

Quantifying the Persistence of Vaccine-related T cell Epitopes in Circulating Swine Influenza A Strains from 2013-2017

Swan Tan ^{1,2}, Andres H. Gutiérrez ⁴, Philip C. Gauger ³, Tanja Opriessnig ^{3,4}, Justin Bahl ^{1,2}, Leonard Moise ^{2,5} and Anne S. De Groot ^{2,5*}

¹ Department of Infectious Diseases, University of Georgia, Athens, GA, USA

² Center for Vaccines and Immunology, University of Georgia, Athens, GA, USA

³ Department of Veterinary Diagnostic and Production Animal Medicine, College of Veterinary Medicine, Iowa State University, Ames, IA, USA

⁴ Roslin Institute, University of Edinburgh, Midlothian, Scotland, UK

⁵ EpiVax Inc., Providence, RI, USA

* Correspondence: dr.annie.degroot@gmail.com; Tel.: +1-401-952-4227 (R.I.)

Abstract: When swine flu vaccines and circulating influenza A virus (IAV) strains are poorly matched, vaccine-induced antibodies may not protect from infection. Highly conserved T cell epitopes may, however, have a disease-mitigating effect. The degree of T cell epitope conservation among circulating strains and vaccine strains can vary, which may also explain differences in vaccine efficacy. Here, we evaluate a previously developed conserved T cell epitope-based vaccine and determine the persistence of T cell epitope conservation over time. We used a pair-wise homology score to define conservation between the vaccine's swine leukocyte antigen (SLA) class I and II-restricted epitopes and T cell epitopes found in 1,272 swine IAV strains sequenced between 2013 and 2017. Twenty-four of the 48 total T cell epitopes included in the epitope-based vaccine were highly conserved and found in >1,000 circulating swine IAV strains over the five-year period. In contrast, commercial swine IAV vaccines developed in 2013 exhibited declining conservation with the circulating IAV strains over the same five-year period. Conserved T cell epitope vaccines may be useful adjunct for commercial swine flu vaccines and to improve protection against influenza when antibodies are not cross-reactive.

Keywords: Swine IAV, Immunoinformatics, T cell epitope conservation

1. Introduction

When a new strain of a pathogen emerges, the first question asked is often whether existing vaccines might be effective against it. In the past, experts have relied on examining the humoral immune response by using antibody assays to determine the potential of existing vaccines to cross-protect [1]. It is now well-established that cell-mediated immunity (CMI) contributes to protection against severe disease even in the absence of antibody response [2–5]. CMI involves cytotoxic T lymphocytes (CTL) and T helper (Th) lymphocytes, which are triggered to respond when their T cell receptors (TCR) recognize T cell epitopes presented by class I or class II major histocompatibility complex (MHC) molecules on the surface of antigen presenting cells or infected cells [6].

In humans, immune responses to conserved T cell epitopes may result in reduced morbidity, despite the lack of cross-reactive antibody to the new strain [7–9]. This is supported by a case-controlled study that investigated the association of the pandemic IAV H1N1 2009 infections with 2008-2009 seasonal trivalent inactivated flu vaccination [10]. Previous seasonal vaccination protected against pandemic H1N1, despite the lack of antibody protection [11]. Using immunoinformatic tools available to us at the time, we defined T cell epitopes that were present in the newly emergent strain (pH1N1 A/California/04/2009; GenBank accession numbers ACP41105 for hemagglutinin or HA and ACP41107 for the neuraminidase or NA), and highly conserved in the existing seasonal influenza

vaccine (containing H1N1 A/Brisbane/59/2007; GenBank accession numbers ACA28844 for HA and ACA28847 for NA). The *in silico* analysis demonstrated that despite the lack of antibody cross-reactivity, more than 50% of T cell epitopes in the novel pH1N1 virus were also present in the seasonal vaccine, supporting the concept that pre-existing T cell response due to vaccination or exposure may have protected in the absence of protective antibody response. Prospective animal studies also confirmed that seasonal H1N1 vaccines that did not induce cross-reactive antibody responses, but induced cross-reactive T cell responses, did not protect against pandemic pH1N1 infection, but greatly reduced morbidity, mortality, virus replication, and viral shedding [12]. Thus, T cell epitopes can be conserved between seasonal vaccines and emerging influenza strains, thereby contributing to protection.

Knowing that IAVs have broad host tropism and are able to infect hosts ranging from wild birds to mammals which include humans, ferrets, dogs, horses, and pigs. Antigenic shift and drift are significant challenges not only to human seasonal vaccination but also to effective swine flu vaccination over time. The segmented IAV genome allows for antigenic shift by reassortment of RNA segments from different viral strains, generating novel viruses [13]. Antigenic drift that is due to gradual accumulation of mutations in the HA and NA surface antigens over time also contributes to the remarkable diversity of IAVs co-circulating among swine populations. This sequence-level diversity can impact T cell response since even single amino acid modifications to T cell epitopes can reduce Human Leukocyte Antigen (HLA) binding or T cell recognition, leading to viral escape and viral camouflage [14, 15] that contribute to lower vaccine efficacy. An additional problem facing swine influenza vaccine developers is the high diversity of circulating IAV genotypes impacting individual pork farms each year [16], making it difficult to know whether a given commercial vaccine will be protective.

To address these challenges in pigs, we applied previously developed computational method for estimating the degree of epitope conservation between vaccines and outbreak strains. Rather than focus on sequence identity, this algorithm identifies individual epitopes, and searches for epitope pairs that share MHC-binding properties and have identical TCR-facing residues, while allowing for amino acid variability at the T cell epitope (the HLA-binding-pocket facing amino acid residues). For swine IAV, the first step is to use SLA prediction matrices (PigMatrix) [17], and once an SLA-binding T cell epitope is predicted, JanusMatrix is applied to isolate the TCR-facing residues for comparison with similar SLA-binding epitopes in other circulating influenza strains [18, 19]. The third step is to use the Epitope Content Comparison (EpiCC) tool, which compiles the similarities and differences between the T cell epitopes in the vaccine and the circulating strain, assigning a score that reflects the degree of conservation [19].

Here, we compare a computationally designed swine flu vaccine based on conserved T cell epitopes called multi-epitope vaccine (MEpiV) with commercially available inactivated full strain vaccines using the computational algorithms described above. The MEpiV is composed of immunoinformatic-identified conserved SLA class I and class II epitopes assembled head-to-tail as class I and class II poly-epitope genes and formulated for delivery in a DNA vaccine vector [20]. The MEpiV was previously shown to be protective in a heterologous prime-boost vaccination and challenge study when combined with whole-inactivated vaccine [21]. In this study, we determine if the conserved T cell epitope-based vaccine would maintain conservation with circulating strain T cell epitopes over time.

To evaluate T cell epitope conservation for the MEpiV vaccine and to compare the conservation of the epitopes selected in 2013 to circulating strains for subsequent years, we used the HA sequence of seasonal inactivated swine flu vaccines as a benchmark for comparison. We then applied EpiCC and demonstrated that the MEpiV vaccine designed using computational tools in 2013 maintains >50% conservation with circulating strains over a five-year period. As can be expected, EpiCC also indicated that T cell epitopes in commercial seasonal vaccines are less well conserved over the same

time period. This evaluation of vaccines using EpiCC shows the approach to understanding T cell epitope conservation and the utility of the tool for comparing vaccines against emerging influenza strains. The analysis also reinforces the utility of designing influenza vaccines based on highly conserved epitopes from circulating viral strains, as these epitopes may be conserved over time.

2. Materials and Methods

Datasets. The sequences of all available H1N1, H1N2 and H3N2 swine IAV genomes circulating during the five year period 2013-2017 were obtained from the NIAID Influenza Research Database (<https://www.fludb.org/>) [9]. All genome sequences were downloaded and pre-processed to remove duplicated sequences. The final set of 1272 whole genome sequences were translated into protein sequences and were compared to the epitope-based DNA vaccine, MEpiV (Supplemental table 1) using an immunoinformatics approach as described below. In order to further evaluate the conservation of MEpiV, the sequence from two standard inactivated swine IAV vaccine antigens (FluSureXP 2016) were included for comparison. HA sequences of inactivated swine IAV vaccine comprised of one H1N1, one H1N2 and two H3N2 strains were provided by Zoetis to facilitate the comparison of T cell epitope conservation in epitope-based and inactivated virus vaccine, applying the same immunoinformatics analysis pipeline. The level of conservation for each of the vaccines (HA from H1N1 to H3N2, the MEpiV as compared to a representative set of IAV strains for each year was measured relative to 2013 to obtain relative changes in the number of T cell epitopes that were conserved.

