Veterinary Immunology and Immunopathology 223 (2020) 110034

Contents lists available at ScienceDirect

Veterinary Immunology and Immunopathology

journal homepage: www.elsevier.com/locate/vetimm

T cell epitope content comparison (EpiCC) analysis demonstrates a bivalent PCV2 vaccine has greater T cell epitope overlap with field strains than monovalent PCV2 vaccines

Meggan Bandrick[a], Andres H. Gutiérrez[b], Prerak Desai[a,1], Gonzalo Rincon[a], William D. Martin[b], Frances E. Terry[b], Anne S. De Groot[b,c], Dennis L. Foss[a,*]

[a] Zoetis, Inc., Kalamazoo, MI, United States
[b] EpiVax, Inc., Providence, RI, United States
[c] Center for Vaccines and Immunology, University of Georgia, Athens, GA, United States

ARTICLE INFO

Keywords:
PCV2
Vaccine
Cell-mediated immunity
T cell epitope content comparison

ABSTRACT

Porcine circovirus type 2 (PCV2) has one of the highest evolutionary rates among DNA viruses. Traditionally, PCV2 vaccines have been based on the 2a genotype as this was the first genotype discovered. Today, eight genotypes of PCV2 viruses have been identified, and, taken together with the rapid evolutionary rate, propensity to recombine, and high rate of vaccination, further variation in PCV2 is expected. For these reasons, there is a growing genetic gap between available vaccines and field strains. When selecting vaccines, it is important to consider vaccines that contain T cell epitopes that are well-matched to the circulating strains. To quantify the relatedness between PCV2 vaccines and field strains, we predicted and compared their T cell epitope content and calculated Epitope Content Comparison (EpiCC) scores using established in silico tools. T cell epitopes predicted to bind common class I and class II swine leukocyte antigen (SLA) alleles were identified from two major structural proteins, the capsid (encoded by ORF2) and the replicase (encoded by ORF1). The T cell epitope content of three commercial PCV2a-based vaccines (a baculovirus expressed PCV2a ORF2 [VacAlt], a PCV1-PCV2a chimeric virus vaccine [VacA] and a combination cPCV2a-cPCV2b chimeric virus vaccine [VacAB]) and an experimental PCV2b ORF2-based chimeric virus vaccine [VacB] (Table 1), were compared to that of 161 PCV2 field strains (representing genotypes a-f).

The T cell epitope content and conservation between vaccine and field strains varied. While all vaccine strains provided broad coverage of the field strains including heterologous genotypes, none of the vaccines covered all the putative T cell epitopes identified in the field strains. PCV2a-based vaccine strains generally scored higher in terms of conserved epitope content against PCV2a field isolates but were not identical. The PCV2b-based vaccine strain had higher scores against PCV2b and PCV2d field strains. The combination PCV2a-PCV2b vaccine (VacAB) had, on average, the highest EpiCC score. PCV2 continues to evolve and EpiCC analysis provides a new tool to assess the possible impact of virus genetic divergence on T cell epitope coverage of vaccine strains. Given that multiple genotypes are currently found and may co-exist on farms, this analysis suggests that a combination of PCV2a and PCV2b vaccine strains may be required to provide optimal coverage of current and future field isolates.

1. Introduction

Porcine Circovirus (PCV) is a non enveloped, single-stranded DNA virus. Three types of PCV have been identified: PCV1, PCV2, and PCV3. PCV1 was originally identified as a cell culture contaminant and is not pathogenic in pigs. PCV3 has recently been tied to cases of reproductive failure in swine but its role as a primary pathogen remains to be understood (Palinski et al., 2017). PCV2 has been present in pigs since 1969 and first reported to be associated with porcine multisystemic wasting syndrome (PMWS) in the 1990s. PCV2-associated disease (PCVAD), or PCV-disease (PCVD), has become one of the most economically important emerging diseases worldwide affecting pigs

* Corresponding author at: Zoetis, 333 Portage Street, Kalamazoo, MI, 49007, United States.
E-mail address: dennis.l.foss@zoetis.com (D.L. Foss).
[1] Present address: Immunology Translational Sciences, Janssen R&D, Spring House, PA, United States.

https://doi.org/10.1016/j.vetimm.2020.110034
Received 2 December 2019; Received in revised form 17 February 2020; Accepted 3 March 2020

Table 1
Analyzed vaccines.

Category	Type	Genotype	Name	Description	Analyzed ORFs	Label	GenBank Reference
Commercial	Monovalent	PCV2a	be-PCV2a	Baculovirus expressed PCV2a ORF2	ORF2	ORF2 VacAlt	Not available
Commercial	Monovalent	PCV2a	cPCV2a	PCV1-PCV2a chimeric whole virus vaccine	ORF1 ORF2	ORF1 PCV1 ORF2 VacA	AU699795 AF264042
Experimental	Monovalent	PCV2b	cPCV2b	PCV1-PCV2b chimeric whole virus vaccine	ORF2	ORF2 VacB	GU799576
Commercial	Bivalent	PCV2a and PCV2b	cPCV2a-cPCV2b	Combination of cPCV2a and cPCV2b	ORF2	ORF2 VacAB	AF264042 GU799576

primarily between 5–18 weeks of age with mortality up to 30 % (Segalés, 2012).

PCV viruses have similar genomic structures and encode two major proteins, the capsid (encoded by ORF2) and the replicase (encoded by ORF1). Today, six PCV2 genotypes have been described PCV2a-f (Bao et al., 2018) and another two, including potential recombinant forms have been proposed (PCV2g-h) (Franzo and Segalés, 2018). Genotypes are delineated by a p-distance of 0.035 (Segalés et al., 2008) and proposed to have a maximum intra-genotype p-distance of 13 % (calculated on the ORF2 gene) (Franzo and Segalés, 2018). PCV2 genotypes can be further organized by cluster (Olvera et al., 2007). A striking feature of PCV2 is its high rate of evolution. Despite being a DNA virus, PCV2 has an evolutionary rate similar to RNA viruses (Firth et al., 2009). In addition to the high rate of point mutations, PCV2 viruses recombine (Hesse et al., 2008; Lefebvre et al., 2009). In fact, approximately 20–35 % of PCV2 viruses are recombinants (Franzo et al., 2016; Franzo and Segalés, 2018). Co-infection of PCV2 viruses is common (Gerber et al., 2013) and allows multiple PCV2 viruses to interact within a single cell and share genetic material.

Since the identification of PCV2a, two major changes in prevalence of circulating strains have occurred, known as “genotype shifts.” PCV2a was displaced as the predominant circulating genotype by PCV2b and PCV2b was later displaced by PCV2d, the current predominant genotype in many geographies (Xiao et al., 2016, 2015; Yang et al., 2017). PCV2a, PCV2b, and PCV2d are considered clinically relevant. PCV2c, e, and f are found at lower prevalence and their clinical significance is unknown. Taken together, due to its inherent nature, there is much diversity within PCV2 viruses.

PCV2 virus evolution has also, at least in part, been shaped by PCV2 vaccination and resultant immune pressure (Franzo et al., 2016; Reiner et al., 2015). The first PCV2 genotype to be identified was PCV2a; subsequently PCV2 vaccines were based on PCV2a. PCV2a-based vaccines have ultimately added to decreased prevalence of PCV2a but have contributed to increased diversity within the PCV2a genotype group, especially focused at PCV2a epitopes shared by vaccine and field strains (Franzo et al., 2016). Despite PCV2a no longer being the predominant circulating genotype, currently available PCV2a-based vaccines have historically been successful in controlling PCV2 disease. PCV2a-based vaccines may be adequate in conferring cross-protection to divergent genotypes under certain circumstances, but homologous vaccine/challenge combinations are better at decreasing viremia (Karuppannan and Opriessnig, 2017).

