Genetic variability of porcine circovirus 2 in vaccinating and non-vaccinating commercial farms

Tuija Kekarainen,1 Angel Gonzalez,2 Anna Llorens1 and Joaquim Segalés1,3

1Centre de Recerca en Sanitat Animal (CReSA), UAB-IRTA, Campus de la Universitat Autònoma de Barcelona, 08193 Bellaterra (Cerdanyola del Vallès), Spain
2Laboratori de Medicina Computacional, Unitat de Bioestadística, Facultat de Medicina, Universitat Autònoma de Barcelona, 08193 Bellaterra (Cerdanyola del Vallès), Spain
3Departament de Sanitat i Anatomia Animals, Universitat Autònoma de Barcelona, 08193 Bellaterra (Cerdanyola del Vallès), Spain

Vaccines against porcine circovirus 2 (PCV2) are now widely used to control the diseases caused by the virus. Although the vaccines protect pigs against the disease, they do not lead to sterilizing immunity and therefore infections with PCV2 continue in farms. It is expected that, due to its high evolutionary rate, PCV2 can adapt quickly to environmental pressures such as vaccination. The goal of this study was to elucidate the molecular variation of PCV2 in relation to vaccination.

PCV2 variability was investigated from samples of infected pigs from five farms where vaccination had never been applied and two farms where pigs had been vaccinated for at least 2 years. For the genetic analysis, full PCV2 genomes were amplified and subsequently pooled by vaccination status from serum of eight vaccinated, infected pigs and 16 non-vaccinated, infected pigs. Variability of viral populations was quantified using next-generation sequencing and subsequent bioinformatics analysis. The number of segregating sites was similar in the non-vaccinated (n=109) and vaccinated pools (n=96), but the distribution of these sites in the genome differed. Most notably, in the capsid gene, the number of segregating sites was observed only in the non-vaccinated population. Based on the structural analysis, it is expected that some low-frequency amino acids result in biologically low-fit viruses. On the contrary, D294 in replicase represents a novel amino acid which was dominant and unique in the vaccinated pool. This work showed that variable PCV2 populations were circulating in commercial farms, and that this variability was different in samples obtained from vaccinating and non-vaccinating farms.

INTRODUCTION

Porcine circovirus type 2 (PCV2; family Circoviridae, genus Circovirus) is one of the smallest DNA viruses (genome size 1.7 kb) infecting mammals. It has been linked to a number of diseases collectively named PCV diseases (PCVDs) (Segalés et al., 2012). PCV2 is regarded as one of the most important pathogens for domestic swine worldwide, causing significant economic losses to the pig industry. In the last 3 years, the vaccines against this virus have become the most used vaccines in pigs, displaying high levels of efficacy and excellent return on investment (Beach & Meng, 2012). Vaccinated animals can still become infected with PCV2, although with lower viraemia loads than non-vaccinated animals. Therefore, current vaccines applied as single-dose protocols do not provide sterilizing immunity (Kekarainen et al., 2010). During recent years, and in spite of vaccination, some reports about PCVD cases have been described and a novel variant called mutant PCV2b (mPCV2b) has been recovered from diseased pigs (Opriessnig et al., 2013). This has raised concerns about possible PCV2 vaccine escape strains, although it has been shown that a current vaccine is effective against mPCV2b challenge under experimental conditions (Opriessnig et al., 2014).

Based on genetic variation in its capsid gene, PCV2 is divided into three major groups designated PCV2 genotypes a, b and c (Segalés et al., 2008). The genetic distance between genotypes is at least 3.5%, but diversities between isolates can reach up to 10% (Cortey et al., 2011a; Grau-Roma et al., 2008). Evolutionary rates for PCV2 are higher than expected for DNA viruses (1.2 x 10⁻³ substitutions per site per year), resembling the rates from RNA viruses (Firth et al., 2009).
Although the molecular variation of PCV2 is known, there are no in-depth molecular analyses on the PCV2 variants circulating in commercial farms. Such knowledge should be of great interest especially in scenarios where non-sterilizing vaccination is applied and viral evolution under vaccination pressure may occur. Therefore, the goal of this study was to characterize the variability of PCV2 in farms where vaccination had never been applied and in farms where pigs had been vaccinated against PCV2 for at least 2 years. For this, next-generation sequencing (NGS) was applied on full-genome PCV2 sequences amplified from pools of infected pigs from vaccinating and non-vaccinating commercial farms located in north-eastern Spain. The results obtained showed that populations of PCV2 variants were circulating in the analysed farms, and that this variability could differ in samples obtained from vaccinating and non-vaccinating farms.