Immunoinformatics tools. Three separate algorithms were used to evaluate conservation of vaccine epitopes contained in the target vaccine against the complete set of swine IAV sequences: (1) PigMatrix, which defines T cell epitopes for swine class I and class II epitopes, (2) JanusMatrix (JMX), a tool for identifying for epitopes that can be compared between strains by looking for epitopes that bind to the same allele and have conserved TCR facing residues which be used to compare strains, and (3) EpiCC, the T cell epitope content comparison algorithm utilizes results generated from PigMatrix and JMX and produces an overall score for class I and/or class II epitopes on a whole antigen level to enable pairwise comparisons between circulating IAV and vaccine strains (See Figure 1). A total of 1272 pairwise comparisons were performed, comparing each 9-mer sequence for possible conservation of SLA binding and TCR face, between MEpiV vaccine and circulating strains. EpiCC examines all of the epitopes in a vaccine against all of the epitopes in a given strain and produces an overall score for all class I or class II epitopes for each strain sequence. In addition to EpiCC, we used JanusMatrix to perform the same comparison on an epitope-by-epitope basis for the 28 class I and 20 class II epitopes in the vaccine.

T cell epitope prediction using PigMatrix. Using the pocket profile method and well-defined EpiMatrix binding preferences for human MHC pockets, we developed PigMatrix prediction matrices as previously described [12 - 13]. Matrices were designed based on the binding preferences of the best-matched Human Leukocyte Antigen (HLA) pocket for each SLA pocket. The contact residues involved in the binding pockets were defined from crystal structures of SLA or HLA supertype alleles for class I and II, respectively. Allele selection was based on prior data indicating their prevalence in outbred swine populations [14 - 15]. Matrices were constructed to predict T cell epitope binding to class I (SLA-1*0101, 1*0401, 1*0801, 1*1201, 1*1301, 2*0101, 2*0401, 2*0501, 2*1201, 3*0401, 3*0501, 3*0601, 3*0701) and class II (SLA-DRB1*0101, 0201, 0401, 0402, 0601, 0602, 0701 and 1001) SLA alleles. PigMatrix raw scores were standardized to Z-scores to compare potential epitopes across multiple SLA alleles. Peptides with Z-scores ≥ 1.64 (the top 5% of any given sample of 9-mers) were identified as likely to be SLA ligands.

Identification of conserved vaccine epitopes in different circulating swine IAV subtypes. JanusMatrix (JMX) is another immunoinformatics algorithm, which was incorporated to prospectively identify conserved vaccine epitopes among prevalent swine IAV [16]. JMX is used to

facilitate epitope to epitope-based comparison between swine IAV protein sequences and the vaccine strain. Conserved peptides at TCR-face were searched against all the circulating strains and hence, the presence of these peptides can be identified when there are matches in each individual strain.

T cell epitope content comparison (EpiCC) analysis. In order to determine conservation of the T cell epitopes in the MEpiV among the three-circulating swine IAV subtypes, we applied EpiCC to facilitate pairwise comparison of protein sequences. This method of comparison is based on an immunological property expressed in terms of T cell epitope content which incorporated JMX computation, rather than sequence identity. Shared (conserved) T cell epitopes between the vaccine target and the circulating swine IAV strains were evaluated. The assumption was based on given an epitope i and j from different strains (the circulating strain, s and the vaccine strain, v), cross-reactive memory T cells can be activated by epitopes with identical TCR-facing residues (TCR f) bind to the same alleles. Potential cross-reactive of class I epitope is calculated by considering identical residues at positions 4, 5, 6, 7, and 8 and for class II, the calculation is taking into account identical residues at positions 2, 3, 5, 7, and 8. Therefore, the probability to induce cross T-cell immunity is computed based on the following equation and p stands for the probability for epitope binding to the HLA allele:

$$S(i, j)_a = p(i)_a \cdot p(j)_a$$

Applying the above equation to calculate the probability of cross T cell epitope between two strains s and v ,

$$E(s, v) = \sum_{i \in s, j \in v} \sum_{a \in A} S(i, j)_a$$

To further compute the shared T cell epitope content between (conserved) two strains, i.e., EpiCC score, the following equation is applied:

$$P(\text{shared}) = \frac{2 \cdot E(s, v)}{E(s, s) + E(v, v)}$$

Area under the curve (AUC) computation. Given that the complexity of multiple comparisons which were done according to years and subtypes, AUC calculation is applied to represent T cell epitope conservation of a subtype in a year. EpiCC scores that were calculated for MEpiV and swine IAV sequences were plotted in radar form (a line plot that is on circular orientation). The area under the radar curve (a numerical integral) was computed by combining spline interpolation and integration with the formula shown below:

$$AUC = \int_a^b f(x) dx$$

The higher the AUC value, the more T cell epitopes against MEpiV were conserved. Normalization of AUC values with respect to the baseline score of MEpiV vaccine was performed as the number of sequences varied across the years. This enabled direct comparison of epitope content conservation across years.

Phylogenetic analysis. T cell epitope conservation was mapped onto phylogenetic tree to correlate T cell epitope conservation with genetic evolution of swine IAV. Phylogenetic trees inferred from the Maximum Likelihood (ML) were constructed based on HA protein (H1 and H3 subtypes) of circulating swine IAV strains with RAXML.v8 using the GTR-GAMMA nucleotide substitution model. Both phylogeny trees were rooted with midpoint. MEpiV vaccine epitopes were evaluated against H1 and H3 tree tips using `ggtree` package version 2.2.4 in R [26].