PCV2 vaccine-induced protection to field challenge strains can be partially attributed to the epitopic determinants common or unique among vaccine and field PCV2 viruses. Minor variations within ORF2, the immune-dominant target, have been shown to result in differences in immune recognition (Constans et al., 2015; Kurtz et al., 2014; Saha et al., 2012). Having a high degree of similarity among epitopes between vaccine and field viruses is essential for conferring broad immune recognition. Given the high level of diversity within PCV2 field strains and the shift away from PCV2a predominance, traditional PCV2a-based vaccines may not provide optimal epitope coverage of current and future field isolates. When selecting vaccine candidates, it is important to consider not only T cell epitope content and density, but also the potential to induce memory T cells that will recognize epitopes contained within circulating strains. In other words, the vaccine must be immunogenic (able to induce an immune response), but also contain T cell epitopes that are well-matched to the circulating strains that may be encountered by the vaccinated animal.

In this study, we used PigMatrix (Gutiérrez et al., 2016), a computer-based predictive tool, to identify putative class I and II SLA epitopes of PCV2 proteins encoded by ORF1 and ORF2. We then used the EpiCC algorithm to compare the T cell epitope content of vaccine sequences and field strains (Gutiérrez et al., 2017). We analyzed sequences of three commercial PCV2a-based vaccines (a baculovirus expressed PCV2a ORF2 [VacAlt], a PCV1-PCV2a chimeric virus vaccine [VacA] and a combination cPCV2a-cPCV2b chimeric virus vaccine [VacAB]) and an experimental PCV2b ORF2-based chimeric virus vaccine [VacB] (Table 1). EpiCC scores were used to identify the vaccine strain sequence that best represents the T cell epitope content of the input set of circulating strains and that may induce the broadest cross-reactive T cell response. In a previous retrospective study, we used PigMatrix and EpiCC to compare swine influenza A vaccines to circulating strains defining a threshold EpiCC score associated with vaccine efficacy (Gutiérrez et al., 2017). Here, EpiCC was used in a prospective manner to examine vaccine and circulating strains, thus establishing its usefulness in the development of new vaccines.

2. Methods

2.1. Phylogenetic and population analysis of PCV2 sequences

A total of 4500 publicly available (as of June 2017) PCV2, ORF2 nucleotide sequences were downloaded from GenBank (Table S1). The nucleotide sequences were aligned using MAFFT (Katoh and Standley, 2013) and a maximum likelihood tree was constructed using FastTree (Price et al., 2010). BAPS (Bayesian analysis of Population Structure) (Cheng et al., 2013) was used to cluster the strains and perform admixture analysis to identify field strains with possible recombination events. The *de novo* identified sequence clusters were then mapped to known sequence clusters as defined by Xiao et al. (Xiao et al., 2015). Within each cluster identified using BAPS, PCA (principal component analysis) for the DNA distance matrix was used to identify the minimum number of PCs (principal components, K) that explain 95 % of the variance in the distance matrix. The cluster was further partitioned into k sub-clusters using K means clustering. A strain nearest to the sub-cluster medoid was selected as a representative for the sub-cluster. Using this iterative process, based on the clustering results, 161 representative strains (Tables S1 and S2) spanning the known taxonomic diversity of PCV2 ORF2 sequences were subjected to a further T cell epitope content comparison (EpiCC) analysis. Strains representing recombinant forms were included within the analysis and across cluster (except for cluster 11, PCV2e, as recombinant viruses were not found within the cluster).

2.2. Development of MHC binding prediction matrices for swine: PigMatrix

As a first step in the comparison of T cell epitope content found in commercial and experimental PCV2 vaccines vs content found in field strains of PCV2, it was necessary to develop tools for predicting T cell

epitopes presented by swine leukocyte antigen (SLA). While many SLA alleles may be present in US swine populations, the prevalence and distribution of these SLA alleles is unknown. Therefore, for this analysis, we used SLA alleles that were present in a cohort of outbred pigs that had been SLA typed in a previous study (Gutiérrez et al., 2016).

The SLA predictive tools used for this analysis were developed using the pocket profile method (Gutiérrez et al., 2015; Sturniolo et al., 1999). By comparing the SLA binding pocket profiles of selected alleles to the HLA binding pocket profiles of well-documented human leukocyte antigen (HLA), it is possible to construct first order predictive matrices for SLA. Briefly, the protein sequences of selected SLA were aligned to HLA. Where SLA and HLA alleles were well-aligned (identity > 80 %), "contact" residues assumed to comprise the pockets of the binding groove were extracted. Contact residues were then group by pocket to create pocket profiles. SLA-derived profiles were then compared to HLA pocket profiles. Where profiles were well-matched (identity > 70 %), human pocket binding preferences were copied and rendered as a vector of 20 numeric binding coefficients (one for each of 20 natural amino acids) with higher values indicating higher binding potential. A set of nine vectors, one for each binding position, defines a matrix. These SLA-binding "PigMatrix coefficient" matrices were loaded into EpiVax's iVAX system.

2.3. MHC binding prediction

Using EpiVax's iVAX system and the PigMatrix coefficient matrices described above, all input protein sequences were parsed into overlapping 9-mer frames and each frame was scored for its binding potential to matrices in a panel of eight SLA class I alleles (SLA-1*0801, 1*1201, 1*1301, 2*0501, 2*1201, 3*0501, 3*0601, and 3*0701) and five SLA class II alleles (SLA-DRB1*0201, 0402, 0602, 0701, and 1001). Raw scores were then normalized against the scores of a large set of randomly compiled peptides. The normalization step renders results as Z-scores that are directly comparable across alleles. Peptide 9-mers with Z-scores above 1.64 (approximately the top 5 % of any given peptide set) are considered potential SLA binders.

2.4. T cell epitope content comparison (EpiCC) analysis

The EpiCC algorithm, contained within the iVAX web site, compares two protein sequences and renders results as an EpiCC score (Fig. 1). The two input sequences are the target sequence and the comparison sequence. The EpiCC score quantifies the relatedness of the putative epitope content shared between a given pair of sequences. More similar shared epitope content between the vaccine and field strain sequences results in a greater EpiCC score. The EpiCC scores can be depicted on a radar plot, which enables the scores of multiple vaccines to be compared to each other and a threshold for putative vaccine efficacy to be set if vaccine efficacy data are available (Gutiérrez et al., 2017). In this case, the EpiCC algorithm was applied iteratively to a set of vaccine (target) sequences comparing each to list of field strain (comparison) sequences and returning an EpiCC score for each comparison.