RESULTS

Low level of PCV2 infection in vaccinating farms

Standard PCR was applied to detect PCV2 in non-vaccinated (n=8) and long-term vaccinated (n=8) farms in animals between 15 and 18 weeks of age, which normally have the highest infection rates (Grau-Roma et al., 2009). At this stage, it was seen that only two vaccinating farms had detectable levels of PCV2 by PCR, whilst the virus was detected in six non-vaccinating farms. The low infection rate was reflected in overall prevalence, which in vaccinating farms was 6% (8/129) and in non-vaccinating farms was 68% (81/119). Infection due to inactivated viral vaccine was excluded as these farms were using the same vaccine, which was based on the capsid protein of PCV2. Next, for NGS, the full-length PCV2 genome was amplified from serum samples of all infected animals from two vaccinating farms and a total of 16 animals from five non-vaccinating farms. Animals from non-vaccinating farms were selected based on similar levels (mean±SD) of PCV2 DNA (log_{10}7.1±1.3 copies ml⁻¹) and anti-PCV2 antibodies (OD_{450} 1.1±0.8) in serum as in non-vaccinated animals (log_{10}7.1±0.7 copies ml⁻¹ and OD_{450} 1.5±0.8). After quantification of the full-genome amplicons, they were pooled in equimolar ratios according to origin (vaccinating versus non-vaccinating farms), resulting in the ‘vaccinated pool’ (samples from animals in vaccinating farms) and ‘non-vaccinated pool’ (samples from animals in non-vaccinating farms).

Sequencing results

To determine the molecular variability in vaccinating and non-vaccinating farms, vaccinated and non-vaccinated pools were sequenced by Ion Torrent technology, and filtered by quality and length before analysis. In total, 27,438 reads were analysed from the non-vaccinated pool and 35,037 reads were analysed from the vaccinated pool, with median lengths of 172 and 165 bp, respectively. To generate the consensus sequence for each pool, reads were first mapped to the PCV2 reference sequence (GenBank accession number GU049341) followed by replacement of any base with the consensus base. Finally, reads were remapped to the consensus sequence and pairwise aligned. The mean numbers of bases aligned per site were 2500 (range 137–6615) and 2950 (range 235–6408) in non-vaccinated and vaccinated pools, respectively (Fig. S1, available in the online Supplementary Material).

PCV2 population diversity in non-vaccinating and vaccinating farms

Polymorphism analysis was performed in the whole genome, ORF1 (replicase) and ORF2 (capsid) genes for non-vaccinated and vaccinated pools reporting sites with frequencies >5%. As summarized in Table 1, the analysis showed that viral populations were heterogeneous in both pools. In the whole genome and the capsid-encoding region, higher numbers of segregating sites and mutations were identified in the non-vaccinated pool than in the vaccinated pool. In the replicase region, the numbers of segregating sites and mutations were comparable between the studied pools, except when the types of mutations were taken into account: synonymous changes were more frequent in the non-vaccinated pool, whilst non-synonymous changes were increased fourfold in the vaccinated pool. Indeed, the percentage of non-synonymous changes from the total amount of mutations in the replicase gene was 7% (2/27) in the non-vaccinated pool and 29% (8/28) in the vaccinated pool. In the capsid gene, non-synonymous changes were 43% (36/83) and 47% (31/66) in the non-vaccinated and vaccinated pools, respectively. However, the numbers of segregating sites or mutations were not statistically different between the non-vaccinated and vaccinated pools.

As the numbers of pigs included in the pools differed, we next analysed whether this could have affected our analysis. According to the Watterson estimate (θ), the mutation rate was similar (θ=0.012) in both analysed pools. Based on this, the different samples numbers did not affect the detection of segregating sites and mutations.