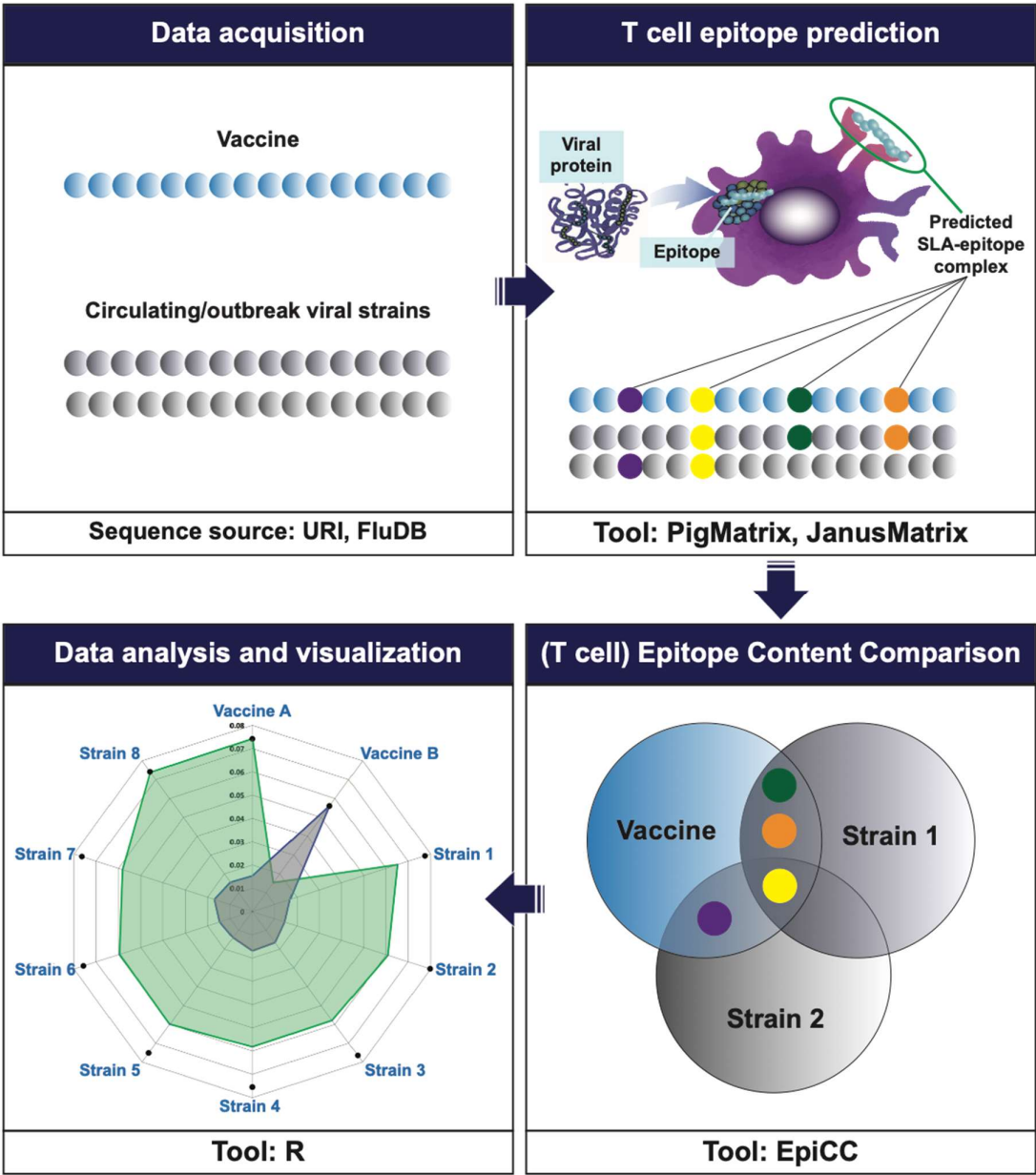


Figure 1. Workflow for typical EpiCC analysis. The vaccine sequence of interest (here, MEpiV) and circulating pathogen strains (swine IAV in this example) are retrieved and pre-processed prior to performing the EpiCC analysis. T cell epitopes are identified in the vaccine and circulating strains (colored beads) using EpiMatrix (for HLA restricted human T cell epitopes) or PigMatrix (for SLA-restricted T cell epitopes). Once all epitopes are identified, comparison and quantification of the T cell epitopes is performed using EpiCC. An overall EpiCC score (area under the curve) is calculated for the combined class I and II epitopes for each vaccine/strain comparison. Greater AUC scores indicate higher numbers of conserved T cell epitopes. EpiCC scores can be compared and contrasted for selected vaccines (here, MEpiV versus seasonal whole inactivated swine IAV vaccines).

3. Results

3.1. Swine IAV dataset from 2013 to 2017

The goal of this study was to determine whether a vaccine designed in 2013 may continue to provide CMI boosting as was illustrated in 2019 [11]. The MEpiV vaccine contains 28 class I and 20

class II T cell epitopes and was produced as a plasmid DNA vaccine and tested in 2015 [8]. Circulating swine IAV whole genome sequences of three major subtypes (H1N1, H1N2 and H3N2) from 2013 to 2017 were computationally screened in the same stepwise process to evaluate their T cell epitope content in an epitope to epitope comparison to circulating strains (Figure 2). A total of 1,272 whole genome swine influenza A sequences were analyzed, comprising 409 (32.2%) H1N1, 388 (30.5%) H1N2 and 475 (37.3%) H3N2 sequences. The highest number of sequences available was from 2016 (407 sequences; 32.0% of the total), while the lowest number was for 2014 (133 sequences; 10.5% of the total).

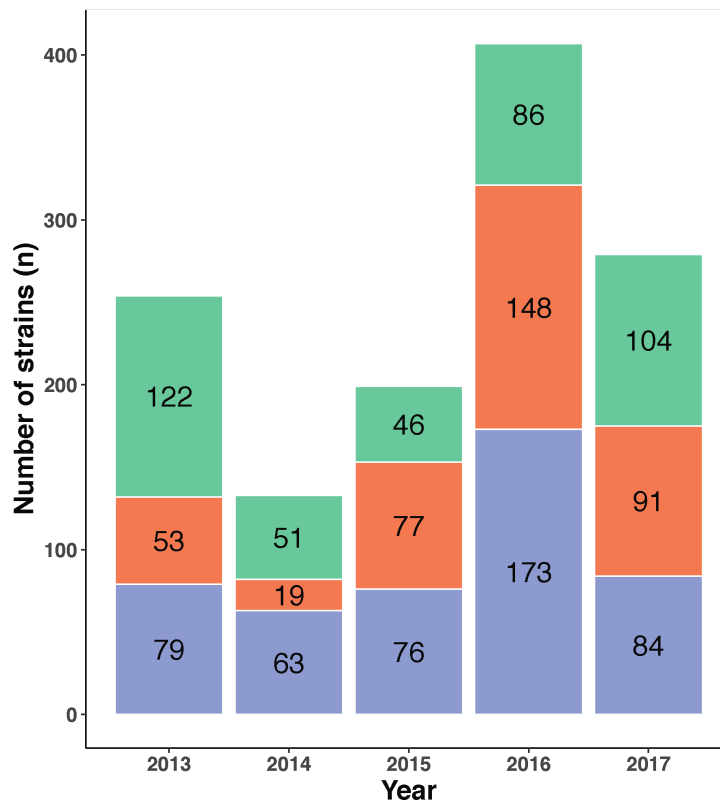


Figure 2. Swine IAV genome sequences from the H1N1, H1N2 and H3N2 subtypes from 2013-2017 included in this analysis. Color-coded stacked bar chart represents the three subtypes; each stacked component shows the number of strains per subtype for that year.

Subtype ■ H1N1 ■ H1N2 ■ H3N2

3.2. T cell epitope content comparison (EpiCC) of swine MEpiV vaccine against H1N1, H1N2 and H3N2 circulating swine IAV

In order to determine conservation of MEpiV vaccine epitopes among the three circulating swine IAV subtypes, we applied EpiCC to facilitate pairwise comparison of protein sequences. This sequence comparison method is based on an immunological property, potential T cell immunogenicity, rather than sequence identity. Shared (conserved) T cell epitopes between the vaccine target and the circulating swine IAV strains were assessed.

Higher EpiCC scores are thought to be associated with greater protection by vaccines against challenge strains [19]. For MEpiV vaccine class I epitopes, the highest EpiCC score is found for H1N1 swine IAVs (EpiCC score of 0.0256 with 98.5% conservation when normalized to MEpiV baseline), and the lowest for H3N2 (EpiCC score of 0.0100 with 38.5% conservation when normalized to MEpiV baseline). Interestingly, on average, EpiCC scores of MEpiV vaccine class II epitopes for all subtypes is 14.3% higher than scores of class I epitopes. The average range difference (in percentage) of class II

EpiCC scores for all three subtypes is 25.6%, while for class I EpiCC scores it is 28.6%. The range difference of class II EpiCC scores is 11.7% smaller than the range difference of class I, indicating that conservation of MEpiV class II epitopes was consistent in all circulating swine IAV that were analyzed. Detailed information for each of the circulating strains and their respective EpiCC scores are tabulated in Table S2.

While detailed lists of EpiCC scores are informative, we also used radar plots to visualize the EpiCC scores. Radar plots were constructed to describe degree of conservation of MEpiV vaccine class I and II T cell epitopes in the three prevalent swine flu subtypes (Figure 3) and the area under the curve for the EpiCC scores (AUC, outlined in color in Figure 3) was used to quantify and compare the T cell epitope conservation between the vaccine and circulating swine IAV each year. As shown in Figure 3 and Figure S1, the AUC described by the EpiCC scores is greater for the MEpiV vaccine against H1N1, than the AUC for H3N2 and H1N2 circulating strains. Thus, the vaccine is predicted to be effective against all circulating H1N1 strains in 2013 to 2017. The MEpiV vaccine is predicted to drive broad CD4 immune response based on data published in Gutierrez et al. and Hewitt et al.

