Spain*

Received 6 June 2006; returned to author for revision 18 July 2006; accepted 31 July 2006

Available online 11 September 2006

Abstract

Porcine circoviruses (PCVs) type 1 (PCV1) and type 2 (PCV2) show high levels of nucleotide similarity, but PCV1 is considered non-pathogenic and PCV2 has been associated with several disease outcomes in pigs, mainly postweaning multisystemic wasting syndrome (PMWS). After exploring different topologies of the origin of PCVs, it was concluded that PCV1 and PCV2 seem to have a common origin. On the other hand, PCV2 could be divided into two groups (1 and 2) and eight clusters (1A to 1C and 2A to 2E), but none of those was apparently associated with disease status or geographic area. When phylogenetic trees constructed with the whole PCV2 genome, the *cap* or the *rep* genes were compared, some incongruence was identified. The possible existence of recombination was evaluated and cluster 1B was found to have a possible recombinant origin. Selective pressure was detected in all parts of the PCV2 genome, especially in the *rep* gene. Finally, the *cap* gene was the more suitable phylogenetic and epidemiological marker for PCV2, despite the fact that the virus can undergo recombination mainly within the first part of the *rep* region.

Keywords: Porcine circoviruses; Phylogeny; Evolution; Clonality; Recombination; Postweaning multisystemic wasting syndrome; Swine

Introduction

Two porcine circoviruses (PCVs), from the *Circoviridae* family, have been described to date, namely porcine circovirus type 1 (PCV1) and type 2 (PCV2) (Allan et al., 1999b). PCVs are the smallest animal viruses known; PCV2 has a 1768 nucleotide genome of circular, single stranded DNA, and the genome of PCV1 is even shorter, with 1759 nucleotides (Hamel et al., 1998; Meehan et al., 1998). Although PCVs are closely related, PCV1 is considered non-pathogenic (Allan et al., 1994; Tischer et al., 1986) while PCV2 has been described as the necessary infectious factor, although not proven as the sole factor, to develop postweaning multisystemic wasting syndrome (PMWS) (Allan et al., 1999a; Krakowka et al., 2001). PMWS is clinically suspected in outbreaks of respiratory clinical signs and wasting in pigs and pathologically characterized by lymphocyte depletion and granulomatous inflammation of lymphoid tissues (Rosell et al., 1999). PCV2 has also been related to a number of disease syndromes in pigs (Segales et al., 2004); the terminology porcine circovirus diseases (PCVD) was proposed to include those conditions that had been linked with PCV2 infection (Allan et al., 2002). Even though PCV1 and PCV2 seem to be very different from a pathogenic point of view, both viruses are very similar at the genomic level since they have about 80% overall nucleotide sequence identity (Hamel et al., 1998; Mankertz et al., 2000; Meehan et al., 1998). The possibility of these PCVs having a common antecessor or one deriving from the other has never been explored.

* Corresponding author.
E-mail address: Joaquim.Segales@cresa.uab.es (J. Segalés).

[1] A. Olvera and M. Cortey contributed equally to the authorship of this work.

0042-6822/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2006.07.047

From the first described PCV2 sequence by Hamel et al. (1998) until September 2005, up to 150 PCV2 complete genomes have been submitted to the GenBank nucleotide database at the National Center for Biotechnology Information, NCBI (http://www.ncbi.nlm.nih.gov). Despite the high number of sequences available from GenBank, no extensive phylogenetic studies have been performed; this is probably due to the high level of homology (≥94%) between PCV2 genomes (Larochelle et al., 2002; Mankertz et al., 2000; Meehan et al., 2001). This high level of similarity between genomes could theoretically prevent the detection of significant differences among strains. Moreover, both PCVs genomes show the typical organization of the *Circoviridae* family, with a replicase (*rep*) gene in the sense strand and a capsid (*cap*) gene in the antisense strand (Cheung, 2003a; Mankertz et al., 2004). Furthermore, the *rep* gene has a splicing alternative *rep′*, in which the central part of *rep* (378 nucleotides) is eliminated as an intron (Cheung, 2003b; Mankertz and Hillenbrand, 2001).

To date, several genotyping studies on PCV2 have been published (De Boisseson et al., 2004; Grierson et al., 2004a, 2004b; Larochelle et al., 2002; Wen et al., 2005), but no evaluation of the utility of the *cap* and the *rep* genes as molecular markers or assessment of potential PCV2 clonality have been undertaken. Furthermore, phylogenetic and evolutionary studies are of increasing importance in molecular epidemiological studies on viral and bacterial pathogens.

Therefore, the main objectives of the present study were (i) to determine the ancestral status of PCV1 or PCV2 within the *Circorviridae* family; (ii) to find out clonal relationships among PCV2 complete genomes; and (iii) to evaluate the capsid (*cap*) and replicase (*rep*) genes as phylogenetic and epidemiological markers for PCV2.

Results

Ancestral status

All phylogenetic trees pointed to the same topology among the assayed *Circoviridae* family sequences. Three main monophylethic lineages supported by high bootstrap values were differentiated: one for Beak and Feather Disease Virus grouped with Columbid, Starling and Canary Circoviruses, another for Duck, Mallard Duck, Muscovy Duck and Goose Circoviruses and a monophyletic branch for PCVs separated from avian circoviruses. Both PCV types were also clearly separated into two different monophyletic clusters (Fig. 1). After evaluation of the six proposed topologies based on the maximum likelihood, maximum parsimony and minimum evolution criteria, log-likelihoods, SBL and the number of steps all undoubtedly pointed to the dichotomy topology as the best tree. These results suggest a common origin as the most reliable possibility for PCVs (data not shown).

Genome variation and utility of PCV2 sequences

Saturation effects in the whole genome and for the first, second and third codon positions of the *cap* and *rep* genes data sets were evaluated. In all cases, the number of observed transversions relative to that of transitions gradually decreased with increasing divergence; moreover, data sets followed a line, not a curve, indicating that transitions and transversions were not saturated and that they can be used in phylogenetic reconstructions. In addition, Xia's test supported little saturation for all data sets.

From 148 PCV2 genomes, 479 out of 1768 nucleotide sites were variable, with 604 mutations detected: 330 transitions, 272

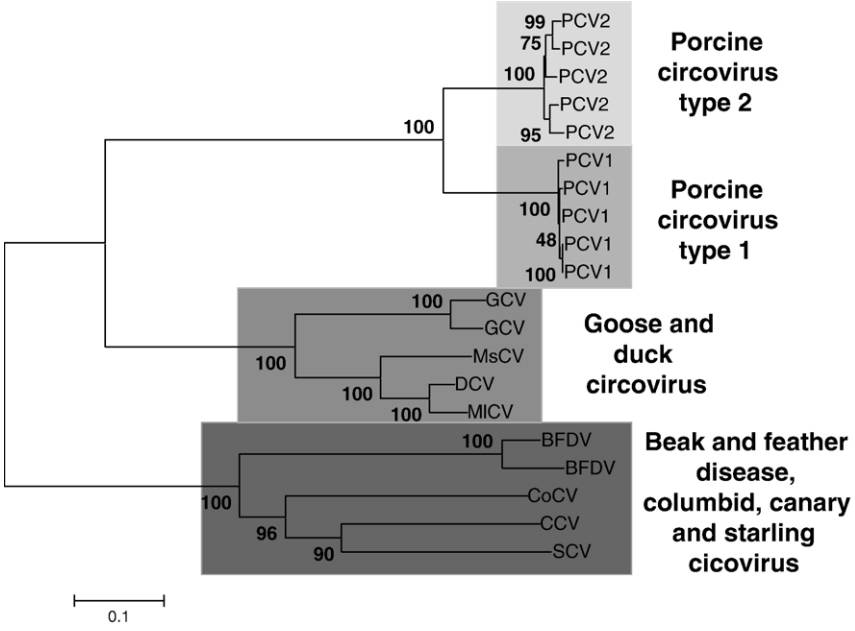


Fig. 1. Evolutive relationships among Circovirus groups. Phylogenetic tree based on the NJ method for the 20 *Circoviridae* sequences used (Appendix A). Identical tree topologies, as well as equal percentage of support for the internal branches, are reported with the ML and MP tree-building methods. Scores of the best topology: −lnL=20,815, SBL=3.136, Tree Length=4855, Consistency Index=0.711.