Next, the sites that segregated in only one of the pools were identified. Altogether, there were 32 variable sites that were unique in one or other pool (Fig. 1), from which most (n=23) were detected in the non-vaccinated animals. Interestingly, in the vaccinated pool, only two unique segregating sites were detected in the capsid region, whilst 14 sites were variable in the non-vaccinated pool. Five out of seven sites in the replicase gene in the vaccinated pool resulted in amino acid changes, but none in the non-vaccinated pool.

Different PCV2 variants were detected in non-vaccinating and vaccinating farms

To identify the major type of circulating viruses in non-vaccinating and vaccinating farms, the main and the
variant full-length PCV2 sequences were generated. This was done based on estimated codon frequencies as explained in Methods. According to the phylogenetic analysis of the capsid gene, the main genotype variants circulating in the non-vaccinating farms was PCV2b (Fig. 2a), whilst the dominating variant in the vaccinating farms was PCV2a. The minor variants of each pool grouped together with the main variant of the other pool when the capsid gene was analysed, indicating that both genotypes were present in both farm types. This was not the case using the replicase gene, where the main variants did group together and apart from the minor variants (Fig. 2b).

Circulation of both PCV2a and b genotypes in the analysed farms was confirmed by analysis of non-synonymous amino acid changes in the viral capsid protein (Fig. 3). The non-synonymous changes in the non-vaccinated and vaccinated pools were mainly in the same location, but differed in their frequencies. Sites with up to three amino acid variants were identified at positions 59, 63, 131 and 190, and 59 and 86 in non-vaccinated and vaccinated pools, respectively. All these positions are located in the heterogenic region overlapping previously assigned antigenic regions of the capsid protein (Lekcharoensuk et al., 2004; Mahé et al., 2000; Shang et al., 2009). In these positions, the less frequent variants were typically found in ~16% of sequences. Three amino acid (28L, 141T and 228E) changes were found that, to the best of our knowledge, have not been reported in PCV2 previously. These variants were found at frequencies of only ~5% in the vaccinated pool (28L and 228E) and non-vaccinated pool (141T), and did not locate in known regions containing conformational or linear epitopes.

In the replicase protein, six amino acids were highly conserved in the non-vaccinated pool, but represented minor variants in the vaccinated pool (Fig. 4). When compared with published PCV2 replicase sequences reported in GenBank, it was noticed that 34H, 153R, 222R and 294D in replicase were novel amino acids found in this study. Interestingly, all were found in the population infecting vaccinated animals and one of them (294D) was the main amino acid in its position, present in 71% of the sequences. Amino acids 153 and 222 are located in the helicase domain of replicase (Gorbalenya et al., 1989).

### Table 1. Numbers of segregating sites and mutations in pools of non-vaccinating and vaccinating farms

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Vaccination status</th>
<th>Genome (nt 1–1795)</th>
<th>Replicase (nt 51–992)</th>
<th>Capsid (nt 1734–1036)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. segregating sites</td>
<td>No</td>
<td>109</td>
<td>27</td>
<td>79</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>96</td>
<td>26</td>
<td>65</td>
</tr>
<tr>
<td>No. different mutations</td>
<td>No</td>
<td>114</td>
<td>27</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>98</td>
<td>28</td>
<td>66</td>
</tr>
<tr>
<td>Synonymous changes</td>
<td>No</td>
<td>72</td>
<td>25</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>55</td>
<td>20</td>
<td>35</td>
</tr>
<tr>
<td>Non-synonymous changes</td>
<td>No</td>
<td>38</td>
<td>2</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>39</td>
<td>8</td>
<td>31</td>
</tr>
</tbody>
</table>

**Fig. 1.** Unique polymorphic sites in non-vaccinated and vaccinated pools. The position in the PCV2 genome is indicated at the left; an asterisk following the number indicates sites with non-synonymous changes. Frequency of polymorphism is shown at the top of the graph.

Three-dimensional (3D) mapping of amino acid variants in the PCV2 virus-like particle (VLP)

Selected amino acid variants detected in non-vaccinated and vaccinated pools were mapped onto the X-ray 3D structure and cryo-electron microscopy (cryo-EM) reconstruction of the PCV2 VLP (Fig. 5) (Khayat et al., 2011). As Fig. 5 shows, many of the residues derived from frequencies of non-synonymous amino acid changes in the PCV2 capsid protein are exposed to the solvent and belong to
epitope regions identified previously by experimental methods.