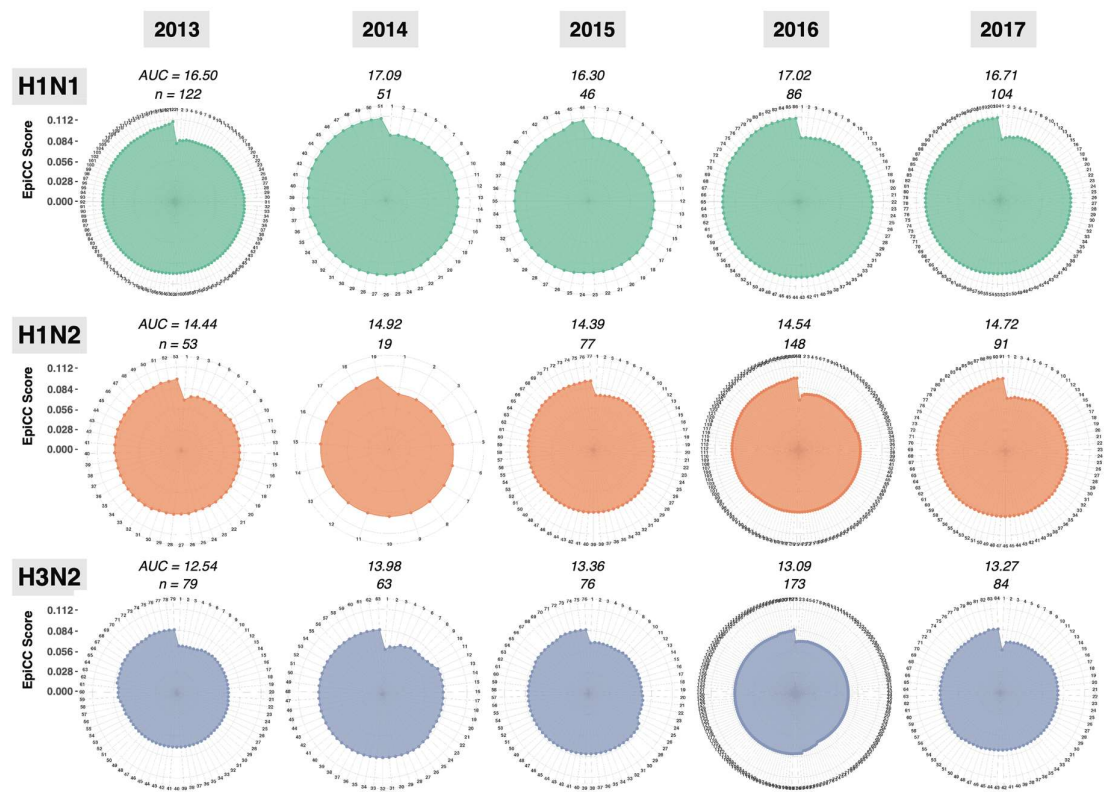


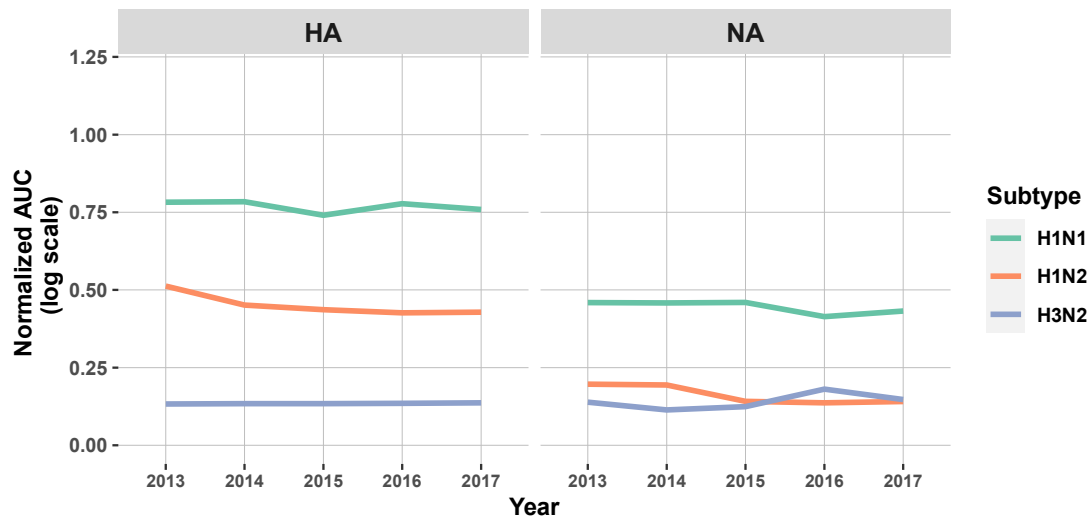
Figure 3. Radar plots enable quantitative analysis of the degree of T cell epitope conservation between the conserved epitopes from all of the IAV proteins contained in a vaccine (here, MEpiV) and the epitopes from all of the IAV proteins contained in whole genome circulating strains for each year. The EpiCC score describing the T cell epitope conservation between the vaccine (MEpiV) against each swine IAV circulating strain is plotted on the radiating axes of radar plot for each year, for a period of five years, left to right. Circulating IAV strains were sorted from lowest to highest EpiCC scores. Radar plots for class II EpiCC scores are shown here and radar plots for class I are provided in supplemental data (Figure S1).

Computing the AUC facilitates qualitative comparison of the vaccine against circulating strains over time. As expected, when considering the MEpiV computer-designed vaccine epitopes, the overall EpiCC scores, compared to circulating viral strains, did not change very much over time. The overall conservation was maintained for all three viral subtypes, although the total EpiCC scores

were lower for H1N2 and H3N2 strains. Class II T cell epitopes were 80.8% more conserved on average, as compared to class I in all subtypes (Figure S1). We visualized these data on the individual antigen level in biaxial plots with the x-axis representing time and the y-axis representing the AUC for vaccine against circulating strains for that year (Figure 4). For the HA antigen, there was 79.5% conservation of MEpiV vaccine (both class I and II HA epitopes) in H1N1 over multiple years whereas HA epitope conservation in H1N2 and H3N2 were 51.7% and 8.6%, respectively.

The overall conservation of NA class I and II epitopes in H1N1 was 45.0% while conservation of vaccine epitopes in H1N2 and H3N2 strains was lower at 10.3% and 9.0%, respectively. Conservation in H1N2 and H3N2 for surface antigens was relatively low compared to H1N1, due to complete lack of conservation (AUC of zero) for H3 and N2 epitopes in the MEpiV vaccine. Internal antigen epitopes were also well conserved across all subtypes (Figure S2), suggesting internal proteins might contribute to vaccine efficacy. While the original MEpiV epitopes were selected from seven representative swine influenza strains, this finding suggests that vaccine epitopes that are highly conserved in one set of sequences for a given year may still be relevant and provide cross-protective immunity in years that follow.

(A) Class I T cell epitopes of surface antigens HA and NA



(B) Class II T cell epitopes HA and NA

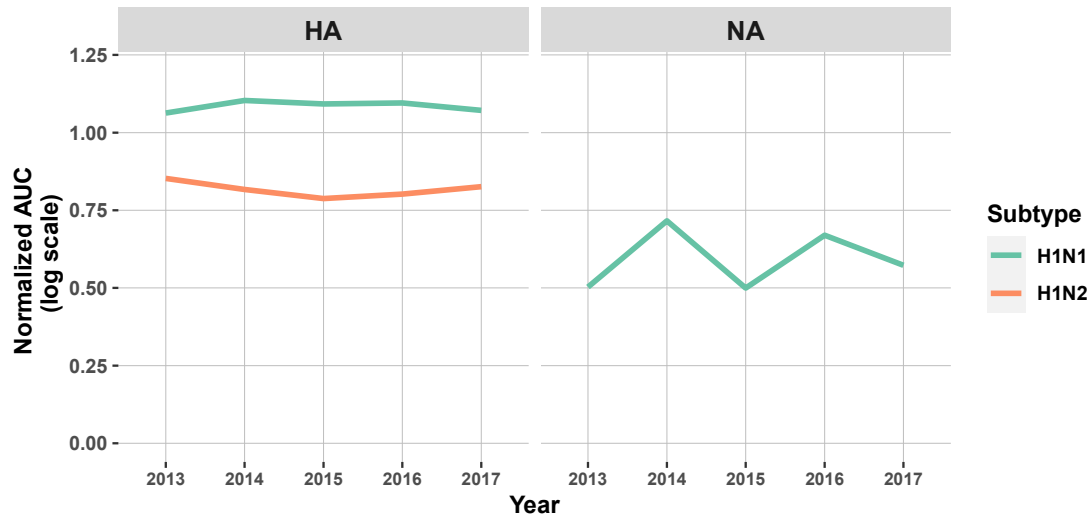


Figure 4. Line plots showed the normalized AUC for the comparison of MEpiV vaccine epitopes to epitopes found in circulating IAV strains for surface antigens HA and NA, by subtypes and by year. The AUC is shown on a normalized scale to enable direct comparison by antigens, subtypes, years and T cell epitopes classes. Lines represent H1N2 (NA) and H3N2 (HA and NA) were removed in (B) as they showed no conservation. Similar line plots (different y-axis scaling) for the internal antigens are shown in Figure S2.

3.3.2. T cell epitope conservation analysis of individual epitopes using JanusMatrix (JMX)

The EpiCC tool gives an overall score for the combined epitope content, rather than assessing and reporting on each epitope in a vaccine. Since the MEpiV is composed of distinct T cell epitopes, we wished to determine the conservation of each epitope over time, and therefore we performed an additional epitope-by-epitope analysis using JMX comparing the vaccine epitopes with their homolog in circulating strains. In this case, JMX searches the circulating swine IAV strains for 9-mers with the same TCR-facing amino acids as those of the input class I and II MEpiV vaccine epitopes [18]. JMX homology score was calculated for every input MEpiV vaccine epitope that appears “homologous” to a given TCR, even though there may be minor variations in the MHC binding residues, as long as the peptide would still be predicted to bind to the same MHC.