transversions and 2 indels. From these mutations, 252 were detected in the *cap* gene and 238 in the *rep* gene. As expected, both genes presented a biased pattern of mutation, as indicated by low gamma parameters ($\alpha_{Cap}=0.33$, $\alpha_{Rep}=0.33$, $\alpha_{Genome}=0.21$); third codon positions showed more variable sites (171) than the first (121) and second ones (100).

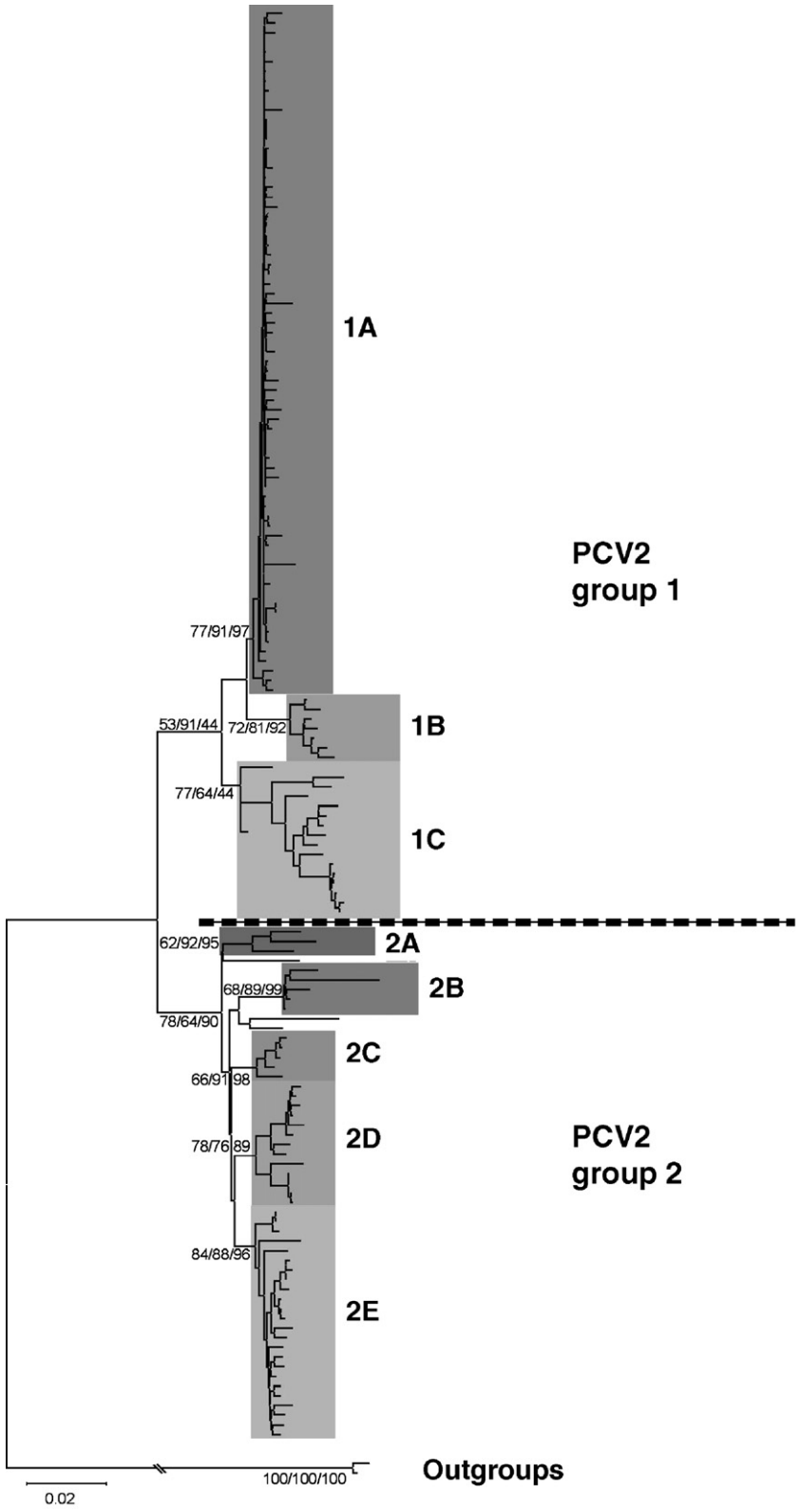


Fig. 2. Phylogenetic tree based on the NJ method for the 148 PCV2 sequences plus 2 outgroups used in the study. Numbers along the branches refer to the percentages of confidence in the ML, MP and NJ analyses. Sequences included in each cluster are listed in Appendix B.

Phylogenetic relationships among PCV2 sequences

A NJ tree for PCV2 genomes with confidence values for ML, MP and NJ methods is shown in Fig. 2. All three algorithms reported congruent results, and the groupings were supported by moderate to high confidence values. Sequences could be divided into two main groups: group 1, which was further subdivided in 3 clusters (1A to 1C), and group 2, subdivided in 5 clusters (2A to 2E). Three PCV2 sequences remained unclustered in group 2, two of which were in the same branch (supported by moderate confidence values). The average distance within groups 1 and 2 was 0.0158 and 0.0234, respectively, and it was 0.0444 between groups. Within clusters, the average distances ranged from 0.005 for cluster 1A to 0.023 for cluster 2D; between clusters, those distances ranged from 0.015 (cluster 1A vs 1B) to 0.049 (cluster 1C vs 2E). Groups were defined by 52 marker positions, the distribution of which was highly skewed: 33 located in the *cap* gene, 13 in the *rep* gene and 6 in the rest of the genome (indel, *cap* and *rep* genes marker positions are indicated in Table 1). In addition, an indel position just after the stop codon of the *cap* gene (genome position 1033, Table 1) differentiated PCV2 genomes of group 2 (1768 nucleotides) from those of group 1 (1767 nucleotides). The genome size was very consistent for each group, and the few disagreements (two out of 148) were attributed to indel positions inside one of the two ORFs. Moreover, cluster 1C was characterized by another indel position in the last codon of the *cap* gene (genome position 1036), changing the reading frame. This *cap* gene finished one codon further in another stop codon, being one lysine longer (Table 1).

No discrepancy in the definition of groups and clusters was observed comparing the three phylogenetic methods assayed between the *cap* gene and the whole PCV2 genome. Nevertheless, partial trees based on the *rep* gene presented some incongruence compared with the *cap* and the complete genome trees. Several genomes ($n=22$), belonging to clusters 1B and 1C, were grouped into group 1 based on the whole genome and the *cap* gene. Surprisingly, those genomes were grouped into group 2 with the *rep* gene. It is noteworthy that this feature was reported in all phylogenetic reconstructions.

