Rapid Communication

Low levels of diversity among genomes of Porcine circovirus type 1 (PCV1) points to differential adaptive selection between Porcine circoviruses

Martí Cortey a, b, Joaquim Segalés b, c, * 

a National Veterinary Institute, Ulls Väg 2B SE-751 89, Uppsala, Sweden
b Centre de Recerca en Sanitat Animal (CReSA), UAB-IRTA, Campus de la Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Catalonia, Spain

c Departament de Sanitat i Anatomia Animals, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Catalonia, Spain

ARTICLE INFO

Article history:
Received 23 August 2011
Returned to author for revision 10 October 2011
Accepted 12 October 2011
Available online 14 November 2011

Keywords:
Porcine circovirus type 1
PCV1
PCV2
Circoviridae
Evolution
Adaptive selection

ABSTRACT

Several features related with the evolutionary patterns among all the PCV1 genomes available at GenBank have been analyzed in the present work (diversity, number of genotypes, recombination, saturation, selection, evolutionary rate). The reported results point to low levels of nucleotide and amino acid diversity, low number of positively selected codons and a slow evolution rate. Compared with the other species of the Circoviridae family, the diversity is the lowest reported. This can be related with the fact that PCV1 is the single non-pathogenic member of the family. Overall, differential levels of adaptive evolution between PCV1 and PCV2 may explain the different diversity levels, and the different evolutionary and selection rates observed.

© 2011 Elsevier Inc. All rights reserved.

Porcine circoviruses (PCVs) are small, non-enveloped, single-stranded DNA viruses showing icosahedral symmetry, which are classified into the family Circoviridae, genus Circovirus. PCVs are the smallest animal viruses known, with a genome length of only 1.7–1.8 kb. Two distinct PCVs species have been defined so far: PCV type 1 (PCV1) and 2 (PCV2); both show high levels of nucleotide identity and similar genomic organizations (Olvera et al., 2007). PCVs contain two major ambisense open reading frames (ORF). ORF1 (rep gene) is transcribed in a clockwise direction, and encodes two viral replication-associated proteins Rep and Rep'. ORF2 (cap gene) is located in the complementary viral strand and follows an anti-clockwise transcription, encoding the immunogenic capsid protein. Between the ORFs there is an intergenic region that comprises the origin of replication and characterized by a putative stem-loop structure (Mankertz et al., 2004). Both PCV1 and PCV2 infections are common in pig populations all over the world (Allan and Ellis, 2000; Beach et al., 2010; Fenaux et al., 2004), despite lower prevalences for PCV1 have been reported compared with PCV2. Consequently, PCV1 has attracted much less attention of researchers compared with PCV2. However, PCV1 has recently gained notoriety because of its presence as contaminant in live human vaccines (Baylis et al., 2011; Beach et al., 2011). Therefore, in order to fill one of the gaps existing in relation to PCV1 research, the main objective of this work was to analyze the evolutive patterns and relationships among PCV1 genomes published worldwide.

All PCV1 complete genomes (n = 36) available at the GenBank (Table 1) were downloaded, aligned with ClustalW (Thompson et al., 1994), and analyzed in the present study. Partial sequences of PCV1 also existing in the GenBank were not considered due to the lack of homogeneity to perform corresponding analyses. Complete genomes, and cap and rep genes separately did not present evidences of saturation or recombination, according to the analyses performed with packages DAMBE (Xia and Xie, 2001) and RDP3 (Martin et al., 2010), respectively. Therefore, both genes and the complete genome were used for phylogenetic inference. Neighbor-Joining trees of the PCV1 genomes, and cap and rep genes separately did not present evidences of saturation or recombination, according to the analyses performed with packages DAMBE (Xia and Xie, 2001) and RDP3 (Martin et al., 2010), respectively. Therefore, both genes and the complete genome were used for phylogenetic inference. Neighbor-Joining trees of the complete genome, and the cap and the rep genes were constructed in MEGA4 (Tamura et al., 2007) based on the p-distance and resampled with 1000 bootstrap replicates. All trees exhibited very low levels of polymorphism, together with the lack of a clear clustering structure (Fig. 1). The lowest p-distance reported among the most divergent PCV1 genomes was as high as 98%. The low diversity levels...
were observed also when mutations were analyzed; 138 (7.8%) mutations in the whole genome, 84 mutations (12%) in the cap and 52 mutations (5.5%) in the rep genes. In addition, more than half (87 out of 138, 63%) of these mutations were point mutations reported in a single PCV1 sequence. Likewise, most amino acid mutations were found in the cap (58%) and the rep (74%) genes. Accordingly, the nucleotide diversity estimates for the complete genome were 0.0083 for the cap and 0.0348 for the rep genes. Interestingly, most (16 out of 38, 42%) of the positively selected codons in the cap gene were located between residues 30 and 78, which roughly coincide with one of the replication mechanisms (i.e. Beach et al., 2010; Finsterbusch et al., 2009). Finally, the rate of substitution was estimated with the BEAST software, which uses a prior distribution of the chain was sampled every 100 generations until convergence, using HKY85 + І (Hasegawa et al., 1985) as a substitution model, and the default parameters in the priors’ panel. The mean genomic substitution rate for PCV1 was estimated as 1.2·10^-1 substitutions per site per year (subs-site^-1 year^-1); Firth et al., 2009).

Table 2 Summary of diversity in the Circoviridae family.