While performing the JMX analysis to compare vaccine epitopes to circulating strain epitopes matched for binding to the same MHC and identical at the TCR-face, we were able to identify the specific TCR-homologous 9-mers in circulating swine IAV strains. We applied JMX homology scores to further examine the level of conservation of individual T cell epitopes in every swine IAV subtype and quantify the overall conservation (Table 1).

Doing so, we were able to identify the most highly conserved T cell epitopes. Among 28 class I peptides, 16 of the peptides were more than 80% conserved in the three-circulating swine IAVsubtypes throughout the five year period (Table 1A). Only two surface epitopes from NA were conserved and were N1-specific. Most of the highly conserved peptides were from internal antigens: PB2 (GTEKLTITY), PB1 (VSDGGPNLY, DTVNRTHQY), PA (QVSRPMFLY), NP (AFDERRNKY, CTELKLSY, ASQGTGRSY, KSCINRCFY, DTVHTRTPY), and M1 (SLLTEVETY, LTEVETYVL, DLLENLQAY, LASCGLIY, LASCGLIY, NTDLEALME). The two most conserved peptides were SLLTEVETY, LTEVETYVL from the M1 protein. These peptides were found in all 1272 IAV strains. Interestingly, the least conserved peptides (GAKEVALSY, NMDKAVKLY) are also from M1, with conservation less than 3% in all subtypes and only being observed in 2013.

In addition, 10 out of 20 class II peptides were highly conserved (> 80%) in circulating swine IAV strains (Table 1B). None of these highly conserved peptides was found in HA and NA, rather they were found in internal antigens such as PB1 (MMGMFNLSTVLGVSI, YRYGFVANFSMELPSFGVSG), PA (EVHIYYLEKANKIKSEKTHIHIF, RSKFLLMDALKLSIEDP), NP (IEDLIFLARSALILRGSAVHKSLCP), M1 (TRQMVHAMRTIGTHPSSA, TYVLSIIPSGPLKAEIA QRLESV, SCMGLIYNRMGTVTTEAAGLVC), and NS2 (FEQITFMQALQLLLEVE, FQDILMRMSKMQL GSSSE). This suggests that epitopes from the internal antigens are well-conserved across strains and over time, and may contribute to vaccine efficacy.

Table 1. Total class I (A) and class II (B) peptides found in circulating IAV strains over a five-year period, sorted from the greatest to the lowest conservation. Antigens are sorted according to viral surface antigens (HA and NA), followed by internal antigens. Shaded rows represent MEpiV epitopes that show conservation equal to or greater than 80% in swine circulating IAV strains.

(A)

No.	Antigen	Class I epitopes	JMX homology score (% of conservation)			Average conservation (%)
			H1N1	H1N2	H3N2	
1	HA	GMVDGWYGY	401.7 (98.2)	384.8 (99.2)	240.7 (50.7)	79.0
2	HA	GMIDGWYGY	401.7 (98.2)	384.8 (99.2)	240.7 (50.7)	79.0
3	HA	SVKNGTYDY	402.8 (98.5)	308.0 (79.4)	0.5 (0.1)	9.2
4	HA	RIYQILAIY	392.8 (96.0)	57.0 (14.7)	0	37.6
5	HA	NADTLCIGY	375.0 (91.7)	22.0 (5.7)	0	22.9
6	HA	TSADQQSLY	352.0 (86.1)	17.0 (4.4)	0	19.5
7	HA	LSTASSWSY	306.5 (74.9)	16.5 (4.3)	0	17.9
8	HA	ITIGKCPKY	58.8 (14.4)	3.5 (0.9)	0	3.6
9	NA	KSCINRCFY	0	384.0 (99.0)	474.0 (99.8)	99.4
10	NA	DTVHDRTPY	0	371.3 (95.7)	468.0 (98.2)	96.9
11	NA	GTIKDRSPY	322.25 (78.8)	0	0	78.8
12	NA	EMNAPNYHY	337.29 (82.5)	0	0	82.5
13	NA	ELDAPNYHY	381.86 (93.4)	0	0	93.4
14	NA	EICPKLAEY	0	98.4 (25.4)	120.6 (25.4)	25.4
15	PB2	GTEKLITY	405.7 (99.2)	379.7 (97.9)	454.7 (95.7)	97.6
16	PB1	VSDGGPNLY	408.4 (99.9)	386.2 (99.5)	472.0 (99.4)	99.6
17	PB1	DTVNRTHQY	409.0 (100.0)	386.7 (99.7)	468.0 (98.5)	99.4
18	PA	QVSRPMFLY	400.6 (97.9)	379.6 (97.8)	433.0 (91.2)	95.7
19	NP	AFDERRNKY	407.8 (99.7)	386.5 (99.6)	471.8 (99.3)	99.5
20	NP	CTELKLSY	406.0 (99.3)	383.5 (98.8)	472.0 (99.4)	99.2
21	NP	ASQGTKRSY	400.0 (97.8)	369.0 (95.1)	464.0 (97.7)	96.9
22	M1	SLLTEVETY	409 (100.0)	388 (100.0)	475.0 (100.0)	100.0
23	M1	LTETVETYVL	409 (100.0)	388 (100.0)	475.0 (100.0)	100.0
24	M1	DLLENLQAY	407 (99.5)	387 (99.7)	468.0 (98.5)	99.2
25	M1	LASCMGLIY	399 (97.6)	388 (100.0)	473.0 (99.6)	99.1
26	M1	NTDLEALME	399 (97.6)	366 (94.3)	463.0 (97.5)	96.5
27	M1	NMDKAVKLY	11 (2.7)	10 (2.6)	16.0 (3.4)	2.9
28	M1	GAKEVALSY	12 (2.9)	9 (2.3)	13.0 (2.7)	2.6

(B)

No.	Antigen	Class II epitopes	JMX homology score (% of conservation)			Average conservation (%)
			H1N1	H1N2	H3N2	
1	HA	YEELREQLSSVSSFER	392.6 (96.0)	365.3 (94.1)	0	63.4
2	HA	STRIYQILAIYSTVASSLVLV	393.1 (96.1)	253.4 (65.3)	0	53.8
3	HA	GDKITFEATGNLVVPRY	348.2 (85.1)	56.4 (14.5)	0	33.2
4	HA	VPRYAFAMERNAGSGIIIS	13.0 (3.2)	1.1 (0.3)	0	1.2
5	NA	CRTEFLTQGALLNDKH	408.4 (99.9)	0	0	33.3

6	NA	SVVSVKLAGNSSLCPV	102.9 (25.2)	0	0	8.4
7	NA	NQTYVNISNTNFAAGQSVVSVKL	66.0 (16.1)	0	0	5.4
8	NA	MANLILQIGNIISIWISHS	62.1 (15.2)	0	0	5.1
9	PB1	MMGMFNMMLSTVLGVSI	409.0 (100.0)	387.4 (99.8)	474.9 (100.0)	99.9
10	PB1	YRYGFVANFSMELPSFGVSG	409.0 (100.0)	388.0 (100.0)	474.5 (100.0)	100.0
11	PA	EVHIYYLEKANKIKSEKTHIHIF	406.3 (99.3)	386.1 (99.5)	472.9 (99.6)	99.5
12	PA	RSKFLLDALKLSIEDP	405.9 (99.2)	381.0 (98.2)	474.7 (99.9)	99.1
13	NP	IEDLIFLARSALILRGSAVHKSCLP	400.3 (97.9)	311.3 (80.2)	454.3 (95.6)	91.2
14	NP	TRGVQIASNENVETMDSNTLELR	346.5 (84.7)	243.8 (62.8)	268.5 (56.5)	68.0
15	NP	IDPFKLLQNSQVVSIMLRP	343.4 (84.0)	270.6 (69.7)	296.9 (62.5)	72.1
16	M1	TRQMVHAMRTIGTHPSSSA	398.6 (97.5)	380.8 (98.1)	462.1 (97.3)	97.6
17	M1	SCMGLIYNRMGTVTTEAAFGVLC	399.3 (97.6)	382.0 (98.5)	462.7 (97.4)	97.8
18	M1	TYVLSIIPSGPLKAEIAQRLESV	395.3 (96.7)	367.2 (94.7)	465.5 (98.0)	96.5
19	NS2	FEQITFMQALQLLLEVE	407.6 (99.7)	384.9 (99.2)	466.6 (98.2)	99.0
20	NS2	FQDILMRMSKMQLGSSSE	364.9 (89.2)	327.2 (84.3)	358.9 (75.6)	83.0

We then used this epitope matching information generated from JMX analysis jointly with HA phylogeny trees to visualize the distribution of MEpiV vaccine class I and II epitopes (Figure 5). Epitopes from both classess were well conserved in most internal proteins, indicated by the presence of small bars adjacent to the tips of the respective HA phylogeny tree. Epitopes in the external proteins such as HA and NA are subtype-specific, demonstrating MEpiV vaccine consists of H1, N1 and N2-specific epitopes. A big blank under HA for the H3 phylogeny tree showing almost absence of H3 epitopes in MEpiV vaccine. Interestingly, although there are subtype-specific epitopes, we would expect H1N1 and H1N2 IAV strains have shared conservation in HA epitopes, however, H1-specific eptiopes are only found conserved in 47 H1N2 swine IAV strains that are of the same clade as H1N1 IAV strains. Six out of eight class I and half of four class II HA epitopes were absent in H1N2 swine IAV subtype.