Recombination analysis

After running the RDP software, 4685 potential recombination events were detected. Considering only statistically significant events reported by more than one algorithm, possible recombination events between cluster 1A and sequences belonging to different clusters of group 2 were consistently reported. All PCV2 genomes from cluster 1B were found to be daughter genomes of recombination events that are listed in Fig. 3. In order to summarize the results, only the cluster of the parent genomes was reported and putative breakpoints (BP) were grouped in 11 areas. Clusters 1A and 2D were mainly identified as the major and minor parents. Nevertheless, clusters 1C, 2B and 2A were occasionally reported as minor parents. Supporting this, a close inspection of marker positions showed that cluster 1B had the same marker positions as cluster 1A

Table 1
Indel, *cap* and *rep* gene marker positions are summarized in this table indicating if they are located in the first, second or third codon positions

Rep in sense

Gene													
Gene position	39	60	102	111	204	288	339	501	666	687	702	711	822
Genome position	89	110	152	161	254	338	389	551	716	737	752	761	872
Cluster 1A	C	T	T	A	G	C	G	T	C	G	A	A	T
Cluster 1B	A	G	G	C	A	T	A	C	C	G	A	A	T
Cluster 1C	A	G	G	C	A	T	A	C	G	C	T	T	C
Group 2	A	G	G	C	A	T	A	C	G	C	A	T	C
Codon Position	3	3	3	3	3	3	3	3	3	3	3	3	3
Region				1						2			3

IG and STOP in antisense

Gene	1024–1032	1033–1035
Genome position	1024 1032	1033 1035
Cluster 1A	TATTATTCA	**-TTA**
Cluster 1B	TATTATTCA	**-TTA**
Cluster 1C	TATTAT**TCA**	CTT-
Group 2	TATTATTCA	**TTTA**

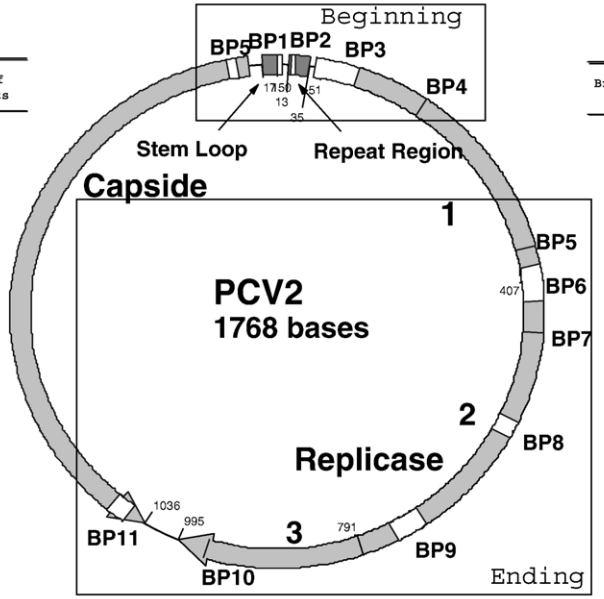
Cap in antisense

Gene																																	
Gene position	699	696	630	567	549	540	531	528	513	510	506	501	480	477	451	408	273	271	268	266	265	264	263	262	258	256	255	246	238	229	222	177	27
Genome position	1036	1040	1106	1169	1187	1196	1205	1208	1223	1226	1230	1235	1256	1259	1285	1328	1463	1465	1468	1470	1471	1472	1473	1474	1478	1480	1481	1490	1498	1507	1514	1559	1709
Cluster 1A	A	G	T	A	C	T	G	T	A	A	G	T	G	G	T	G	C	C	A	C	G	G	G	G	T	A	G	T	G	T	G	T	C
Cluster 1B	A	G	T	A	C	T	G	T	A	A	G	T	G	G	T	G	C	C	A	C	G	G	G	G	T	A	G	T	G	T	G	T	C
Cluster 1C	G	T	G	A	A/CC		A	T	A	A	G	A	A	G	G	T	C	C	T	C	G	G	G	G	T	A	G	T	G	T	A	A	T
Group 2	G	T	G	G	G	T	A	G	G	T	C	A	G	A	G	T	T	T	T	A	T	T	T	T	G	T	C	C	C	C	A	A	G
Codon Position	3	3*	3	3	3	3	3	3	3	3	2*	3	3	3	1*	3	3	1*	1*	2*	1*	3*	2*	1*	3*	1*	3	3	1*	1*	3	3	3
Region	D						C						B																	A			

Indel positions are located between the *cap* and *rep* gene (intergenic region, IG) after the *cap* stop codon in antisense. Those indel positions are indicated presenting the consensus sequences between positions 1024 and 1035. Marker positions leading to an amino acid change are indicated with an asterisk. Marker positions of group one are indicated in dark gray, of cluster 1C in light gray and of group two in white. The positions in each gene and in the genome are indicated. The three replicase regions (numbers) and the four capsid putative antigenic regions (letters) are indicated at the bottom.

Ending

Breakpoint Region	Major Parent Cluster	Minor Parent Cluster	Identified by	Position of Breakpoints	N° of events
BP4	1A	2D	RDP, Maxchi, Geneconv, Bootscan	167	7
BP6	1A	2D(17), 1C (9), 2E(1), 2A(1)	RDP, Maxchi, Geneconv, Bootscan, Siscan	389(13), 407(5), 412(3), 420(3), 429(2)	26
BP7	1A	1C	Bootscan, Siscan	466(4)	4
BP8	1A	2D(2), 2E(2), 1C(1)	RDP, Geneconv, Siscan	567(1), 575(1), 578(1), 580(1), 587(1)	5
BP9	1A	2G	RDP, Geneconv	716(1), 751(1)	2
BP10	1A	2D(4), 2E(1)	RDP, Maxchi	963	5
BP11	1A	2E(3)	Maxchi	1057(1), 1078(1), 1167(1)	3



Beginning

Breakpoint Region	Major Parent Cluster	Minor Parent Cluster	Identified by	Position of Breakpoints	N° of events
BP1	1A	2D	Geneconv, Bootscan	1(4), 6(1)	5
BP2	1A	2D(4), 1C(5)	RDP, Bootscan, Siscan	16(4), 20(4), 34(1)	9
BP3	1A	2D(11), 2E(7), 1C(8)	RDP, Maxchi, Geneconv, Bootscan, Siscan	42(2), 56(1), 58(1), 62(14), 67(1), 76(2), 79(1), 89(3)	25
BP4	1A	2D	RDP, Maxchi, Geneconv, Bootscan	167 (6)	6
BP5	1A	2D(5), 2A(1)	RDP	1723(4), 1727(1), 1712(1)	6

Fig. 3. Results of the search for recombination events among PCV2 genomes using RDP, MaxChi, GeneConv, BootScan, SiScan and Chimera algorithms. The location of beginning and ending breakpoints is indicated in a PCV2 genome scheme. Numbers represent the three regions defined for the *rep* genes. Putative recombination breakpoint regions are also indicated (BP). Following the annotation of genome PCV2-B (Accession Number AF112862), BP1 is located after the origin of replication, BP2 at the repeat region, BP3, BP4 and from BP6 to BP10 in the *rep* gene, finally, BP5 and BP11 are located in the *cap* gene. Beginning and ending breakpoint regions are indexed in two tables. Cluster of the parent genomes are indicated in the tables with the algorithms with significant results and the breakpoints in each region. The number of events detected is indicated in brackets.