Residues 59, 63 and 190 are located on the surface of the particle, and display a significant variability between PCV2 viral species (Fig. 3). These amino acids are also included in regions important for epitope recognition of mAbs (Lekcharoensuk et al., 2004). Amino acids 86 and 131, which display high variability, have also been identified as part of antigenic regions (Mahé et al., 2000; Meng, 2013), whilst the C-terminal residue 230 has been included in several studies as part of a conformational epitope (Lekcharoensuk et al., 2004; Meng, 2013; Shang et al., 2009). These amino acids, together with 133, 136 and 228, are located in close proximity to each other in the subunit interface at the icosahedral threefold axes (Fig. 5a), suggesting their possible implication in a conformational epitope.

Finally, according to the structural analysis, some amino acid changes are predicted to produce a defective VLP. This is particularly the case for the low-frequency substitution P110R (Fig. 3). P110 is located in the loop region, forming a negative electrostatic potential surface at the icosahedral fivesfold axes interface (Fig. 5b). Thus, the substitution by the voluminous positively charged arginine residue would interfere with capsomer association, disrupting capsid assembly.

**DISCUSSION**

Vaccination against PCV2 has been widely used in pig farms during the last 7 years. Current vaccines protect animals from disease, but do not result in sterilizing immunity. PCV2 is still circulating in farms despite vaccination and therefore its evolution could be shaped by vaccination. In this study, a detailed molecular characterization of PCV2 was performed in farms that had never applied PCV2 vaccination and in farms that had been vaccinating for at least 2 years.

Traditionally, Sanger sequencing has been applied for molecular studies of PCV2. It has been shown that evolutionary rates for PCV2 are higher than expected for a DNA virus ($1.2 \times 10^{-3}$ substitutions per site per year), resembling the rates from RNA viruses (Firth et al., 2009). Co-infections with PCV2 genotypes have been described in nature (Cai et al., 2012; Khaiseb et al., 2011; Zhai et al., 2011), and infection by PCV2 may comprise the main sequence and several different variants closely related to it (Grau-Roma et al., 2008). By applying NGS on PCV2, we were able to confirm that, at any given moment, several different variants are circulating in farms. With this technique, low-frequency variants were also identified, which would not have been efficient with traditional techniques that require cloning of the PCR product and sequencing of individual clones. It is expected that some low-frequency amino acids result in biologically low-fit viruses. Indeed, one such amino acid change, P110R, was found in the capsid that, based on the structural analysis, would result in defective virus particle formation.

A shift in genotype prevalence from PCV2a to PCV2b has been described to coincide with severe outbreaks of PCV2 systemic disease, although there are insufficient data to support higher pathogenicity for PCV2b (Carman et al., 2006; Cheung et al., 2007; Cortey et al., 2011b; Dupont et al., 2009; Opriessnig et al., 2008). Currently, the PCV2b genotype infects pigs predominantly worldwide (Cortey et al., 2011b; Dupont et al., 2008; Kim et al., 2011; Shen et al., 2009; Opriessnig et al., 2008).
et al., 2012). PCV2 genotyping is based on the capsid nucleotide sequences, and according to phylogenetic analysis both genotypes were found in the pools of vaccinated and non-vaccinated animals analysed in this study. Phylogenetic analysis of the capsid gene revealed that the dominant PCV2 genotype in the vaccinated pool belonged to PCV2a, whilst in the non-vaccinated pool it belonged to PCV2b. This is especially interesting as all current vaccines are based on the PCV2a genotype, although these vaccines have been shown to induce cross-protective immunity (Fort et al., 2008; Opriessnig et al., 2009).