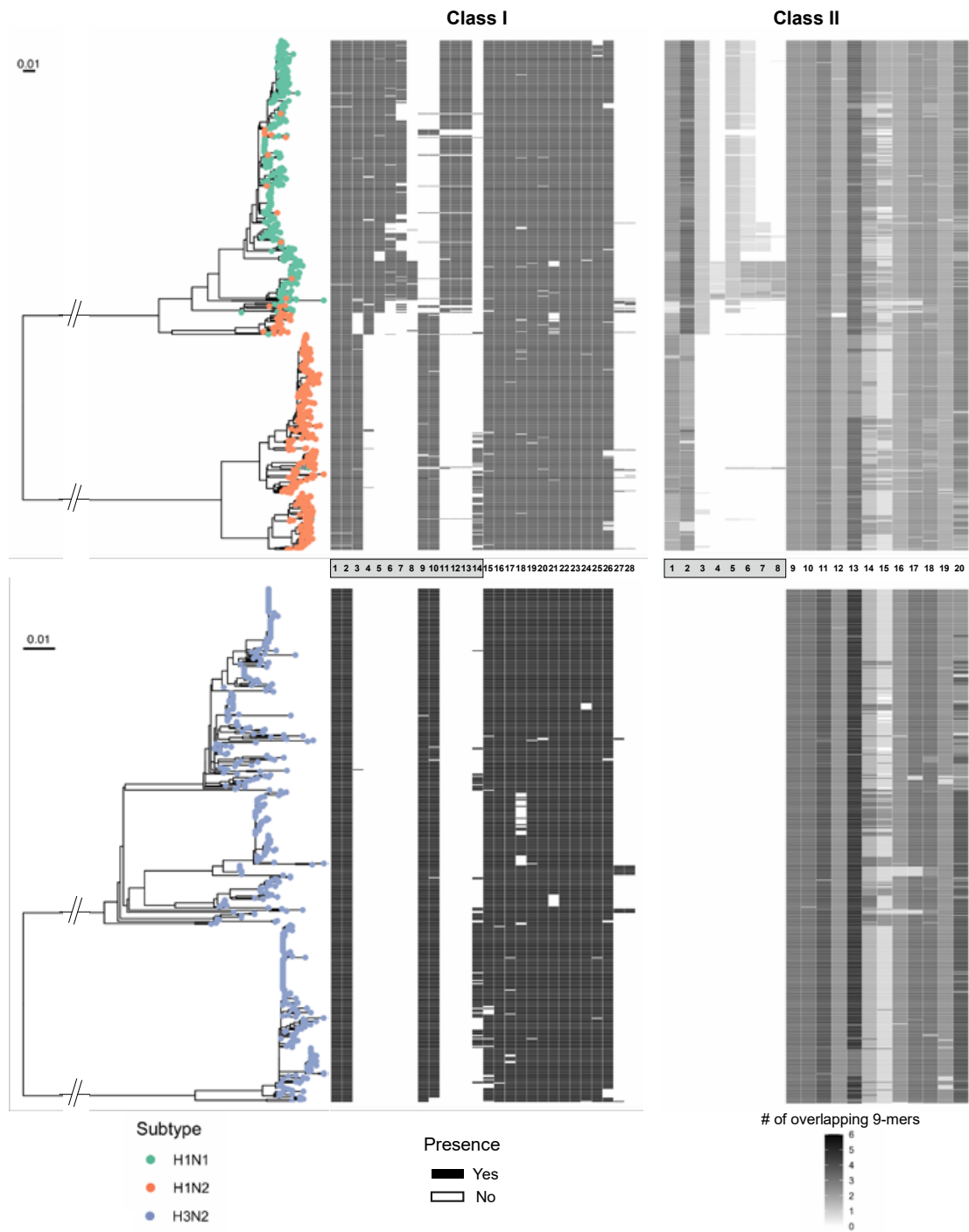


Figure 5. Phylogenetic-epitope mapping for class I and class II of MEpiV vaccine epitopes. HA of the circulating swine IAV strains were color-coded by subtypes and each of the strains were shown as tips on the tree. MEpiV epitopes were listed in the central panel sorted by external (grey box) and internal proteins arrangement (numberings refer to Table 1). The small black and white bars mapped adjacent to the phylogeny tree implying the presence or absence of respective MEpiV vaccine epitopes in these circulating swine IAV strains.

3.4. Strains identification for conserved peptides

The epitope to epitope-based comparison can also be used to identify strains that have the most, or the least conserved T cell epitopes (Table S2), which may be important when selecting strains for a recombinant or inactivated whole antigen vaccine. Forty-four IAV sequences were shown to be highly conserved against the MEpiV prototype vaccine, with conservation at 75%. The majority of the sequences (42/44) belong to the H1N1 subtype, while two belong to the H1N2 subtype. In contrast, four swine IAV sequences had very few epitopes conserved with the prototype vaccine (46.4%); all of these strains were H3N2 subtypes. This is expected as most of the T cell epitopes included in MEpiV were HA H1-specific, and conservation across subtypes is not optimal, indicating that truly universal vaccines must include epitopes from more than one subtype.

This study demonstrates how EpiCC and JMX can be applied in complement for surveillance and analysis of epitope evolution and/or escape. One of the direct applications of the EpiCC program is to enable the selection of challenge IAV strains for vaccine studies. Furthermore, this work also serves a retrospective analysis that provides a baseline strain coverage estimate for MEpiV, but it can easily be applied to other (new or old) vaccines against large numbers of new viruses.

3.5. Comparison of MEpiV and commercial swine flu vaccine

Immunity induced by inactivated virus vaccines usually wanes over time when it is no longer a close match to circulating strains. To further investigate whether the T cell epitope conservation in a vaccine that was computationally designed to contain such epitopes was advantageous as compared to commercial swine flu vaccines, we compared AUC computed from EpiCC analysis for HA antigens of the MEpiV vaccine and the HA found in commercial vaccine which comprises four HA vaccine strains of the major swine IAV seasonal subtypes, one H1N1; one H1N2; two H3N2. The H1 components of this commercial vaccines were included since 2011 while the H3 components were introduced in 2016.

A year-to-year comparison was made relative to 2013 for HA antigens of all vaccine strains except for the H3 components of the commercial vaccine strains that were introduced in 2016. Changes in the conservation of the vaccines against the baseline year were calculated as a ratio, meaning that a score of 1.00 would indicate no change in T cell epitope conservation (in AUC values); greater than 1.00 indicates increasing T cell epitope conservation relative to 2013, and a ratio less than 1.00 implies loss of T cell epitope conservation. The ratio of T cell epitope content (class I and II) for MEpiV over time, remains consistent or increases (except for H1N2) (Figure 6). Specifically, the H3 HA class I epitopes in MEpiV vaccine showed a gradual increase of conservation in circulating swine IAV strains. In contrast, the ratio of conservation for the H3 conventional vaccine strains (FSXP.NC and FSXP.MN) decline over time. The same trend for FSXP.NC and FSXP.MN were observed in class II, however, there was no class II epitope conservation change for MEpiV vaccine as there were no H3-specific class II epitopes selected for MEpiV vaccine. This result is consistent with EpiCC and JMX analyses shown above.