except for *rep* gene up to marker position 501 (Table 1). At the beginning of the *rep* gene, this cluster showed the same marker positions as group 2. No recombination events pointing to the same direction were detected for cluster 1C genomes. Looking at the marker positions, cluster 1C had the same marker positions as group 2 PCV2 genomes in the *rep* gene, but it had different marker positions in the *cap* gene, where those were more abundant. Furthermore, putative beginning BP were placed at the beginning of the *cap* gene (BP5), the region adjacent to the origin of replication (BP1), the repeat region (BP2) and the beginning of the *rep* gene until position 167 (BP3 and BP4) as annotated for the PCV2-B genome (Accession Number AF112862). Possible ending BP were found between positions 167 and 963 of the *rep* gene or at the ending of the *cap* gene (Fig. 3). Only one putative beginning and one putative ending BP were reported by five algorithms. The putative recombination beginning BP was BP1, between positions 42 and 89, at the beginning of the *rep* gene. The best supported putative ending BP was BP3 between positions 389 and 429; curiously this region includes the splice donor for the *rep′* transcript (Mankertz and Hillenbrand, 2001). This result points to the first part of the *rep* gene, corresponding to the first exon of the *rep′* transcript, which has the most likely recombinant region. Potential recombination beginning BP were concentrated at the origin region and the beginning of the *rep* gene while the putative recombination ending BP were spread over a wide part of the *rep* gene, more precisely the intron and the second exon of the *rep′* transcript until the *cap* gene. Finally, position 167 was detected as a BP by four algorithms, although it could be a beginning as well as an ending region.

Selective pressure

In the difference between non-synonymous and synonymous substitution rates (d*N*−d*S*) for the whole genome, the *cap* and the *rep* genes were all negative (−0.0174±0.0053, −0.0744±0.0144 and −0.0604±0.0097, respectively), indicating that the PCV2 genome was under purifying selection. In order to identify regions under strong negative selection, the d*N*−d*S* for every codon was plotted against the amino acid entropy, at the same point, for both *cap* and *rep* genes (Fig. 4). As already said before, the *rep* gene has a splicing alternative, *rep′*, and it was

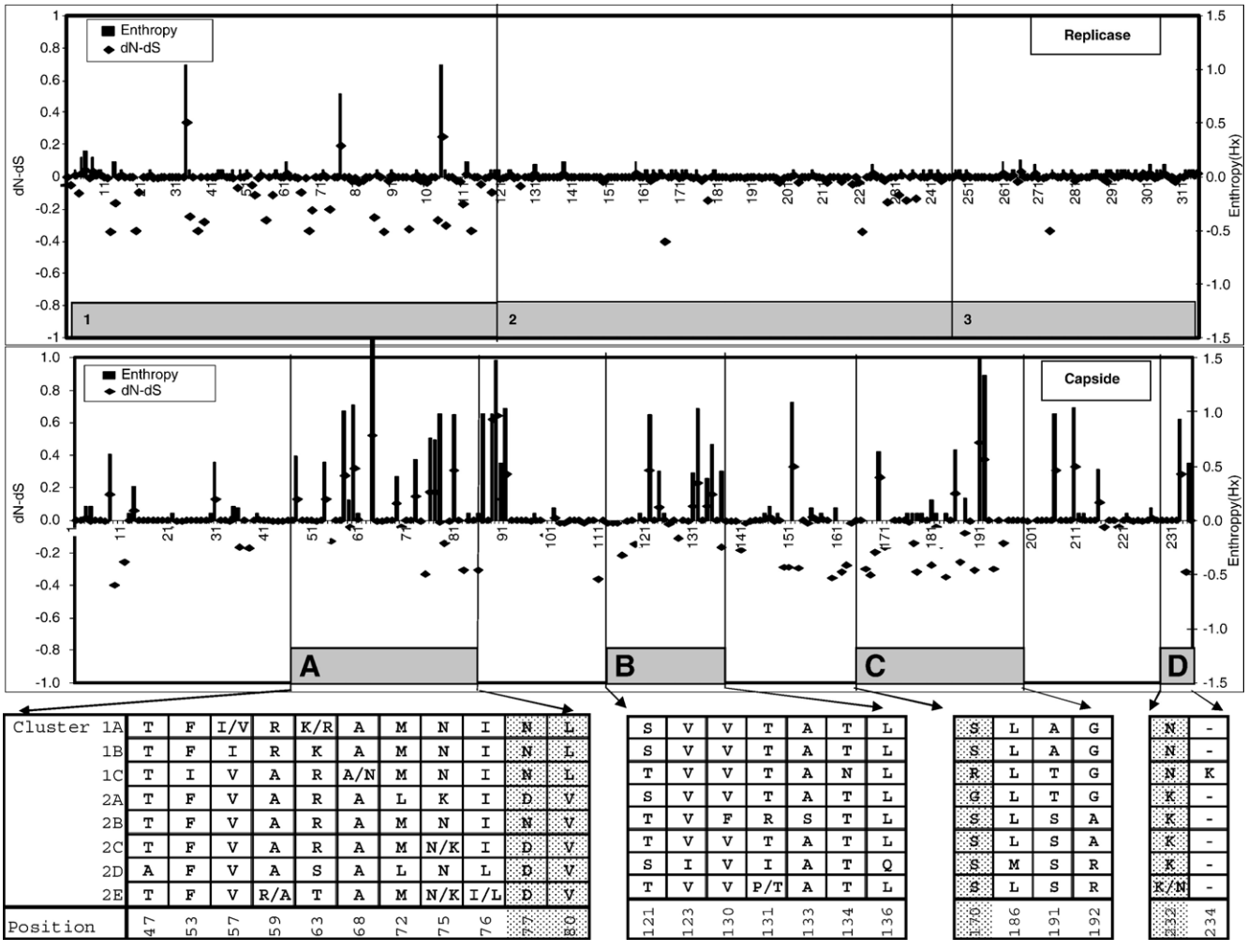


Fig. 4. Difference between non-synonymous and synonymous rates (d*N*−d*S*) and amino acid entropy rates plot for the replicase and the capsid. For the replicase gene, the three regions are indicated. For the capsid gene, the variable amino acid for each cluster is also indicated. Amino acid changes corresponding to a marker position are highlighted in gray.

divided into three regions. Region 1 corresponded to the common part of the *rep* and *rep′* transcripts; region 2 was the intron eliminated by splicing to create the *rep′* transcript; and region 3 was the final fragment of *rep* and *rep′*. The latter region is translated in different reading frames for *rep* and *rep′* (Mankertz et al., 2004; Mankertz and Hillenbrand, 2002). The gamma parameters for those three regions were clearly different (0.19, 0.32 and 97.03), indicating a highly skewed substitution pattern for the first two regions and a more uniform substitution pattern for the last one. The *cap* gene was divided in immunogenic and not immunogenic regions, based on four antigenic domains or epitopes already defined by the experiments performed by Lekcharoensuk et al. (2004) and Mahe et al. (2000) (Fig. 3). Most amino acid positions in all epitopes were under neutral selection. Amino acid positions under selection in epitopes A and B showed more amino acid positions under positive selection (10 and 6 respectively) than under negative selection (6 and 5 respectively). Besides, in epitope C, almost all amino acid positions were under negative selection (16). Finally, one amino acid position was neutral, one positively selected and one negatively selected in epitope D, which was one amino acid longer in cluster C. The region between epitopes was mainly under neutral selection. Curiously, five amino acid positions adjacent to epitope A were positively selected. As said above, few amino acid changes were detected inside the epitopes: 11 out of 37 in epitope A, 7 out of 26 in epitope B, 4 out of 35 in epitope C and 2 out of 4 in epitope D. It has to be pointed out that four of those amino acid changes (two in epitope A, one in epitope C and one in epitope D) were associated with marker positions (Fig. 3).