In a recent study, it was shown that in a farm suffering from apparent vaccine failure, 90 % of pigs were positive for PCV2a, 87 % were positive for PCV2b and 77 % were co-infected with both genotypes by PCR (Gerber et al., 2013). Interestingly, by sequencing PCR products from three animals, the authors identified five amino acid changes that had occurred in PCV2a. Using NGS of the vaccinated pool, we identified the same positions as variable sites and in three of these, the main amino acids (59A, 133V and 136Q) coincided. Interestingly, these residues are located in solvent-exposed regions on the PCV2 VLP, suggesting their possible implication in

---

**Fig. 3.** Frequencies of non-synonymous amino acid changes in the capsid protein detected in non-vaccinated (NV, black bars) and vaccinated (V, grey bars) pools. Below the graph, amino acid variants and their positions according to the numbering of the protein sequence of GenBank accession number GU049341. Novel amino acids found in this study are indicated by asterisks.

**Fig. 4.** Frequencies of non-synonymous amino acid changes in the replicase protein detected in non-vaccinated (NV, black bars) and vaccinated (V, grey bars) pools. Below the graph, amino acid variants and their positions according to the numbering of the protein sequence of GenBank accession number GU049341. Novel amino acids found in this study are indicated by asterisks.
conformational epitopes. It has to be kept in mind that in our study, pools generated by mixing equimolar ratios of PCR products were sequenced by NGS, resulting in short fragments that were assembled using bioinformatics tools to full-length PCV2 genomes. Such an approach is useful to analyse circulating viral populations en masse, but does not provide information on the animal or farm level. However, co-infection with different genotypes and identification of minor variants is feasible.

Altogether, seven novel amino acid changes were identified; 294D was especially interesting as it was found more frequently in the replicase sequence than 294E, which is typically found in PCV2 replicase. Indeed, replacement of E

Fig. 5. Structure of the PCV2 virus-like particle and capsid protein association (Protein Data Bank ID: 3R0R). (a) Surface representation of antigenic domains recognized by anti-PCV2 antibodies and amino acids with non-synonymous changes detected in our study (red). (b) Electrostatic potential map of the PCV2 cryo-EM structure calculated using the program APBS with the non-linear Poisson–Boltzmann equation and contoured at \( \pm 10kT/e \) (negatively and positively charged surface areas in red and blue, respectively). On the right, ribbon representations of capsid proteins contained in the dashed sections. Selected amino acids are shown in the solvent surface with their corresponding sequence number. Projections of the icosahedral viral particle are shown down: (a) threefold and (b) fivefold axes.
by D in nature could likely be based on scoring by the point accepted mutation matrix (PAM125). However, the biological significance of this variant remains unknown as it is not located in any known functional domain.

This study showed that PCV2a and PCV2b genotypes are circulating in both non-vaccinating and vaccinating farms. The PCV2 population in the farms examined showed a high degree of variability. In the capsid gene, in particular, the vaccinated pool lacked several segregating sites that were observed in the non-vaccinated pool. Novel amino acid mutations were found in the capsid. Due to its high proportion in the vaccinated pool, an especially interesting amino acid mutation was 294D in replicate. Determining the biological importance of novel amino acid changes will be a subject for future experiments.

It is still unknown how vaccination may affect PCV2 evolution. This study resulted in detailed genetic characterization of the virus in the field at one time point in the presence and absence of vaccination, but similar analyses under more controlled and elaborate settings, such as experimental infections of non-vaccinated and vaccinated animals, are under way.

METHODS

Farm and animal selection. Conventional pig farms were selected based on their history of PCV2 vaccination: farms that had never vaccinated against PCV2 (non-vaccinated, n=8) and farms that had been vaccinating animals against PCV2 for at least 2 years (vaccinated, n=8). Vaccination in the farms was done according to the manufacturer’s recommendations with currently registered PCV2 vaccines. Two of the farms were vaccinated with Porcilis PCV (MSD Animal Health), three with Ingelvac Circoflex (Boehringer Ingelheim) and no information about the vaccine used was available from three farms. Two of the farms were vaccinated with Porcilis PCV (MSD Animal Health), three with Ingelvac Circoflex (Boehringer Ingelheim) and no information about the vaccine used was available from three farms. All farms were commercial and located in north-eastern Spain, which is one of the main pig-producing regions in the country. At the time of sampling, all farms were considered healthy, with no evidence of pigs suffering from clinical signs of PCV2 systemic disease. Between 15 and 20 serum samples were collected from each farm from healthy pigs aged 15–18 weeks. Age selection was based on the usual highest prevalence of PCV2 infection in this age range (Grau-Roma et al., 2002). Quantification of PCV2 DNA in serum was done as described (Olvera et al., 2004) and anti-PCV2 antibodies were measured by a commercial Ingezim Circo IgG ELISA kit as recommended (Inmunologia y Genetica Aplicada).