Sequences of several commercial and experimental vaccines were analyzed for comparison to the field strain epitopes (Table 1). For ORF1, we compared one vaccine (ORF1 sequence derived from the PCV1 in the PCV1-PCV2a chimeric virus vaccine, cPCV2a; ORF1 PCV1 (Fenaux et al., 2004)), to 91 field strain sequences (1 partial and 90 full-length sequences). This PCV1, ORF1 sequence is also contained in the chimeric PCV2b vaccine (cPCV2b) and the bivalent chimeric vaccine (cPCV2a-cPCV2b). For ORF2, we compared three monovalent vaccines (a baculovirus expressed PCV2a ORF2 [VacAlt], a cPCVa [VacA], and an experimental PCV2b ORF2-based vaccine [VacB]), and one bivalent vaccine (a combination cPCV2a-cPCV2b vaccines [VacAB]) to 161 field strain sequences (15 partial and 146 full-length sequences). The vaccine and field strain sequences were examined for "shared" epitopes using JanusMatrix (Moise et al., 2013). Two putative epitopes are considered to be shared between the vaccine and field strain sequences if they were exactly matched or if both epitopes were predicted to bind the same SLA allele and both epitopes shared exactly matched TCR-facing residues.

For each pair of sequences in this analysis, the shared EpiCC scores were calculated as previously described (Gutiérrez et al., 2017). The same approach was adopted to compare the T cell epitope content of bivalent vaccines and field strains. For this scenario, the calculation assumed that the vaccine strain contains all the putative epitopes from each of its two components, but no extra score is awarded for identical epitopes that appear in both components of a bivalent vaccine.

Shared EpiCC scores were also calculated for each vaccine and field strain as compared to itself. Shared EpiCC scores derived from the comparison of any sequence to itself are referred to as "baseline" EpiCC scores. The greater the baseline score, the greater the epitope content of the sequence. Since no sequence can be better matched to a given sequence than itself, the maximum value for any comparison between a target sequence and a comparison sequence (in this case a vaccine sequence against a field strain) can only be less than or equal to their baseline EpiCC scores.

An average baseline score was also calculated by averaging the baseline scores for all field strains and subsets of field strains grouped by PCV2 genotype (Fig. 1A). The average baseline score represents the T cell epitope content of the field strain sequences. For each vaccine strain, an average shared EpiCC score was calculated by averaging the shared EpiCC scores for all vaccine-field strains comparisons and subsets of vaccine-field strain comparisons grouped by PCV2 genotype. The greater the average shared EpiCC score, the better a vaccine represents the T cell epitope content of the circulating strains and may indicate the vaccine potential to induce a broad T cell response against diverse challenge strains. The score of partial field strain sequences were excluded from the calculation of these averages.

The T cell epitope content of a field strain is considered to be well-matched or "covered" by a given vaccine sequence if the strain baseline and the shared EpiCC score are similar. To quantify vaccine T cell epitope coverage, the shared EpiCC score of each vaccine-field strain comparison was divided by that field strain's baseline EpiCC score and expressed as a percentage (Fig. 1B). The average vaccine T cell epitope coverage was determined by averaging the vaccine coverage for all field strain comparisons. Similar to the average shared EpiCC score, the greater the average T cell epitope coverage, the better the vaccine matches or covers the T cell epitope content of the circulating strain sequences. The scores of partial field strain sequences were excluded from the calculation of the average.

3. Results

3.1. Phylogenetic and population structure of PCV2 ORF2 sequences

BAPS analysis (Cheng et al., 2013) predicted presence of 11 distinct clusters within the analyzed 4500 field strains. The BAPS defined clusters were mapped to known PCV2 clusters (Table S1 and Fig. 2). Using the iterative clustering approach as described in the methods, 161 representative strains that span the known taxonomic diversity (except for PCV2c due to the paucity of published strains) were selected to do a further T cell epitope content analysis (Tables S1 and S2). Recombinant viruses were found in each cluster (and representative strains were included within the T cell epitope content analysis) except for cluster 11 (PCV2e). No recombinant PCV2e viruses were identified, likely due to the low prevalence and unconfirmed virulence of PCV2e. The strains used for T cell epitope content analysis included viruses from global locations.

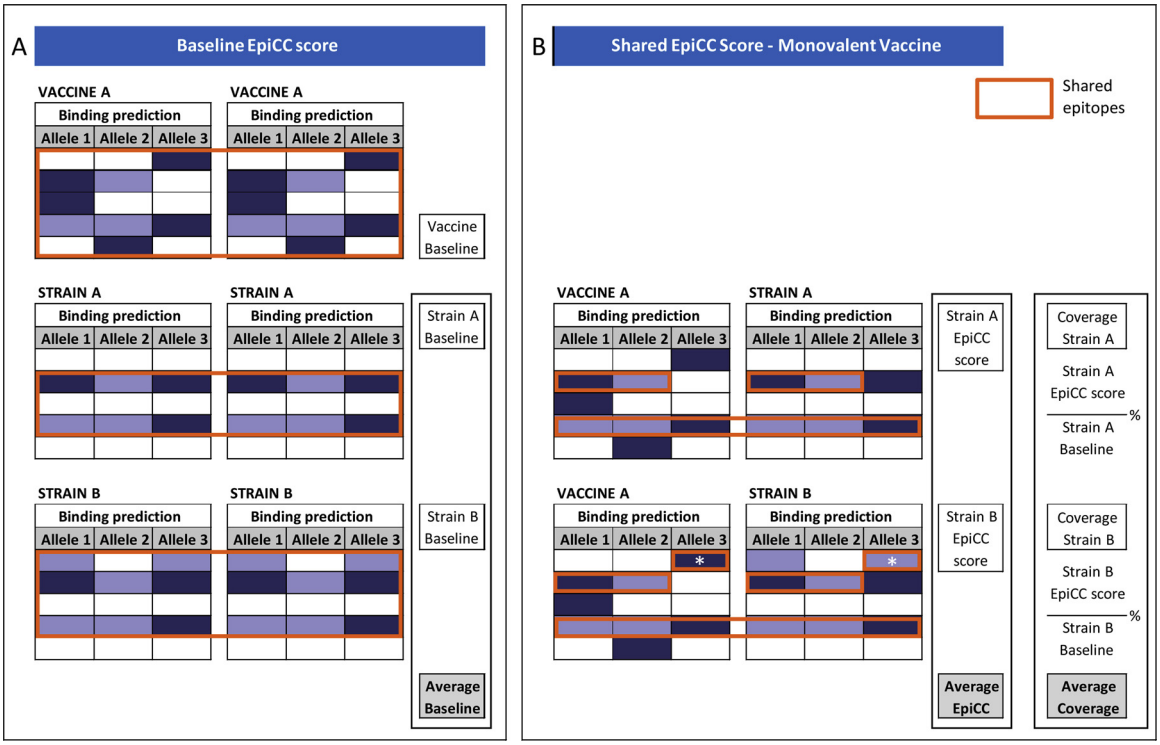

Fig. 1. T cell epitope content comparison (EpiCC) analysis. Vaccine A and Strains A and B were screened for binding likelihood to a set of MHC alleles; 9-mers (rows) predicted to bind to specific MHC alleles (columns) are shown in light (top 5 %) or dark (top 1 %) blue. The baseline EpiCC score is defined by comparing the epitope content of each sequence to itself (A). An average baseline score was calculated for the representative field strains. For the comparison between Vaccine A and Strains A and B, the EpiCC score was based on shared T cell epitope content (B). Note that cross-conserved T cell epitopes, as defined by JanusMatrix, were considered shared (*). The average EpiCC score and a vaccine T cell epitope average coverage score, for this analysis, were also calculated.