Discussion

PCVs are very small animal viruses, with single stranded DNA genomes of less than 2 kb, which show a high level of homology despite their apparently different pathogenicity (Allan et al., 1995). PCV1 is the only non-pathogenic member of the *Circoviridae* family (Allan et al., 1995; Tischer et al., 1986), while PCV2 is involved in the PCVD (Allan et al., 2002). The high degree of similarity between PCVs raised the question about their origin and the possibility of PCV1 derived from an ancestral pathogenic PCV type or PCV2 derived from a non-pathogenic ancestor. Unfortunately, with the sequences used in this study, an ancestral status for PCVs could not be established and topological evaluations gave support to a common origin for both types.

The lowest identity reported in this study between two PCV2 genomes was 0.93. This high sequence homology probably discouraged detailed phylogenetic studies in the past in order to find clonal lineages among PCV2 genomes. Nevertheless, phylogenetic analysis of the genome clearly indicated that PCV2 sequences can be divided into two major groups, with several clusters into each one. Both PCV2 groups were homogeneous and several marker positions, mainly located in the *cap* gene, were identified. Furthermore, both groups showed differences in two indel positions and, as a consequence, genomes in group 1 and in group 2 differed in length by one nucleotide. Unfortunately, no correlation regarding groups and health status could be performed, although this is probably due to the lack of this information in some genome submissions. Furthermore, no relationship between group and country of isolation could be defined. Only a relationship between group and year of submission was found (data not shown); however, it is known that a genome sequence can be published a long time after detection. Genomes of group 1 were mainly published in NCBI after 2003 (87 out of 94) and genomes of group 2 were mainly published before 2003 (33 out of 53), indicating that group 1 could be more recent than group 2.

In the comparison among phylogenetic trees constructed with the whole PCV2 genome, the *cap* and *rep* genes indicated a minor incongruence regarding clusters 1B and 1C. One possible explanation for this incongruence is the potential existence of recombination events involving genomes from both PCV2 groups. However, only cluster 1B was clearly identified as having undergone recombination. In the six genomes included in this cluster, recombination events were detected in five out of the six algorithms used, indicating the strength of this observation. These results indicate that genomes owning to cluster 1B are the product of a recombination event between a genome of cluster 1A as the major parent and a genome of group 2 (most probably belonging to cluster 2D) as the minor parent. The best supported recombinant region was the genome fragment between the beginning of the *rep* gene (BP3) and the splicing donor for *rep′* (BP6) including region 1 of the *rep* gene, the first exon of the *rep′* transcript. Putative recombinant ending BP were diffused in a wide part of the PCV2 genome; the latter is probably due to the high level of conservation of regions 2 and 3 of the *rep* gene complicating the determination of an ending BP.

Furthermore, our results showed that the whole PCV2 genome, the *rep* and the *cap* genes were under purifying selection, as indicated by the negative $dN-dS$ values (Nei, 2000). Looking at the codon selection pressure for both genes, the *rep* gene was highly conserved; in contrast, the *cap* gene showed a variety of positively and negatively selected positions. These findings were additionally supported by a low number of marker positions (33 vs 13), and a high proportion of mutations detected in single sequences (65% vs 45%) in the *rep* gene compared with the *cap* gene. Interestingly, very few mutational changes were detected in regions 2 and 3 of the *rep* gene. These differences could be caused by both neutral selection or, most likely, by a very strong purifying selection. Moreover, region 2 is probably limited in sequence changes, not only by the selection on the amino acid sequence, but also by its implication in a splicing process. Furthermore, the promoter of the *cap* gene (P_{cap}) has also been mapped in this region for PCV1, more exactly among nucleotide positions 431 and 507 (Mankertz and Hillenbrand, 2002). In region 3, splicing results in the utilization of an alternative reading frame for *rep′* compared with *rep* (Cheung, 2004; Mankertz and Hillenbrand, 2001). Probably this frame shift between *rep* and *rep′* in region 3 is affecting the usual degeneration of the genetic code, which allows the third codon position to have more changes than the other two. This is corroborated by an unusually high value of the gamma

parameter (97.03) for this region, indicating a homogeneous pattern of nucleotide substitution in region 3, unlike the other two *rep* gene regions. The high selective pressure on this region of the genome could also be an explanation to the high level of conservation inside PCV2. This virus genome is a circular molecule, and the selection on one region can drag a whole genome due to a hitch-hiking effect (Barton, 2000).

The selective pressure on the *cap* gene was of special interest since it is the most exposed part of the virus and is the most likely to interact with the host immune system (Roitt and Male, 1998). As stated before, this gene shows several positions under positive selection and more amino acid changes than in the *rep* gene are possible. In some viruses, epitopes are under positive selection since they are submitted to high immunologic pressure and have a high rate of amino acid changes to evade the immune response (Grenfell et al., 2004). In PCV2, epitope amino acid sequences showed little variable positions although the four proposed epitopes concentrated the majority of the variation of the *cap* gene. Controversially, epitopes A and B had more positively selected positions, but epitope C was mainly under negative selection. In fact, no high rate of amino acid replacement was found, indicating no clear immunological selection (Grenfell et al., 2004). Interestingly, four marker positions that lead to amino acid changes were in one of the four epitopes and, in cluster C, the epitope D is one lysine longer. This is of special interest and could be indicative of potential immunogenic differences among strains from different groups or clusters.

As a whole, we consider that the *cap* gene is a reliable phylogenetic marker for PCV2 strains since it was able to reconstruct the same tree as the whole viral genome. This is because the *cap* gene is under less selective pressure compared with the *rep* gene, accumulating much more variability, and it probably weighs more in the phylogenetic tree reconstructions. The effects of recombination seem to be limited to the first region of the *rep* gene, having little impact on the phylogenetic reconstructions performed using the *cap* gene. Furthermore, since the *cap* gene is shorter, it would be a more suitable molecular marker for epidemiological and phylogenetic studies since its sequencing is less labor extensive than the sequencing of the whole genome. Furthermore, the use of gene sequences in epidemiological studies is of special interest since they are easy to share among laboratories (Foxman et al., 2005).

In conclusion, PCV1 and PCV2 seem to have a common origin. PCV2 genomes can be divided in different clonal lineages (2 groups and 8 clusters), although the possible existence of virulence or immunogenic differences between strains has to be further studied. Finally, the *cap* gene is a suitable phylogenetic and epidemiological marker since it was not affected by recombination, in contrast to the *rep* gene.

Materials and methods

Ancestral status of PCVs

Twenty complete genomes of different animal circoviruses available in GenBank (http://www.ncbi.nlm.nih.gov) were used. Five PCV2 and five PCV1 genomes randomly selected were aligned with representatives of the other members of the *Circoviridae* family, including Goose Circovirus, Canary Circovirus, Beak and Feather Disease Virus and yet unclassified circoviruses (Columbid Circovirus, Duck Circovirus, Mulard Duck Circovirus, Muscovy Duck Circovirus and Starling Circovirus) (Appendix A). To determine the ancestral status of PCV1 or PCV2 within the *Circorviridae* family, six alternative topologies were constructed considering the initial divergence of PCV1 sequences (two topologies), the initial divergence of the PCV2 sequences (three topologies) and a final topology considering a dichotomy between PCV1 and PCV2 groups of sequences. These six topologies were tested using minimum evolution, maximum likelihood and parsimony criteria. Initially, topologies were tested by the modified Kishino and Hasegawa test (Shimodaira, 1999), computing the log-likelihoods per site for each tree and comparing the total log-likelihoods among topologies. Secondly, the sum of branch lengths (SBL) for each proposed topology was computed. Under the minimum evolution criteria, the topology with the smallest SBL was considered the best phylogenetic tree. Finally, the number of steps for each topology was computed, and the one that required the smallest number of steps was chosen to be the best phylogenetic tree.