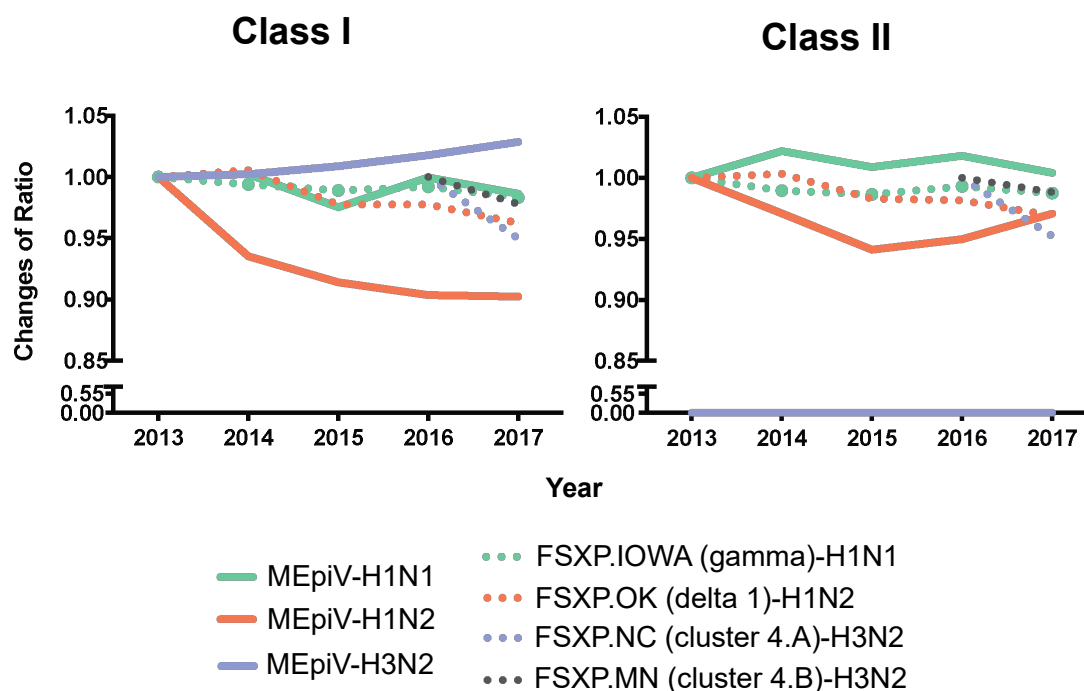


Figure 6. MEpiV compared to commercial seasonal vaccine. EpiCC analysis for the HA epitopes in MEpiV and HA from commercial (inactivated, whole) vaccines were calculated and then normalized to the EpiCC AUC determined for circulating strains for the vaccine in 2013, to show changes in AUC over time. Solid lines represent HA antigen of MEpiV, epitope-based vaccine and dotted lines illustrate HA components from commercial swine vaccine.

4. Discussion

In general, vaccine efficacy assessment methods are lacking for swine IAV. More specifically, in lieu of challenge studies, there is no method available for evaluating new vaccines against circulating strains for cross-protection by T cell epitopes. Here, we used the EpiCC tool to approximate the potential T cell epitope cross-protection between MEpiV and circulating strains. In previously published studies, we established a threshold of cross-conservative epitope protection, using EpiCC to compare one vaccine against IAV strains circulating in one year [20]. We have also demonstrated the utility of EpiCC tool applied for another pathogen, Porcine circovirus 2 (PCV2) in a study evaluating multiple vaccines against circulating PCV2 strains [27]. In this study, we demonstrate how EpiCC can be used for longitudinal analysis against evolving strains circulating in swine populations.

The current analysis applies the EpiCC tool to a computationally designed T cell epitope vaccine and compares the vaccine with circulating strains over a five-year period. Having established the longitudinal conservation of the H1N1 T cell epitopes in the subunit vaccine, we then compared the 5-year trajectory of the epitope vaccine with that of a typical commercial swine IAV vaccine. The MEpiV retained conservation of T cell epitope content over time, while antigenic 'drift' was evident for the commercial vaccine, resulting in lower EpiCC scores for the epitopes contained in the HA antigen over time, as expected. Consistency of the area under the curve (AUC) over years (for the MEpiV) suggests that the T cell epitopes in the prototype vaccine could reliably drive robust immune responses in swine regardless of drift, and that a conserved epitope-driven vaccine may be a valuable adjunct to vaccination with whole, inactivated seasonal vaccine as was shown in Hewitt et al.

Comparing T cell epitope conservation can contribute to assessing the projected efficacy of a vaccine. This study illustrates how EpiCC might be applied to evaluate several different vaccines,

and to select the best vaccine strain (based on T cell epitope conservation) for any given year. This is as relevant for IAV as it may be relevant for other emerging viruses such as COVID-19.

The analysis also demonstrates the use of JMX, a novel tool that searches for conserved T cell epitopes using TCR facing residues. JMX may make more accurate comparisons between T cell epitopes contained in vaccines as compared to circulating strains over time. By quantifying conservation using JMX, we are also able to examine which T cell epitopes are conserved in which strains of IAV. This type of analysis may be useful for the selection of challenge strains in vaccine studies. Not surprisingly, epitopes from M1 and PB1 proteins were better conserved with circulating strain epitopes over the 5-year period studied in this example, and as expected, epitopes from HA and NA protein were much less well conserved.

Compared to the commercial whole antigen killed vaccine, MEpiV T cell epitopes were highly conserved over time. This finding is particularly relevant for influenza, since cross-reactive antibodies may not be present when influenza strains shift, rather than drift [27]. Experts in the field have advocated for the development of ‘universal influenza vaccines’ that can boost immune responses in the absence of antibody cross-reactivity for this reason. The fact that lower levels of conservation were observed for H1N2 and H3N2 over time suggests that conserved epitope-based vaccines should be designed for each IAV subtype. We have explored the use of MEpiV-type vaccines given by heterologous prime-boost with a commercial swine influenza vaccine (which contains whole HA antigen) and found increased immunogenicity by priming with the MEpiV vaccine over homologous commercial vaccine prime-boost, equivalent body temperature control one day after pH1N1 challenge, and reduced lung lesions and influenza antigen, as illustrated in Hewitt et al. Reducing the overall viral burden and increasing the average daily gain, distributed across large populations of swine, may prove cost-effective for pork producers. One application of “universal” T cell epitope-based vaccines being explored in humans is to combine them with seasonal vaccines, a topic which might also be of interest to the animal health community [28–30].

In conclusion, we applied the EpiCC tool to enable research on the impact of conserved T cell epitope-based vaccines as compared to whole antigen vaccines for influenza and other pathogens. As shown here, the EpiCC tool enables a comparison between vaccines and circulating field strains, but it is also useful for identifying whether existing vaccines might have efficacy (at the T cell epitope level) against emerging infection. The EpiCC tool is likely to be useful for application to other viral populations such as emerging G4 influenza, African Swine Fever, as well as to human influenza strains versus candidate vaccines.

6. Patents

The MEpiV vaccine design is intellectual property that is shared between EpiVax and University of Rhode Island.

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1, Figure S1: Radar plots showing degree of class I T cell epitope conservation between the conserved epitopes in the MEpiV for the HA antigen and epitopes contained in HA from circulating strains for each year, Figure S2: Line plots showing the degree of conservation by antigens, subtypes and years, Table S1: MEpiV vaccine class I and II peptides, Table S2: EpiCC scores of all swine IAV circulating strains, Table S3: List of swine IAV circulating strains that have the most and the least conserved class I (A) and II (B) epitopes.

Author Contributions: For conceptualization, A.S.D.G., L.M. and S.T.; methodology, S.T., A.H.G., A.S.D.G., L.M. and J.B.; formal analysis, S.T.; investigation, S.T., A.S.D.G. and L.M.; resources, P.C.G.; writing—original draft preparation, S.T.; writing—review and editing, A.S.D.G., L.M., S.T., A.H.G., P.C.G., T.O. and J.B.; visualization, S.T.; supervision, A.S.D.G., L.M. and J.B.

Funding: This research received no external funding.

Acknowledgments: We acknowledge Will Martin and Matt Ardito for their technical input and support, Dr. Douglas S. Pearce and Dr. Lucina Galina for their input and critical review.

Conflicts of Interest: A.S.D.G. is a senior officer and shareholder, and L.M. and A.H.G are employees of EpiVax, Inc., a privately owned biotechnology company located in Providence, RI. These authors acknowledge that there is a potential conflict of interest related to their relationship with EpiVax and attest that the work contained in this research report is free of any bias that might be associated with the commercial goals of the company. The iVAX toolkit is currently available for use by commercial developers by subscription or for specific projects under a fee-for-service arrangement. Academic researchers are invited to contact the authors at University of Georgia (UGA), or the technology transfer office at the UGA School of Veterinary Medicine, for access to the iVAX Toolkit for research purposes.