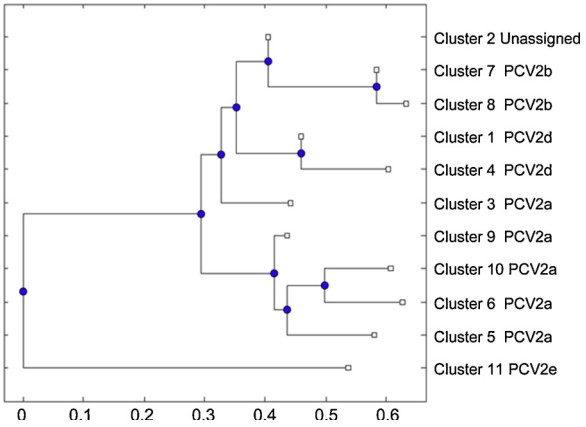

Fig. 2. Taxonomic relationship across the 11 PCV clusters as identified by BAPS clustering. BAPS (Bayesian Analysis of Population Structure) was used to cluster 4500 PCV2 ORF2 sequences. The genotype of each cluster is indicated, while cluster 2 was not assigned to a specific genotype.

3.2. T cell epitope content shared between vaccines and representative field strains is variable

The relatedness of the T cell epitope content shared between PCV2 vaccines and field strains was evaluated using the EpiCC algorithm. First, we calculated the maximum (baseline) EpiCC score for each sequence by comparing its T cell epitope content to self. Second, we compared the putative SLA class I and class II T cell epitope content of one ORF1 vaccine sequence (ORF1 PCV1) derived from the three chimeric virus vaccines, three monovalent ORF2 vaccine sequences (ORF2 VacAlt, ORF2 VacA and ORF2 VacB) and the combined T cell epitope content of ORF2 VacA and ORF2 VacB (VacAB) against that of ORF1 or ORF2 sequences from PCV2 field strains. The greater the similarity of the epitope content shared between the vaccine and field strain sequences, the greater the EpiCC score. Third, for each vaccine, we calculated average EpiCC scores considering sequences of all full-length field strains and by genotype. Fourth, for each vaccine-field strain comparison, the EpiCC score was divided by the field strain baseline EpiCC score, as a measure of vaccine T cell epitope coverage. To generalize vaccine T cell epitope coverage at the population level, the vaccine coverage for all field strains was averaged.

Total shared EpiCC scores (combined MHC class I and II) were plotted on radar plots. Fig. 3 presents the total shared EpiCC scores for ORF1 PCV1 compared to 91 field strain sequences and Fig. 4 shows EpiCC scores of the four ORF2 vaccine sequences (VacAlt, VacA, VacB, and VacAB) compared to 161 field strains. The EpiCC scores are displayed on these plots as a distance along a radiating line and each axis corresponds to one strain sequence.

3.2.1. ORF1

Class I and class II baseline EpiCC scores (4.37 and 4.45, respectively) were similar for ORF1 PCV1 vaccine (Table 2). Compared to the total average of the field strains (9.14), ORF1 PCV1 had a lower total baseline EpiCC score (8.81), which is driven by its lower than average class II baseline score (4.45 compared to 5.15 average). In terms of total shared EpiCC score, the average calculated for the set of field strains (5.09) represented 55.68 % of the average baseline (9.14), which was lower than the baseline of ORF2 vaccines. This result suggests that, on average, the vaccine covered more than half of the T cell epitope content found in field strains. Specifically, on average, ORF1 PCV1 shared 101.96 epitopes (class I and class II combined) with field strains (Table S3).

The T cell epitope content variability of ORF1 sequences was limited (average baseline sd = 0.14). This was also confirmed by the range of EpiCC scores for the vaccine strain to field strain comparison (average

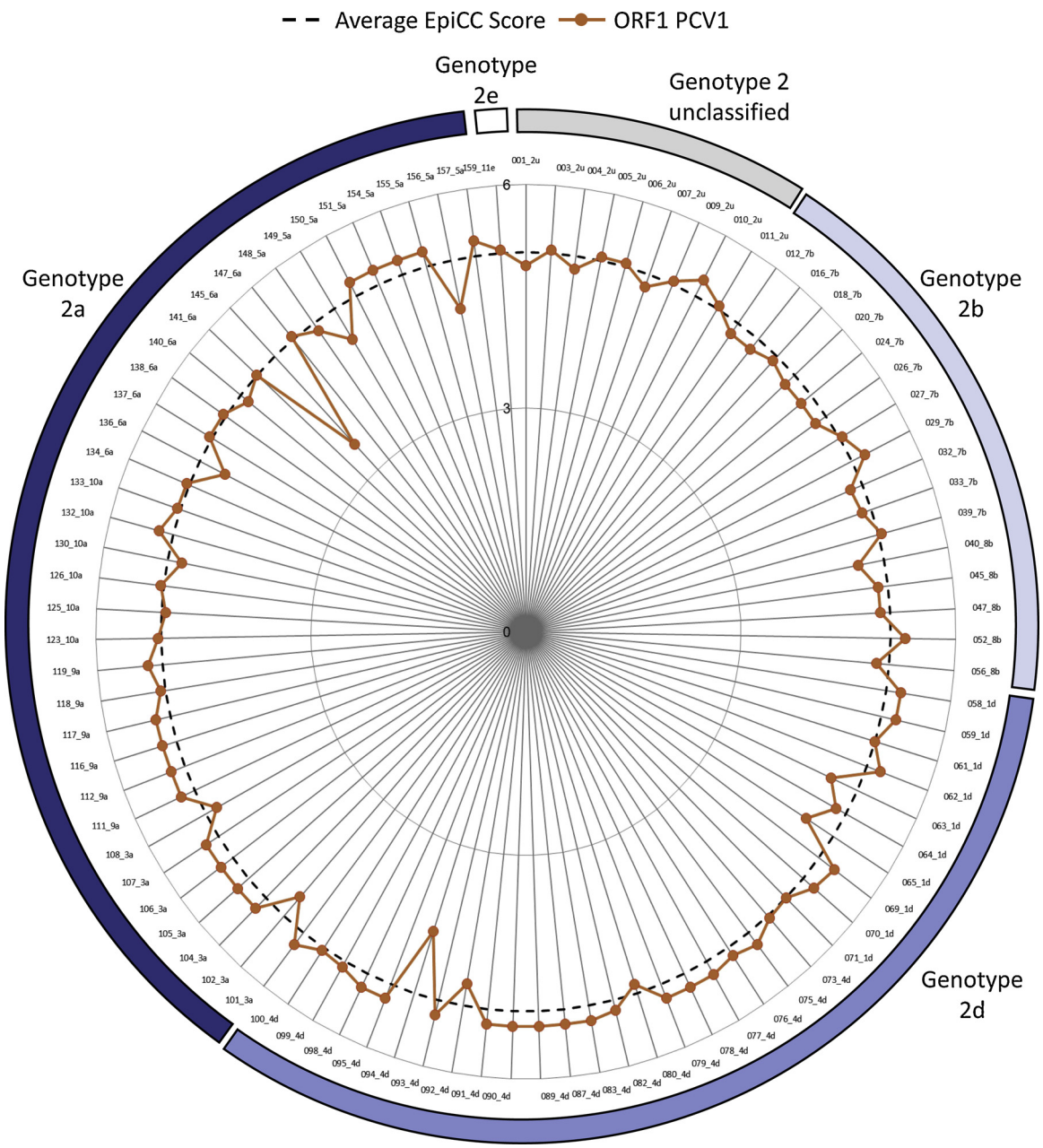

Fig. 3. ORF1 total shared EpiCC score. Radar plot shows EpiCC scores for comparison between ORF1 PCV1 vaccine against challenge strains and field PCV2 viruses. Each axis corresponds to the ORF1 sequence of one strain. Labels include sequence id number (001-161), cluster (1 - 11) and genotype 2 (a, b, d, e, or unclassified (u)). Average EpiCC score calculated including full-length sequences is also shown. Field strains are sorted by genotype.