PCV2 genome sequence data set

The one hundred and forty-eight PCV2 and two PCV1 complete genomes (for rooting purposes) present at the NCBI nucleotide database (http://www.ncbi.nlm.nih.gov) in September 2005 were used to study the clonality of PCV2 and the utility of *cap* and *rep* genes as potential epidemiological markers. All sequences were downloaded, linearized at the same point and aligned with the annotated genome sequence for PCV2-B (Accession Number AF112862, (Hamel et al., 2000)). NCBI accession numbers of the sequences used in this study are listed in Appendix B. Utility of *cap* and *rep* genes was studied only with 137 genomes by deleting eleven sequences containing indel positions in any of the two ORFs (Appendix B) from the database. Those indel positions moved the reading frame, making the comparison among amino acid sequences difficult.

Phylogenetic utility of PCV2 sequences

The loss of phylogenetic information due to substitution saturation in the PCV2 genome was evaluated. The level of saturation was studied by plotting the pairwise number of observed transitions and transversions versus genetic distance. In addition, substitution saturation was evaluated with Xia et al.'s (2003) test. Both analyses were done for the entire genome, the *rep* and the *cap* genes, as well as for the first, second and third positions of the codons for both genes separately; all these studies were performed using the DAMBE program (Xia and Xie, 2001).

Phylogenetic analyses of PCV2 sequences

Phylogenetic relationships among sequences were analyzed by parsimony and nucleotide distance methods. Firstly, the heuristic search option of PAUP 4.0.b (Swofford, 1998), considering a single stepwise addition procedure and a tree bisection-reconnection (TBR = 100) branch swapping algorithm, was used for unweighted maximum parsimony analysis (MP). A majority rule consensus tree was then generated from the 100 most parsimonious trees found in each of the 1000 bootstrap replicates of the analysis. Secondly, we computed a nucleotide distance matrix between sequences to infer phylogenies by a Neighbor-Joining (NJ) and a Maximum Likelihood (ML) trees using respectively MEGA 2.1 (Kumar et al., 2001) and TreePuzzle 5.0 (Schmidt et al., 2002). Confidence in the NJ tree was estimated by 1000 bootstrap replicates. The tree search quartet puzzling algorithm directly assigned estimations of support to each internal branch of the ML tree. Sequences were compared using the gamma Tamura–Nei model (Nei, 2000) considering different gamma parameters for the whole PCV2 genome, the *rep* and the *cap* genes. A phylogenetic tree based on the NJ method with the percentages of confidence in the ML, MP and NJ analyses along the branches is shown in Fig. 2.

Recombination among PCV2 sequences

To detect putative recombination breakpoints in the PCV2 genome and to identify sequences possibly originated from a recombination event, six methods implemented in the RDP v1.08 program (Martin and Rybicki, 2000) were evaluated (RDP, GeneConv, BootScan, MaxChi, Chimaera and SiScan). The six methods used the following general settings: window size = 20, highest acceptable P value = 0.001 and Bonferroni correction. Only putative recombination events detected by more than one method were considered. In addition, marker positions for the identified clusters were indexed (Table 1).

Selection pressure on the PCV2 genome

The existence of selective pressures along the genome was assessed first by calculating the difference between non-synonymous (dN) and synonymous (dS) rates (dN − dS) for the whole genome, the *cap* and the *rep* genes by the modified Nei–Gojobori method using MEGA 2.1 software. Furthermore, the amino acid sequence entropy calculated using BioEdit (Hall, 1998) was plotted versus the difference between dN and dS (Fig. 3). These differences were calculated with the SNAP web utility (http://hiv-web.lanl.gov/content/hiv-db/SNAP/WEBSNAP/SNAP.html).

Acknowledgments

A. Olvera and M. Cortey would like to thank Mr. August Kuentzmann Damm for technical support.

Appendix A. Accession Numbers, virus name and host of the sequences used in Fig. 1

Accession Number	Virus	Host
NC001944	Beak and feather disease virus	Psittaciformes
AF311296	Beak and feather disease virus	Psittaciformes
NC003410	Canary circovirus	*Serinus canaria*
NC002561	Columbid circovirus	*Columba fasciata*
NC008033	Starling circovirus	*Sturnus vulgaris*
NC003054	Goose circovirus	*Anser sp.*
AF536933	Goose circovirus	*Anser sp.*
NC007220	Duck circovirus	*Anas domesticus*
NC005053	Mulard duck circovirus	*Anas platyrhynchos domesticus* × *Cairina moschata domesticus*
NC006561	Muscovy duck circovirus	*Cairina moschata domesticus*
NC005148	Porcine circovirus type 2	*Sus scrofa*
AF109399	Porcine circovirus type 2	*Sus scrofa*
AF112862	Porcine circovirus type 2	*Sus scrofa*
AF109398	Porcine circovirus type 2	*Sus scrofa*
AF117753	Porcine circovirus type 2	*Sus scrofa*
NC006266	Porcine circovirus type 1	*Sus scrofa*
AY660574	Porcine circovirus type 1	*Sus scrofa*
AY184287	Porcine circovirus type 1	*Sus scrofa*
AY219836	Porcine circovirus type 1	*Sus scrofa*
AY193712	Porcine circovirus type 1	*Sus scrofa*

Appendix B. Accession Numbers, clade assigned to the sequence, country of origin of the isolated, year of the isolation and known characteristics of the pig host

Accession Number	Clade	Country	Year	Health status
DQ141322	1A	Shandong, China	2005	Unknown
DQ104422	1A	China	2005	Unknown
DQ104420*	1A	China	2005	Unknown
AY188355	1A	Zhejiang, China	2005	Unknown
AY969004	1A	Henan, China	2005	Unknown
DQ017036	1A	Guangzhou, China	2005	Unknown
AY849938	1A	Changsha, China	2005	Unknown
AY916791	1A	Guangzhou, China	2005	Unknown
AY484416	1A	Netherlands	2004	PMWS
AY484415	1A	Netherlands	2004	PMWS
AY484414	1A	Netherlands	2004	PMWS
AY484413	1A	Netherlands	2004	PMWS
AY484412	1A	Netherlands	2004	Healthy
AY484411	1A	Netherlands	2004	Healthy
AY484409	1A	Netherlands	2004	Healthy
AY484408	1A	Netherlands	2004	Healthy
AY682997*	1A	Guangzhou, China	2004	Unknown
AY682993	1A	Dongguan, China	2004	Unknown
AY682990	1A	Boluo, China	2004	Unknown
AY732494	1A	Jiangxi, China	2004	Unknown
AY691679	1A	China	2004	Unknown
AY686764	1A	Zhejiang, China	2004	Unknown
AY686762	1A	Jiangxi, China	2004	Unknown
AY651850	1A	Zhejiang, China	2004	PMWS
AY641542	1A	China	2004	PMWS
AY604430	1A	China	2004	Unknown
AY613854	1A	Guangdong, China	2004	PMWS
AY596823	1A	Guangzhou, China	2004	PMWS
AY596822*	1A	Guangzhou, China	2004	PMWS
AY578327*	1A	China	2004	PMWS

(continued on next page)

Appendix B (*continued*)