DNA extraction and PCV2 PCRs. DNA was extracted from serum samples and PCV2 infection was assessed by in-house PCR as described previously (Quintana et al., 2002). After standard PCR had been performed, five non-vaccinating and two vaccinating farms were under way.

DNA sequencing and bioinformatics analysis. Deep sequencing was performed using the Ion Torrent platform (Life Technologies) in the Centre for Research in Agricultural Genomics (Spain), according to the manufacturer’s specifications. Briefly, 100 ng purified pools of DNA were used for library construction with the Ion Xpress Plus Fragment Library kit, Ion Xpress Barcode adapters and emulsion PCR performed with the Ion PGM Template OT2 200 kit (Life Technologies). The quality of the obtained library was evaluated with the Bioanalyzer instrument with an Agilent High Sensitivity DNA kit. Sequencing was run on the Ion Torrent Personal Genome Machine (Life Technologies) loaded with a 314 chip as recommended with the Ion PGM Sequencing 200 kit (Life Technologies).

Sequences were filtered by quality (cut-off 20 per base quality score) and length (minimum length 30 bp) using FastQ tools implemented in Galaxy (Blankenberg et al., 2010). Filtered sequences were mapped and aligned to the reference genome PCV2b (GenBank accession number GU049341) using Segminator II (Archer et al., 2012a, b) – a tool specifically developed for characterization of viral NGS data. Probabilistic read error detection was utilized together with the known platform error rate to determine the probability of an observed nucleotide being biologically relevant and not due to the platform error. Sites were considered polymorphic if the alternative base frequency was >5% and if the probability was >0.97. The per-site coverage, codon frequencies and consensus sequence were generated by Segminator II. The consensus genomic sequence was obtained for each pool by replacing any nucleotide of the reference sequence GU049341 with the most common nucleotide in the reads. Reads were then remapped using the generated consensus sequence of each pool. Numbering of the amino acid positions and nucleotide positions throughout the text refers to numbering of a Spanish PCV2b genotype (GenBank accession number GU049341).

Full-length viral (minor) variants were generated based on the codon frequencies obtained by Segminator II. Briefly, each variable position was checked manually and the codons with the highest frequencies were first used to construct the main PCV2 type in the analysed pool. Then, variant PCV2 types were constructed containing codons with frequencies >5% (cut-off, as explained above) and below the most frequent codon. These main and variant PCV2 types were used in the phylogenetic analysis. The Jukes–Cantor p-distance model was used to reconstruct the phylogenetic tree with 500 resamples (Tamura et al., 2011). Polymorphic sites and mutation rates were calculated using DnaSp (Librado & Rozas, 2009).

The structure of the PCV2 VLP and capsid protein has been published (Protein Data Bank ID: 3R0R) (Khayat et al., 2011) and was used to map the amino acid changes. An electrostatic potential map of the PCV2 cryo-EM structure was calculated using the program APBS with the non-linear Poisson–Boltzmann equation and contoured at ± 10kT/e. Structural figures were prepared with the PyMOL molecular graphics system.

The χ2 test was used to compare the number of segregating sites and mutations between non-vaccinated and vaccinated pools. The Watterson estimate (θ) was used to estimate the mutation rates between the non-vaccinated and vaccinated pools using the number of synonymous changes. Synonymous changes were used based on the assumption that they present neutral mutations.

ACKNOWLEDGEMENTS

This work was supported by the Boehringer Ingelheim European Animal Health PCV2 award. T. K. is supported by the Ramón y Cajal...
Programme. Special thanks go to Professor Esteban Domingo (Centro de Biologia Molecular 'Severo Ochoa', Madrid, Spain) and Antonio Barbadilla (Institut de Biotecnologia i de Biomedicina and Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, Barcelona, Spain).

REFERENCES