EpiCC score sd = 0.28; Table 2). This limited variability is visible in Fig. 3, which shows the total shared EpiCC scores of ORF1 PCV1 compared to each analyzed ORF1 sequence. Furthermore, averages of EpiCC scores calculated by genotype were highly similar for ORF 1, only varying between 8.75 and 8.97 across the four genotypes (Table 3). This limited variability of ORF1 T contrasts with the higher variability of ORF2 EpiCC scores across genotypes (Table 3).

3.2.2. ORF2

For the three tested ORF2 monovalent vaccines (VacAlt, VacA and VacB), class I baseline EpiCC scores were higher than those of class II (Table 2). Compared to the average total baseline of the field strains (10.3), ORF2 VacA had a higher total baseline EpiCC score (10.42), which is explained by higher than average class I and class II baseline scores (6.64 and 3.78, respectively). ORF2 VacB had higher average class I and class II EpiCC scores (5.05 and 2.76, respectively) than ORF2 VacAlt and ORF2 VacA. Consequently, ORF2 VacB had the highest average total shared EpiCC score (7.81). This result showed that on average, ORF2 VacB covered 75.83 % of the T cell epitope content identified in field strains. In terms of number of epitopes, ORF2 sequences from field strains had on average 167.87 predicted epitopes (class I and class II) and 127.1 of them were shared with ORF2 VacB (Table S3).

For the T cell epitope content comparison of the bivalent vaccine ORF2 VacAB, the EpiCC approach assumes that this vaccine contains all the epitopes of ORF2 VacA and ORF2 VacB. Therefore, the shared epitope content between ORF2 VacAB and a field strain can only be equal to or higher than that of the individual component of the bivalent vaccine with the highest EpiCC score against the strain. Thus, compared to the monovalent ORF2 vaccines, the bivalent vaccine had the highest total baseline (15.35) and average total EpiCC score (9.04) (Table 2). Thus, VacAB covered on average, 87.49 % of the epitope content

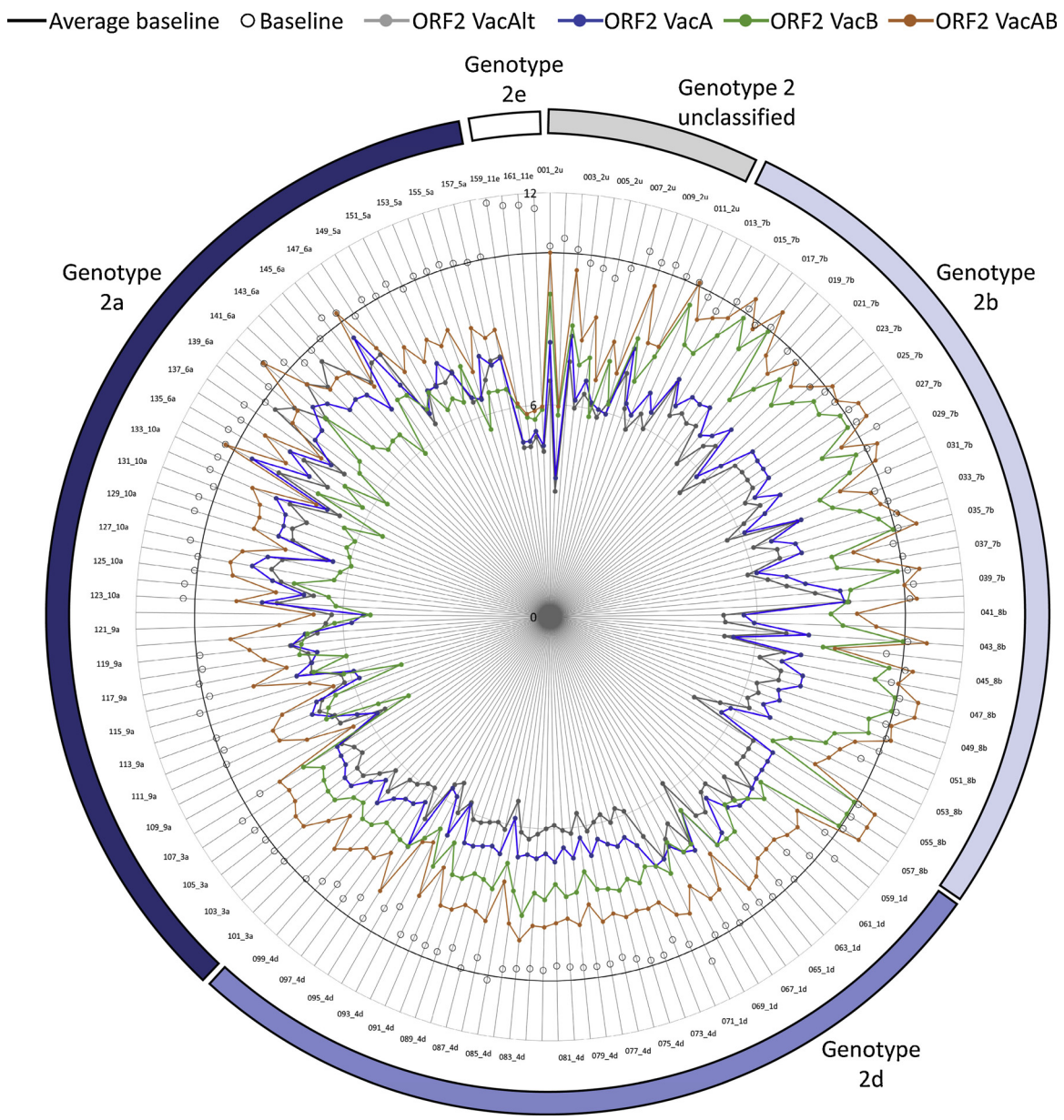

Fig. 4. ORF2 total shared EpiCC score. The radar plot compares EpiCC scores between ORF2 PCV2 vaccines (ORF2 VacAlt, ORF2 VacA, ORF2 VacB, and ORF2 VacAB) against field PCV2 viruses. Each axis corresponds to the ORF2 sequence of one strain. Labels include sequence id number (001-161), cluster (1 - 11) and genotype 2 (a, b, d, e, or unclassified (u)). Only labels for field strains with even id numbers are shown for simplicity. Baseline EpiCC scores for each sequence (open circles) and average EpiCC score calculated including full-length sequences are shown. Baseline EpiCC scores of partial sequences are not shown. Field strains are sorted by genotype.

identified in the field strains. Moreover, ORF2 VacAB shared, on average, 146.86 putative epitopes with the field strains (Table S3).

The differences in total shared EpiCC scores among the four ORF2 vaccines compared to each analyzed strain can be visualized in the radar plot presented in Fig. 4. ORF2 VacAB had the highest average total shared EpiCC score. However, ORF2 VacAlt had higher EpiCC scores than ORF2 VacAB for six strains from cluster 6 (PCV2a, 138,139, and 142-145). For the monovalent vaccines analyzed here, we observed that while ORF2 VacB had the highest average total shared EpiCC score of the monovalent vaccines, ORF2 VacA and ORF2 VacAlt had, in several instances, higher scores than those of ORF2 VacB. Additionally, there were discernible clusters of sequences for which ORF2 VacB or ORF2 VacA and ORF2 VacAlt had higher scores. For example, ORF2 VacB had higher scores for sequences from clusters 3, 4, 7, and 8. On the other hand, ORF2 VacA and ORF2 VacAlt had higher scores for sequences from clusters 6 and 10. In terms of scores calculated by genotype (Table 3), PCV2a vaccines (ORF2 VacA and VacAlt) had higher average EpiCC scores for PCV2a field strains, as expected. ORF2 VacA had higher scores than ORF2 VacAlt for all the genotypes. ORF2 VacB had higher average EpiCC scores not only for PCV2b strains, but also PCV2d and PCV2e. The bivalent vaccine ORF2 VacAB had the highest scores by genotype.