Accession Number	Clade	Country	Year	Health status
AY579893	1A	China	2004	PMWS
AY536756	1A	Zhejiang, China	2004	PMWS
AY536755	1A	Zhejiang, China	2004	PMWS
AY322003	1A	France	2004	Healthy
AY322002	1A	France	2004	Healthy
AY322001	1A	France	2004	Healthy
AY322000	1A	France	2004	PMWS
AY321999	1A	France	2004	PMWS
AY321998	1A	France	2004	Unknown
AY321997	1A	France	2004	PMWS
AY321996	1A	France	2004	Unknown
AY321995	1A	France	2004	PMWS
AY321994	1A	France	2004	Healthy
AY321993	1A	France	2004	Healthy
AY321992	1A	France	2004	Healthy
AY321991	1A	France	2004	PMWS
AY321990	1A	France	2004	PMWS
AY321989	1A	France	2004	PMWS
AY321988	1A	France	2004	PMWS
AY321987	1A	France	2004	Healthy
AY321986	1A	France	2004	PMWS
AY321985	1A	France	2004	Unknown
AY321984	1A	France	2004	PMWS
AY321983	1A	France	2004	Healthy
AY321982	1A	France	2004	Healthy
AY424405	1A	Austria	2003	Unknown
AY424404	1A	Austria	2003	Unknown
AY391729	1A	Zhejiang, China	2003	PMWS
AY256460	1A	Hungary	2003	PDNS
AY256457	1A	Hungary	2003	PMWS
AY291318*	1A	Shang Hai, China	2003	PMWS
AY291316	1A	Qing Dao, China	2003	PMWS
AY217743	1A	Zhejiang, China	2003	Unknown
AY181945	1A	Tianjin, China	2003	PMWS
AY177626	1A	Guangdong, China	2003	PMWS
AF201897	1A	Netherlands	2000	PMWS
AF538325	1A	China	2003	Unknown
AY288134	1A	China	2003	Unknown
AY288133	1A	China	2003	Unknown
AY122275	1A	China	2002	Unknown
AF201311	1A	France	2000	PMWS
AY847748*	1B	Beijing, China	2005	PMWS
AY484407	1B	Netherlands	2004	Healthy
AY682995	1B	Qingyuan, China	2004	Unknown
AY682992	1B	Conghua, China	2004	Unknown
AY691169	1B	China	2004	PMWS
AY678532	1B	Zhejiang, China	2004	PMWS
AY556475	1B	Guangxi, China	2004	PMWS
DQ151643*	1C	China	2005	Unknown
AY713470	1C	Germany	2005	Healthy
AY943819	1C	Yongzhou, China	2005	Unknown
AY484410	1C	Netherlands	2004	Healthy
AY682996	1C	Shantou, China	2004	Unknown
AY682994	1C	Guangzhou, China	2004	Unknown
AY682991	1C	Conghua, China	2004	Unknown
AY686765	1C	Jiangxi, China	2004	Unknown
AY686763	1C	Shanghai, China	2004	Unknown
AY556477	1C	Hunan, China	2004	PMWS
AY556476	1C	Hainan, China	2004	PMWS
AY556473	1C	Shandong, China	2004	PMWS
AY510375	1C	Zhejiang, China	2004	PMWS
AY291317	1C	Hu Bei, China	2003	PMWS
AY181947	1C	Shandong, China	2003	PMWS
AY181946	1C	Tianjin, China	2003	PMWS

Appendix B (*continued*)

Accession Number	Clade	Country	Year	Health status
AY035820	1C	China	2001	PMWS
AY556474	2	Fujian, China	2004	PMWS
AY146992*	2	Taiwan	2002	Unknown
AF201307	2	Germany	2000	PMWS
AB072302	2A	Japan	2001	PMWS
AF109398	2A	Canada	2000	Unknown
AF117753	2A	Canada	2000	Unknown
AY180397	2B	Taiwan	2003	Unknown
AY180396	2B	Taiwan	2003	Unknown
AY146993	2B	Taiwan	2002	Unknown
AY146991	2B	Taiwan	2002	Unknown
AF364094*	2B	Taiwan	2001	Unknown
AY256459	2C	Hungary	2003	PMWS
AY256455	2C	Hungary	2003	PDNS
AF201310	2C	Spain	2000	PMWS
AF201309	2C	Spain	2000	PMWS
AF201308	2C	Spain	2000	PMWS
AY322004	2D	France	2004	Healthy
NC005148	2D	Austria	2003	Unknown
AY424403	2D	Austria	2003	Unknown
AY424402	2D	Austria	2003	Unknown
AY424401	2D	Austria	2003	Unknown
AY256458	2D	Hungary	2003	PMWS
AY256456	2D	Hungary	2003	PMWS
AY288135	2D	China	2003	Unknown
AF381176	2D	China	2001	Unknown
AF109399	2D	Canada	2000	Unknown
AF264043	2D	USA	2000	PMWS
AF201305	2D	Germany	2000	PMWS
AF201306	2D	Germany	2000	PMWS
DQ104423	2E	China	2005	Unknown
DQ104421	2E	China	2005	Unknown
DQ104419*	2E	China	2005	Unknown
AY699793	2E	USA	2004	Unknown
AY325495	2E	South Africa	2003	PMWS
AY094619	2E	USA	2003	Unknown
AY181948	2E	Censen, China	2003	PMWS
AF465211	2E	Taiwan	2002	Unknown
AF544024	2E	Korea	2002	Unknown
AF520783	2E	South Korea	2002	PDNS
AF454546	2E	South Korea	2002	Unknown
AB072303	2E	Japan	2001	PMWS
AB072301	2E	Japan	2001	PMWS
AF408635	2E	Canada	2001	PMWS
AF381177	2E	China	2001	Unknown
AF381175	2E	China	2001	Unknown
AF112862*	2E	Canada	2000	Unknown
AF118097	2E	Canada	2000	PMWS
AF118095	2E	Canada	2000	PMWS
AF264042	2E	USA	2000	PMWS
AF264041	2E	USA	2000	PMWS
AF264040	2E	USA	2000	PMWS
AF264039	2E	Iowa, USA	2000	PMWS
AF264038	2E	Iowa, USA	2000	PMWS

* Sequences removed in the recombination study.

References

Allan, G.M., Phenix, K.V., Todd, D., McNulty, M.S., 1994. Some biological and physico-chemical properties of porcine circovirus. Zentralbl. Veterinarmed., B 41 (1), 17–26.

Allan, G.M., McNeilly, F., Cassidy, J.P., Reilly, G.A., Adair, B., Ellis, W.A., McNulty, M.S., 1995. Pathogenesis of porcine circovirus; experimental

infections of colostrum deprived piglets and examination of pig foetal material. Vet. Microbiol. 44 (1), 49–64.

Allan, G.M., Kennedy, S., McNeilly, F., Foster, J.C., Ellis, J.A., Krakowka, S.J., Meehan, B.M., Adair, B.M., 1999a. Experimental reproduction of severe wasting disease by co-infection of pigs with porcine circovirus and porcine parvovirus. J. Comp. Pathol. 121 (1), 1–11.

Allan, G.M., Mc Neilly, F., Meehan, B.M., Kennedy, S., Mackie, D.P., Ellis, J.A., Clark, E.G., Espuna, E., Saubi, N., Riera, P., Botner, A., Charreyre, C.E., 1999b. Isolation and characterisation of circoviruses from pigs with wasting syndromes in Spain, Denmark and Northern Ireland. Vet. Microbiol. 66 (2), 115–123.

Allan, G.M., McNeilly, F., Meehan, B., Kennedy, S., Johnston, D., Ellis, J., Krakowka, S., Fossum, C., Wattrang, E., Wallgren, P., 2002. Reproduction of PMWS with a 1993 Swedish isolate of PCV-2. Vet. Rec. 150 (8), 255–256.