References

- [1] J. C. Pedersen, "Hemagglutination-inhibition assay for influenza virus subtype identification and the detection and quantitation of serum antibodies to influenza virus," *Methods Mol. Biol.*, vol. 1161, 2014, doi: 10.1007/978-1-4939-0758-8_2.
- [2] A. Grifoni *et al.*, "Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals," *Cell*, vol. 181, no. 7, 2020, doi: 10.1016/j.cell.2020.05.015.
- [3] S. Sridhar *et al.*, "Cellular immune correlates of protection against symptomatic pandemic influenza," *Nat. Med.*, vol. 19, no. 10, pp. 1305–1312, 2013, doi: 10.1038/nm.3350.
- [4] T. M. Wilkinson *et al.*, "Preexisting influenza-specific CD4 + T cells correlate with disease protection against influenza challenge in humans," *Nat. Med.*, vol. 18, no. 2, 2012, doi: 10.1038/nm.2612.
- [5] J. Zhao, J. Zhao, and S. Perlman, "T Cell Responses Are Required for Protection from Clinical Disease and for Virus Clearance in Severe Acute Respiratory Syndrome Coronavirus-Infected Mice," *J. Virol.*, vol. 84, no. 18, 2010, doi: 10.1128/jvi.01049-10.
- [6] J. L. Sanchez-Trincado, M. Gomez-Perosanz, and P. A. Reche, "Fundamentals and Methods for T- and B-Cell Epitope Prediction," *Journal of Immunology Research*, vol. 2017, 2017, doi: 10.1155/2017/2680160.
- [7] J. B. Ulmer *et al.*, "Protective CD4+ and CD8+ T Cells against Influenza Virus Induced by Vaccination with Nucleoprotein DNA," *J. Virol.*, 1998, doi: 10.1128/jvi.72.7.5648-5653.1998.
- [8] T. J. Powell *et al.*, "Priming with Cold-Adapted Influenza A Does Not Prevent Infection but Elicits Long-Lived Protection against Supralethal Challenge with Heterosubtypic Virus," *J. Immunol.*, 2007, doi: 10.4049/jimmunol.178.2.1030.
- [9] T. Wu *et al.*, "Quantification of epitope abundance reveals the effect of direct and cross-presentation on influenza CTL responses," *Nat. Commun.*, 2019, doi: 10.1038/s41467-019-10661-8.
- [10] L. Garcia-Garcia *et al.*, "Partial protection of seasonal trivalent inactivated vaccine against novel pandemic influenza A/H1N1 2009: Case-control study in Mexico City," *BMJ*, vol. 339,

no. 7725, p. 847, 2009, doi: 10.1136/bmj.b3928.

- [11] A. S. De Groot, M. Ardito, E. M. McClaine, L. Moise, and W. D. Martin, "Immunoinformatic comparison of T-cell epitopes contained in novel swine-origin influenza A (H1N1) virus with epitopes in 2008-2009 conventional influenza vaccine," *Vaccine*, vol. 27, no. 42, pp. 5740–5747, 2009, doi: 10.1016/j.vaccine.2009.07.040.
- [12] A. Vincent *et al.*, "Review of Influenza A Virus in Swine Worldwide: A Call for Increased Surveillance and Research," *Zoonoses and Public Health*, vol. 61, no. 1. pp. 4–17, 2014, doi: 10.1111/zph.12049.
- [13] L. He, A. S. De Groot, A. H. Gutierrez, W. D. Martin, L. Moise, and C. Bailey-Kellogg, "Integrated assessment of predicted MHC binding and cross-conservation with self reveals patterns of viral camouflage," *BMC Bioinformatics*, vol. 15, 2014, doi: 10.1186/1471-2105-15-S4-S1.
- [14] R. Liu *et al.*, "H7N9 T-cell Epitopes that mimic human sequences are less immunogenic and may induce Treg-mediated tolerance," *Hum. Vaccines Immunother.*, vol. 11, no. 9, pp. 2241–2252, 2015, doi: 10.1080/21645515.2015.1052197.
- [15] A. H. Gutiérrez, W. D. Martin, C. Bailey-Kellogg, F. Terry, L. Moise, and A. S. De Groot, "Development and validation of an epitope prediction tool for swine (PigMatrix) based on the pocket profile method," *BMC Bioinformatics*, vol. 16, no. 1, 2015, doi: 10.1186/s12859-015-0724-8.
- [16] A. L. Vincent, K. M. Lager, and T. K. Anderson, "A brief introduction to influenza a virus in swine," *Methods Mol. Biol.*, 2014, doi: 10.1007/978-1-4939-0758-8_20.
- [17] L. Moise *et al.*, "Ivax: An integrated toolkit for the selection and optimization of antigens and the design of epitope-driven vaccines," *Hum. Vaccines Immunother.*, vol. 11, no. 9, pp. 2312–2321, 2015, doi: 10.1080/21645515.2015.1061159.
- [18] L. Moise *et al.*, "The two-faced T cell epitope: Examining the host-microbe interface with JanusMatrix," in *Human Vaccines and Immunotherapeutics*, 2013, vol. 9, no. 7, pp. 1577–1586, doi: 10.4161/hv.24615.
- [19] A. H. Gutiérrez *et al.*, "T-cell epitope content comparison (EpiCC) of swine H1 influenza A virus hemagglutinin," *Influenza Other Respi. Viruses*, vol. 11, no. 6, pp. 531–542, 2017, doi: 10.1111/irv.12513.
- [20] A. H. Gutiérrez *et al.*, "In vivo validation of predicted and conserved T cell epitopes in a swine influenza model," *PLoS One*, vol. 11, no. 7, 2016, doi: 10.1371/journal.pone.0159237.
- [21] J. S. Hewitt *et al.*, "A prime-boost concept using a T-cell epitope-driven DNA vaccine followed by a whole virus vaccine effectively protected pigs in the pandemic H1N1 pig challenge model," *Vaccine*, 2019, doi: 10.1016/j.vaccine.2019.06.044.

- [22] Y. Zhang *et al.*, "Influenza Research Database: An integrated bioinformatics resource for influenza virus research," *Nucleic Acids Res.*, vol. 45, no. D1, pp. D466–D474, 2017, doi: 10.1093/nar/gkw857.
- [23] T. Sturniolo *et al.*, "Generation of tissue-specific and promiscuous HLA ligand databases using DNA microarrays and virtual HLA class II matrices," *Nat. Biotechnol.*, 1999, doi: 10.1038/9858.
- [24] R. R. R. Rowland *et al.*, "Molecular characterization of swine leucocyte antigen class I genes in outbred pig populations," *Anim. Genet.*, 2009, doi: 10.1111/j.1365-2052.2009.01860.x.
- [25] C. S. Ho *et al.*, "Molecular characterization of swine leucocyte antigen class II genes in outbred pig populations," *Anim. Genet.*, 2010, doi: 10.1111/j.1365-2052.2010.02019.x.
- [26] G. Yu, "ggtree: an R package for visualization of tree and annotation data," *bioRxiv*, 2020.
- [27] E. B. Clemens, C. Van de Sandt, S. S. Wong, L. M. Wakim, and S. A. Valkenburg, "Harnessing the power of T cells: The promising hope for a universal influenza vaccine," *Vaccines*, vol. 6, no. 2, 2018, doi: 10.3390/vaccines6020018.
- [28] E. J. Erbeling *et al.*, "A universal influenza vaccine: The strategic plan for the national institute of allergy and infectious diseases," *J. Infect. Dis.*, vol. 218, no. 3, 2018, doi: 10.1093/infdis/jiy103.
- [29] E. Van Doorn *et al.*, "Evaluating the immunogenicity and safety of a BiondVax-developed universal influenza vaccine (Multimeric-001) either as a standalone vaccine or as a primer to H5N1 influenza vaccine," *Med. (United States)*, vol. 96, no. 11, 2017, doi: 10.1097/MD.00000000000006339.
- [30] G. H. Lowell, S. Ziv, S. Bruzil, R. Babecoff, and T. Ben-Yedidia, "Back to the future: Immunization with M-001 prior to trivalent influenza vaccine in 2011/12 enhanced protective immune responses against 2014/15 epidemic strain," *Vaccine*, vol. 35, no. 5, 2017, doi: 10.1016/j.vaccine.2016.12.063.

1. Author 1, A.B.; Author 2, C.D. Title of the article. *Abbreviated Journal Name* **Year**, *Volume*, page range.
Author 1, A.B. Title of Thesis. Level of Thesis, Degree-Granting University, Location of University, Date of Completion.
2. Title of Site. Available online: URL (accessed on Day Month Year).