4. Discussion

We compared the T cell epitope content of four PCV2 vaccines (Table 1) (ORF1 for the PCV1-PCV2a chimeric virus vaccines only [cPCV2a; VacA] and ORF2 for all vaccines (baculovirus expressed PCV2a ORF2 [VacAlt], cPCV2a [VacA], experimental PCV2b ORF2-based vaccine [VacB], and combination cPCV2a-cPCV2b vaccine [VacAB]) against that of 161 field PCV2 strains. PCV2 field strains were selected to represent the diversity within PCV2. Overall, we identified

Table 2
Summary of EpiCC scores.

		VACCINES				
		ORF1 PCV1	ORF2 VacAlt	ORF2 VacA	ORF2 VacB	ORF2 VacAB
Class I	Vaccine baseline[a]	4.37	6.68	6.64	6.62	9.85
(8 alleles)	Average baseline (sd)[b]	3.99 (0.08)		6.62 (0.16)		
	Average EpiCC (sd)[c]	2.21 (0.12)	4.48 (0.68)	4.65 (0.61)	5.05 (0.8)	5.74 (0.68)
	Average coverage[d]	55.38 %	67.71 %	70.27 %	76.29 %	86.77 %
Class II	Vaccine baseline	4.45	3.62	3.78	3.62	5.50
(5 alleles)	Average baseline (sd)	5.15 (0.09)		3.69 (0.33)		
	Average EpiCC (sd)	2.88 (0.17)	2.32 (0.4)	2.54 (0.35)	2.76 (0.51)	3.3 (0.49)
	Average coverage	55.91 %	62.83 %	68.71 %	74.79 %	89.38 %
Total (class I and II alleles)	Vaccine baseline	8.81	10.29	10.42	10.25	15.35
	Average baseline (sd)	9.14 (0.14)		10.3 (0.37)		
	Average EpiCC (sd)	5.09 (0.28)	6.8 (1.02)	7.19 (0.89)	7.81 (1.26)	9.04 (1.11)
	Average coverage	55.68 %	66.03 %	69.78 %	75.83 %	87.79 %

[a] EpiCC score calculated for the vaccine compared to itself.
[b] Average baseline EpiCC score (and standard deviation) of full-length field strains.
[c] Average EpiCC score (and standard deviation) of the vaccine compared to full-length field strains.
[d] Average coverage of each field strain's baseline EpiCC score expressed as a percentage.

Table 3
Total (class I and II alleles) average shared EpiCC score by genotype.

Genotype	Number of full-length sequences		VACCINES[a]				
	ORF1	ORF2	ORF1 PCV1	ORF2 VacAlt	ORF2 VacA	ORF2 VacB	ORF2 VacAB
PCV2a	33	49	8.74	7.66	7.74	6.88	8.83
PCV2b	25	50	8.84	6.55	7.04	8.99	9.79
PCV2d	29	43	8.78	6.29	6.92	7.68	8.71
PCV2e	1	4	8.97	4.87	5.03	5.80	5.92

[a] Average EpiCC score was calculated including only full-length field strains within a given genotype. Higher average shared EpiCC scores represent more similarity between T cell epitope content of a vaccine and field strains.

sequences matching the most prevalent strains (a, b, and d) as well as those at lower prevalence (c, e, and f) across global locations. The field strains generally clustered into 2 major genogroups, the PCV2a and PCV2d/PCV2b groups (Fig. 2). PCV2c was not included in the EpiCC analysis due to its low prevalence. To complete our analysis, we performed individual comparisons for PCV2 ORF1 and ORF2 sequences using the EpiCC algorithm.

The results showed that the ORF1 PCV1 sequence contained in the chimeric PCV1-PCV2 vaccine (VacA) shared 55.68 % of the average total T cell epitope content found in the field strains (Table 2). ORF1 PCV1 shared on average 101.96 epitopes (class I and class II combined) with the field strains (Table S3). Moreover, we also found limited variability in the T cell epitope content shared between the ORF1 PCV1 vaccine and that of the ORF1 sequences from PCV2 field strains. While PCV2 viruses have a high evolutionary rate, estimated to be 1.2×10^{-3} substitutions on the whole genome/site/year (Firth et al., 2009), the PCV1 genome has relatively less diversity and a correspondingly lower evolutionary rate, estimated to be 1.15×10^{-5} substitutions on the whole genome/site/year (Cortey and Segalés, 2012). Furthermore, compared to the epitope content of ORF2 sequence, ORF1 was less variable, in agreement with observations by others (Franzo et al., 2016; Xiao et al., 2015). This may be due to variability within *replicase* leading to unfit viruses or lethal mutants. ORF1 has been shown to be an important source of epitopes for T cells (Stevenson et al., 2007).

Considering ORF2 sequences, ORF2 VacB (PCV2b) had the highest average total shared EpiCC score of the monovalent vaccines for the set of field strains. ORF2 VacB shared, on average, 75.83 % of the putative epitope content identified in the field strains (Table 2). ORF2 VacB shared on average 127.1 epitopes with the field strains and more T cell epitope content with PCV2d and PCV2e strains than ORF2 VacA and VacAlt (Table S3). This suggests that the T cell epitope content of the ORF2 VacB vaccine may have the broadest coverage of the three analyzed ORF2 monovalent vaccine sequences (VacAlt, VacA, and VacB).

In addition to the monovalent vaccines (VacAlt, VacA, and VacB), we compared the T cell epitope content of one bivalent vaccine, ORF2 VacAB, a combination of two of the monovalent vaccines, ORF2 VacA (PCV2a) and ORF2 VacB (PCV2b), against the field PCV2 strains. For this comparison, we utilized the same calculation applied for our EpiCC analysis of monovalent vaccines. ORF2 VacAB shared, on average, 87.79 % epitope content with field strains (Table 2), which translates to 146.86 shared epitopes (Table S3). The bivalent vaccine shared on average more T cell epitope content with strains from all the different genotypes described here than monovalent vaccines. Therefore, compared to the monovalent vaccines, ORF2 VacAB, the bivalent vaccine, may confer broader T cell epitope coverage for this set of field strains. Thus, for the set of analyzed field strains, the vaccines that may confer broader cross-reactive cell-mediated immune response and protection in descending order are: ORF2 VacAB, ORF2 VacB, ORF2 VacA, and ORF2 VacAlt.