<table>
<thead>
<tr>
<th>Genus/species</th>
<th>N</th>
<th>Length</th>
<th>Variable positions (%)</th>
<th>Nucleotide diversity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Circovirus</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beak and feather disease virus</td>
<td>87</td>
<td>1981–2019</td>
<td>956 (46.1%)</td>
<td>π = 0.0808</td>
</tr>
<tr>
<td>Canary circovirus</td>
<td>2</td>
<td>1952</td>
<td>0</td>
<td>–</td>
</tr>
<tr>
<td>Duck circovirus</td>
<td>50</td>
<td>1988–1996</td>
<td>515 (25.7%)</td>
<td>π = 0.0974</td>
</tr>
<tr>
<td>Finch circovirus</td>
<td>2</td>
<td>1962</td>
<td>0</td>
<td>–</td>
</tr>
<tr>
<td>Goose circovirus</td>
<td>25</td>
<td>1820–1821</td>
<td>325 (17.8%)</td>
<td>π = 0.0609</td>
</tr>
<tr>
<td>Gull circovirus</td>
<td>2</td>
<td>2035</td>
<td>0</td>
<td>–</td>
</tr>
<tr>
<td>Pigeon circovirus</td>
<td>12</td>
<td>2032–2040</td>
<td>486 (23.7%)</td>
<td>π = 0.0897</td>
</tr>
<tr>
<td>Porcine circovirus type 1</td>
<td>36</td>
<td>1758–1760</td>
<td>138 (7.8%)</td>
<td>π = 0.0083</td>
</tr>
<tr>
<td>Porcine circovirus type 2</td>
<td>772*</td>
<td>1767–1768</td>
<td>479 (27.1%)</td>
<td>π = 0.0348</td>
</tr>
<tr>
<td>Starling circovirus</td>
<td>2</td>
<td>2063</td>
<td>0</td>
<td>–</td>
</tr>
<tr>
<td>Swan circovirus</td>
<td>2</td>
<td>1783–1785</td>
<td>249 (13.9%)</td>
<td>π = 0.1399</td>
</tr>
<tr>
<td>Gyrovirus</td>
<td>47</td>
<td>2286–2319</td>
<td>354 (15.2%)</td>
<td>π = 0.0229</td>
</tr>
</tbody>
</table>

* 772 PCV2 complete genomes are nowadays available at the GenBank, but diversity calculations are based on the 148 full-length genomes published by Olvera et al. (2007).
prevalence and diagnostic studies (i.e. Calsamiglia et al., 2002; Puvanendiran et al., 2011), and lately on vaccine development and contamination (i.e. Baylis et al., 2011; Beach et al., 2011; Quintana et al., 2006; Tanzer et al., 2011) and presence in food of pork origin (Li et al., 2010). On the contrary, few evolutionary and phylogenetic topics have been analyzed (Cságola et al., 2008; Muhling et al., 2006).

In the present study, saturation and recombination levels for PCV1 have been evaluated, pointing to useful sequences for phylogeny. Very low levels of variability have been detected at the nucleotide and amino acid level, the lowest among the Circoviridae family. Hence, no clear clusters or groups have been identified in the phylogenetic inferences and the PASC analysis. Also, most codons are under neutral or negative selection, and the estimations of the substitution rate are low, around 10^{-5} subsite^{-1} year^{-1}.

The Circoviridae family comprises 12 species organized in two genera: Circovirus (11 species) and Gyrovirus (1 species). Infection with circoviruses is associated with immune-suppression or immune-compromisation (Faurez et al., 2009), with the single exception of PCV1, which is considered non-pathogenic. Genetic diversity is one of the most important features that allow a population evolving in an ever-changing environment with shifting selecting pressures (Sanz-Ramos et al., 2008). Higher variability may let viral populations adapt and survive in the different intra-host environments, including the selective pressures generated by the host immune response (Pfeiffer and Kirkegaard, 2005). For instance, a higher level of genetic heterogeneity within the capsid gene explains evolution to pathogenicity in mice Parvovirus (López-Bueno et al., 2008). The maintenance of advantageous replacement mutations to evade the immune system of the host is defined as adaptive selection or adaptive evolution. Codons under adaptive evolution are positively selected. In this work, very low levels of diversity and few positively selected codons have been reported for PCV1. In contrast, PCV2 shows higher levels of diversity, and about 30% of the codons are positively selected. Positively selected codons in the PCV2 capsid were concentrated in the four proposed epitopes (Olvera et al., 2007), but only in a single region for PCV1.

A growing number of substitution rate estimations for ssDNA viruses indicate that they can evolve as fast as ssRNA viruses (reviewed by Duffy et al., 2008), despite their use of host’s DNA polymerase, which is less error prone than other polymerases. Reasons for these fast rates of evolution in ssDNA viruses (Duffy and Holmes, 2008; van der Walt et al., 2008) are still unclear and challenging. The rate of evolution for PCV2 has been estimated to be as high as 1.2 · 10^{-3} subsite^{-1} year^{-1} (Firth et al., 2009). Intriguingly, PCV1 estimations are two orders of magnitude lower than PCV2, closer to dsDNA estimations. The reported results in the Circoviridae family, and specifically in the Circovirus genus, seem to relate pathogenicity with a higher genome diversity, as long as all pathogenic members of the family show at least ten times more variability than the non-pathogenic PCV1.

In summary, the non-pathogenic PCV1 has lower diversity levels, slower substitution rate and fewer positively selected codons when compared with the pathogenic PCV2. Overall, differential levels of adaptive evolution between PCV1 and PCV2 may explain the different diversity levels, and the different evolutionary and selection rates observed. Also, these differences between PCV1 and PCV2 may account for their differential pathogenicity.

**Conflict of interest**

None of the authors of this paper has a financial or personal relationship with other people or organizations that could inappropriately influence or bias the content of the paper.

**Acknowledgments**

M. Cortey would like to thank Mr. Louis Moritz Trautmann and Mr. August Kuentzmann Damm for technical support.

**References**