Barton, N.H., 2000. Genetic hitchhiking. Philos. Trans. R. Soc. London, Ser. B Biol. Sci. 355 (1403), 1553–1562.

Cheung, A.K., 2003a. The essential and nonessential transcription units for viral protein synthesis and DNA replication of porcine circovirus type 2. Virology 313 (2), 452–459.

Cheung, A.K., 2003b. Transcriptional analysis of porcine circovirus type 2. Virology 305 (1), 168–180.

Cheung, A.K., 2004. Identification of the essential and non-essential transcription units for protein synthesis, DNA replication and infectious virus production of porcine circovirus type 1. Arch. Virol. 149 (5), 975–988.

de Boisseson, C., Beven, V., Bigarre, L., Thiery, R., Rose, N., Eveno, E., Madec, F., Jestin, A., 2004. Molecular characterization of porcine circovirus type 2 isolates from post-weaning multisystemic wasting syndrome-affected and non-affected pigs. J. Gen. Virol. 85 (Pt. 2), 293–304.

Foxman, B., Zhang, L., Koopman, J.S., Manning, S.D., Marrs, C.F., 2005. Choosing an appropriate bacterial typing technique for epidemiologic studies. Epidemiol. Perspect. Innov. 2, 10.

Grenfell, B.T., Pybus, O.G., Gog, J.R., Wood, J.L., Daly, J.M., Mumford, J.A., Holmes, E.C., 2004. Unifying the epidemiological and evolutionary dynamics of pathogens. Science 303 (5656), 327–332.

Grierson, S.S., King, D.P., Sandvik, T., Hicks, D., Spencer, Y., Drew, T.W., Banks, M., 2004a. Detection and genetic typing of type 2 porcine circoviruses in archived pig tissues from the UK. Arch. Virol. 149 (6), 1171–1183.

Grierson, S.S., King, D.P., Wellenberg, G.J., Banks, M., 2004b. Genome sequence analysis of 10 Dutch porcine circovirus type 2 (PCV-2) isolates from a PMWS case–control study. Res. Vet. Sci. 77 (3), 265–268.

Hall, T., 1998. BioEdit. Biological sequence alignment editor for Windows. [http://www.mbio.ncsu.edu/BioEdit/bioedit.html] North Carolina State University, North Carolina, USA.

Hamel, A.L., Lin, L.L., Nayar, G.P., 1998. Nucleotide sequence of porcine circovirus associated with postweaning multisystemic wasting syndrome in pigs. J. Virol. 72 (6), 5262–5267.

Hamel, A.L., Lin, L.L., Sachvie, C., Grudeski, E., Nayar, G.P., 2000. PCR detection and characterization of type-2 porcine circovirus. Can. J. Vet. Res. 64 (1), 44–52.

Krakowka, S., Ellis, J.A., McNeilly, F., Ringler, S., Rings, D.M., Allan, G., 2001. Activation of the immune system is the pivotal event in the production of wasting disease in pigs infected with porcine circovirus-2 (PCV-2). Vet. Pathol. 38 (1), 31–42.

Kumar, S., Tamura, K., Jakobsen, I.B., Nei, M., 2001. MEGA2: molecular evolutionary genetics analysis software. Bioinformatics 17 (12), 1244–1245.

Larochelle, R., Magar, R., D'Allaire, S., 2002. Genetic characterization and phylogenetic analysis of porcine circovirus type 2 (PCV2) strains from cases presenting various clinical conditions. Virus Res. 90 (1–2), 101–112.

Lekcharoensuk, P., Morozov, I., Paul, P.S., Thangthumniyom, N., Wajjawalku, W., Meng, X.J., 2004. Epitope mapping of the major capsid protein of type 2 porcine circovirus (PCV2) by using chimeric PCV1 and PCV2. J. Virol. 78 (15), 8135–8145.

Mahe, D., Blanchard, P., Truong, C., Arnauld, C., Le Cann, P., Cariolet, R., Madec, F., Albina, E., Jestin, A., 2000. Differential recognition of ORF2 protein from type 1 and type 2 porcine circoviruses and identification of immunorelevant epitopes. J. Gen. Virol. 81 (Pt. 7), 1815–1824.

Mankertz, A., Hillenbrand, B., 2001. Replication of porcine circovirus type 1 requires two proteins encoded by the viral *rep* gene. Virology 279 (2), 429–438.

Mankertz, A., Hillenbrand, B., 2002. Analysis of transcription of porcine circovirus type 1. J. Gen. Virol. 83 (Pt. 11), 2743–2751.

Mankertz, A., Domingo, M., Folch, J.M., LeCann, P., Jestin, A., Segales, J., Chmielewicz, B., Plana-Duran, J., Soike, D., 2000. Characterisation of PCV-2 isolates from Spain, Germany and France. Virus Res. 66 (1), 65–77.

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, D., Rybicki, E., 2000. RDP: detection of recombination amongst aligned sequences. Bioinformatics 16 (6), 562–563.

Meehan, B.M., McNeilly, F., Todd, D., Kennedy, S., Jewhurst, V.A., Ellis, J.A., Hassard, L.E., Clark, E.G., Haines, D.M., Allan, G.M., 1998. Characterization of novel circovirus DNAs associated with wasting syndromes in pigs. J. Gen. Virol. 79 (Pt. 9), 2171–2179.

Meehan, B.M., McNeilly, F., McNair, I., Walker, I., Ellis, J.A., Krakowka, S., Allan, G.M., 2001. Isolation and characterization of porcine circovirus 2 from cases of sow abortion and porcine dermatitis and nephropathy syndrome. Arch. Virol. 146 (4), 835–842.

Nei, M.a.S.K., 2000. Molecular Evolution and Phylogenetics. Oxford Univ. Press, New York.

Roitt, I.B., Male, D.J., 1998. Immunología, Trans. Diorki. 4th ed. Harcourt Brace de España, S.A., Madrid.

Rosell, C., Segales, J., Plana-Duran, J., Balasch, M., Rodriguez-Arrioja, G.M., Kennedy, S., Allan, G.M., McNeilly, F., Latimer, K.S., Domingo, M., 1999. Pathological, immunohistochemical, and in-situ hybridization studies of natural cases of postweaning multisystemic wasting syndrome (PMWS) in pigs. J. Comp. Pathol. 120 (1), 59–78.

Schmidt, H.A., Strimmer, K., Vingron, M., von Haeseler, A., 2002. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics 18 (3), 502–504.

Segales, J., Rosell, C., Domingo, M., 2004. Pathological findings associated with naturally acquired porcine circovirus type 2 associated disease. Vet. Microbiol. 98 (2), 137–149.

Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16, 1114–1116.

Swofford, D., 1998. PAUP* Phylogenetic Analysis Using Parsimony (* and Other Methods). 4. Sinauer Associates, Sunderland, MA.

Tischer, I., Mields, W., Wolff, D., Vagt, M., Griem, W., 1986. Studies on epidemiology and pathogenicity of porcine circovirus. Arch. Virol. 91 (3–4), 271–276.

Wen, L., Guo, X., Yang, H., 2005. Genotyping of porcine circovirus type 2 from a variety of clinical conditions in China. Vet. Microbiol. 110 (1–2), 141–146.

Xia, X., Xie, Z., 2001. DAMBE: software package for data analysis in molecular biology and evolution. J. Hered. 92 (4), 371–373.

Xia, X., Xie, Z., Salemi, M., Chen, L., Wang, Y., 2003. An index of substitution saturation and its application. Mol. Phylogenet. Evol. 26 (1), 1–7.