The combination PCV2a-PCV2b vaccine had, on average, the highest EpiCC score, thus the best potential to confer the broadest cross-reactive cell-mediated immunity and protection. Given that the PCV2 field viruses grouped into two major clusters, namely PCV2a and PCV2d/b (Fig. 2), having strains representing epitopes from both clusters in the vaccines reasonably should broaden coverage, on average, to all field viruses within the data set. This prediction of improved coverage compared to monovalent PCV2a-based vaccines is especially apparent against PCV2b, as well as PCV2d. Increased coverage offered to PCV2d based on the PCV2b addition is supported by the greater sequence identity and experimental epitope conservation among PCV2b and PCV2d compared to PCV2a and PCV2d (Xiao et al., 2015) and (Constans et al., 2015), respectively. Despite being the same PCV2a strain as in the Vac A, the bivalent vaccine also has a positive impact on coverage of PCV2a strains and this is likely due to the large diversity and number of recombinant PCV2a viruses.

Our current EpiCC algorithm was developed for comparisons against monovalent vaccines. This calculation does not account for the differences in delivered weights of each component in a bivalent vaccine and target strains. Future models will explore the potential effects of epitopes unique to each component of a multivalent vaccine and epitopes shared within the components. While results of these models might differ from the current scores, they will not alter our conclusion regarding ORF2 VacAB.

EpiCC scores depend on SLA-specific predictions, and it is certain that different sets of SLA alleles will produce different EpiCC scores. However, given the large difference in total shared EpiCC scores among vaccines, we speculate that different sets of SLA alleles may produce consistent results. The impact of SLA alleles on EpiCC scores will be assessed as data about SLA frequencies in North American swine populations become available.

Molecular epidemiological investigations into the evolution of PCV2 genotypes suggest that divergence directed at the shared antigenic determinants of PCV2a-based vaccines and field strains is clear for PCV2a field strains and that there is less divergence for PCV2b and PCV2d field strains (Franzo et al., 2016). It has also been shown that the main areas of divergence are located in regions (Franzo et al., 2016; Xiao et al., 2015) of epitopic determinants (Lekcharoensuk et al., 2004; Saha et al., 2012; Trible et al., 2012). In other words, immune pressure has at least contributed to the selection for and directed the evolution of PCV2 away from vaccine strains. Since less diversifying and directional evolution has been observed for PCV2b and PCV2d compared to PCV2a, it is reasonable that cross-protection offered by PCV2a-based vaccines is only partial. Animal studies have confirmed that PCV2 vaccines can offer cross-protection even in the scenario where vaccine and challenge genotypes are mismatched (Fort et al., 2008; Opriessnig et al., 2014); however, protection may be superior when multiple genotypes exist, or the vaccine and challenge genotypes are matched (Beach et al., 2010; Opriessnig et al., 2013). Since sequence differences have been shown to translate into differential immune recognition (Kurtz et al., 2014; Saha et al., 2012), incomplete cross-protection makes sense. Despite disease control and PCV2 vaccination success (PCV2 vaccines represent the best-selling vaccine within the global swine market), PCV2 infection is still widespread (Karuppannan and Opriessnig, 2017). Cross-protection being only partial, the rapid evolutionary rate of PCV2 viruses, and other reasons including modern animal management and movement, helps to explain why PCV2a, PCV2b, and PCV2d are all still circulating globally even in vaccinated animals.

Updating PCV2 vaccines or developing multivalent PCV2 vaccines representing greater epitope similarity to circulating strains, and in-turn broader immune coverage, are necessary to protect against circulating and emerging strains of PCV2 (Meng, 2013; Segalés, 2015; Ssemadaali et al., 2015). The combination of PCV2a and PCV2b offers significant immunological coverage against all PCV2 genotypes and not only PCV2a and PCV2b but also to PCV2d, i.e. broad cross-protection. In light of the current situation, where the genetic gap among field strains and vaccines is increasing and where having multiple genotypes offers broader protection to any strain, utilizing a vaccine containing PCV2a and PCV2b is appropriate.

This study did not directly assess the ability of the PCV2a-PCV2b combination vaccine to confer broad cross-protection in vivo. A high EpiCC score can be predicted to correlate to protection. EpiCC scores (among vaccine and field strains) that result in protection have been identified for influenza (Gutiérrez et al., 2017). In the case of influenza, the EpiCC scores that inform efficacy are based on efficacy studies. The goal of future studies will be to identify the EpiCC level associated with protection for PCV2 as well.

Declaration of Competing Interest

Funding for this project was provided by Zoetis. MB, PD, GR and DLF were employees of Zoetis at the time of the study. Zoetis scientists provided the vaccine sequences and conducted the initial analysis and clustering of field sequences (obtained from GenBank). The PigMatrix and EpiCC analyses and interpretation were performed independently by AHG, WDM, FET and ADG. EpiVax authors report that they are employees and/or founders and shareholders of EpiVax and therefore can be considered to have a potential conflict of interest, however, the study was performed in a blinded, prospective fashion with no foreknowledge of the potential relevance of the results.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.vetimm.2020.110034.

References

Bao, F., Mi, S., Luo, Q., Guo, H., Tu, C., Zhu, G., Gong, W., 2018. Retrospective study of porcine circovirus type 2 infection reveals a novel genotype PCV2f. Transbound. Emerg. Dis. 65, 432–440.

Beach, N.M., Ramamoorthy, S., Opriessnig, T., Wu, S.Q., Meng, X.-J., 2010. Novel chimeric porcine circovirus (PCV) with the capsid gene of the emerging PCV2b subtype cloned in the genomic backbone of the non-pathogenic PCV1 is attenuated in vivo and induces protective and cross-protective immunity against PCV2b and PCV2a subtypes in pigs. Vaccine 29, 221–232.

Cheng, L., Connor, T.R., Siren, J., Aanensen, D.M., Corander, J., 2013. Hierarchical and spatially explicit clustering of DNA sequences with BAPS software. Mol. Biol. Evol. 30, 1224–1228.

Constans, M., Ssemadaali, M., Kolyvushko, O., Ramamoorthy, S., 2015. Antigenic determinants of possible vaccine escape by porcine circovirus subtype 2b viruses. Bioinform. Biol. Insights 1–12.

Cortey, M., Segalés, J., 2012. Low levels of diversity among genomes of porcine circovirus type 1 (PCV1) points to differential adaptive selection between Porcine circoviruses. Virology 422, 161–164.

Fenaux, M., Opriessnig, T., Halbur, P.G., Elvinger, F., Meng, X.J., 2004. A chimeric porcine circovirus (PCV) with the immunogenic capsid gene of the pathogenic PCV type 2 (PCV2) cloned into the genomic backbone of the nonpathogenic PCV1 induces protective immunity against PCV2 infection in pigs. J. Virol. 78, 6297–6303.

Firth, C., Charleston, M.A., Duffy, S., Shapiro, B., Holmes, E.C., 2009. Insights into the evolutionary history of an emerging livestock pathogen: porcine circovirus 2. J. Virol. 83, 12813–12821.

Fort, M., Sibila, M., Allepuz, A., Mateu, E., Roerink, F., Segalés, J., 2008. Porcine circovirus type 2 (PCV2) vaccination of conventional pigs prevents viremia against PCV2 isolates of different genotypes and geographic origins. Vaccine 26, 1063–1071.

Franzo, G., Segalés, J., 2018. Porcine circovirus 2 (PCV-2) genotype update and proposal of a new genotyping methodology. PLoS One 13 e0208585–e0208585.

Franzo, G., Cortey, M., Segalés, J., Hughes, J., Drigo, M., 2016. Phylodynamic analysis of porcine circovirus type 2 reveals global waves of emerging genotypes and the circulation of recombinant forms. Mol. Phylogenet. Evol. 100, 269–280.

Gerber, P.F., Johnson, J., Shen, H., Striegel, D., Xiao, C.-T., Halbur, P.G., Opriessnig, T., 2013. Association of concurrent porcine circovirus (PCV) 2a and 2b infection with PCV associated disease in vaccinated pigs. Res. Vet. Sci. 95, 775–781.

Gutiérrez, A.H., Martin, W.D., Bailey-Kellogg, C., Terry, F., Moise, L., De Groot, A.S., 2015. Development and validation of an epitope prediction tool for swine (PigMatrix) based on the pocket profile method. BMC Bioinformatics 16 290–290.

Gutiérrez, A.H., Loving, C., Moise, L., Terry, F.E., Brockmeier, S.L., Hughes, H.R., Martin, W.D., De Groot, A.S., 2016. In vivo validation of predicted and conserved t cell epitopes in a swine influenza model. PLoS One 11, 1–19.

Gutiérrez, A.H., Rapp-Gabrielson, V.J., Terry, F.E., Loving, C.L., Moise, L., Martin, W.D., De Groot, A.S., 2017. T-cell epitope content comparison (EpiCC) of swine H1 influenza A virus hemagglutinin. Influenza Respir. Viruses 11, 531–542.

Hesse, R., Kerrigan, M., Rowland, R.R.R., 2008. Evidence for recombination between PCV2a and PCV2b in the field. Virus Res. 132, 201–207.

Karuppannan, A.K., Opriessnig, T., 2017. Porcine circovirus type 2 (PCV2) vaccines in the context of current molecular epidemiology. Viruses 9 (1999–4915), 99.

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.

Kurtz, S., Grau-Roma, L., Cortey, M., Fort, M., Rodríguez, F., Sibila, M., Segalés, J., 2014. Pigs naturally exposed to porcine circovirus type 2 (PCV2) generate antibody responses capable to neutralise PCV2 isolates of different genotypes and geographic origins. Vet. Res. 45 443–443.

Lefebvre, D.J., Van Doorsselaere, J., Delputte, P.L., Nauwynck, H.J., 2009. Recombination of two porcine circovirus type 2 strains. Arch. Virol. 154, 875–879.

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, 8135–8145.

Meng, X.J., 2013. Porcine circovirus type 2 (PCV2): pathogenesis and interaction with the immune system. Annu. Rev. Anim. Biosci. 1, 43–64.

Moise, L., Gutierrez, A.H., Bailey-Kellogg, C., Terry, F., Leng, Q., Abdel Hady, K.M., VerBerkmoes, N.C., Sztein, M.B., Losikoff, P.T., Martin, W.D., Rothman, A.L., De Groot, A.S., 2013. The two-faced T cell epitope: examining the host-microbe interface with JanusMatrix. Hum. Vaccin. Immunother. 9, 1577–1586.

Olvera, A., Cortey, M., Segalés, J., 2007. Molecular evolution of porcine circovirus type 2 genomes: phylogeny and clonality. Virology 357, 175–185.

Opriessnig, T., O'Neill, K., Gerber, P.F., de Castro, A.M.M.G., Gimenéz-Lirola, L.G., Beach, N.M., Zhou, L., Meng, X.-J., Wang, C., Halbur, P.G., 2013. A PCV2 vaccine based on genotype 2b is more effective than a 2a-based vaccine to protect against PCV2b or combined PCV2a/2b viremia in pigs with concurrent PCV2, PRRSV and PPV infection. Vaccine 31, 487–494.

Opriessnig, T., Gerber, P.F., Xiao, C.-T., Halbur, P.G., Matzinger, S.R., Meng, X.-J., 2014. Commercial PCV2a-based vaccines are effective in protecting naturally PCV2b-

infected finisher pigs against experimental challenge with a 2012 mutant PCV2. Vaccine 32, 4342–4348.
Palinski, R., Piñeyro, P., Shang, P., Yuan, F., Guo, R., Fang, Y., Byers, E., Hause, B.M., 2017. A novel porcine circovirus distantly related to known circoviruses is associated with porcine dermatitis and nephropathy syndrome and reproductive failure. J. Virol. 91, e01879–1816.
Price, M.N., Dehal, P.S., Arkin, A.P., 2010. FastTree 2–approximately maximum-likelihood trees for large alignments. PLoS One 5, e9490.
Reiner, G., Hofmeister, R., Willems, H., 2015. Genetic variability of porcine circovirus 2 (PCV2) field isolates from vaccinated and non-vaccinated pig herds in Germany. Vet. Microbiol. 180, 41–48.
Saha, D., Huang, L., Bussalleu, E., Lefebvre, D.J., Fort, M., Van Doorsselaere, J., Nauwynck, H.J., 2012. Antigenic subtyping and epitopes' competition analysis of porcine circovirus type 2 using monoclonal antibodies. Vet. Microbiol. 157, 13–22.
Segalés, J., 2012. Porcine circovirus type 2 (PCV2) infections: clinical signs, pathology and laboratory diagnosis. Virus Res. 164, 10–19.
Segalés, J., 2015. Best practice and future challenges for vaccination against porcine circovirus type 2. Expert Rev. Vaccines 14, 473–487.
Segalés, J., Olvera, A., Grau-Roma, L., Charreyre, C., Nauwynck, H., Larsen, L., Dupont, K., McCullough, K., Ellis, J., Krakowka, S., Mankertz, A., Fredholm, M., Possum, C., Timmusk, S., Stockhofe-Zurwieden, N., Beattie, V., Armstrong, D., Grassland, B., Baekbo, P., Allan, G., 2008. PCV-2 Genotype Definition and Nomenclature, Veterinary Record. British Veterinary Association, London, UK, pp. 867–868.
Ssemadaali, M.A., Ilha, M., Ramamoorthy, S., 2015. Genetic diversity of porcine circovirus type 2 and implications for detection and control. Res. Vet. Sci. 103, 179–186.
Stevenson, L.S., Gilpin, D.F., Douglas, A., McNeilly, F., McNair, I., Adair, B.M., Allan, G.M., 2007. T lymphocyte epitope mapping of porcine circovirus type 2. Viral Immunol. 20, 389–398.
Sturniolo, T., Bono, E., Ding, J., Raddrizzani, L., Tuereci, O., Sahin, U., Braxenthaler, M., Gallazzi, F., Protti, M.P., Sinigaglia, F., Hammer, J., 1999. Generation of tissue-specific and promiscuous HLA ligand databases using DNA microarrays and virtual HLA class II matrices. Nat. Biotechnol. 17, 555–561.
Trible, B.R., Ramirez, A., Suddith, A., Fuller, A., Kerrigan, M., Hesse, R., Nietfeld, J., Guo, B., Thacker, E., Rowland, R.R.R., 2012. Antibody responses following vaccination versus infection in a porcine circovirus-type 2 (PCV2) disease model show distinct differences in virus neutralization and epitope recognition. Vaccine 30, 4079–4085.
Xiao, C.T., Halbur, P.G., Opriessnig, T., 2015. 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. 96, 1830–1841.
Xiao, C.-T., Harmon, K.M., Halbur, P.G., Opriessnig, T., 2016. PCV2d-2 is the predominant type of PCV2 DNA in pig samples collected in the U.S. during 2014–2016. Vet. Microbiol. 197, 72–77.
Yang, S., Yin, S., Shang, Y., Liu, B., Yuan, L., Zafar Khan, M.U., Liu, X., Cai, J., 2017. Phylogenetic and genetic variation analyses of porcine circovirus type 2 isolated from China. Transbound. Emerg. Dis. 65, e383–e392.