T cell Epitope Content Comparison (EpiCC) of swine H1 influenza A virus hemagglutinin to identify best vaccine match

Andres H. Gutierrez[1], Matthew Ardito[1], William D. Martin[1], Anne S. De Groot[1,2]

[1]EpiVax, Inc., Providence, RI, United States; [2]Center for Vaccines and Immunology, University of Georgia, Athens, GA, United States

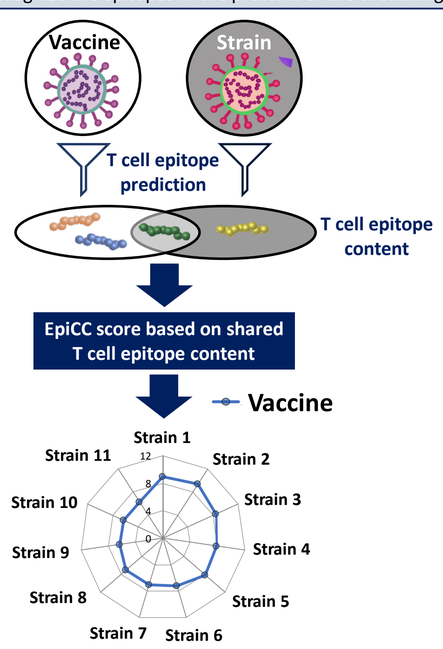
Introduction

EpiCC is an algorithm that compares the putative T cell epitope content shared between commercial vaccines and field isolates to identify the best vaccine match based on T cell epitope relatedness and coverage. The vaccine that covers more of the T cell epitope content of an isolate may confer broader cross-reactive cell-mediated immune response. Based on EpiCC and leveraging an analysis of a large population of field strains, we developed a web application for porcine circovirus type 2 (PCV2). Similar web applications can be developed for other swine pathogens (e.g., FluMatch™, PRRSVMatch™). To illustrate the application of EpiCC, we compared the putative T cell epitope content of hemagglutinin (HA) from circulating H1 swine influenza A virus strains and randomly selected strains from different phylogenetic clusters as potential vaccines to identify the best match. For each cluster, we also identified the best vaccine candidate.

Methods

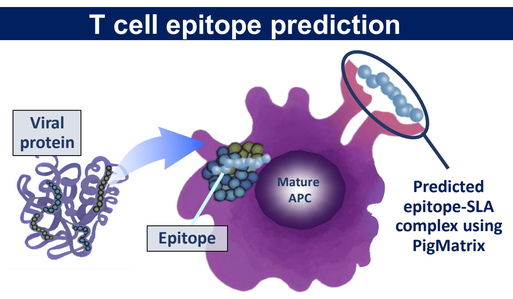
T cell Epitope Content Comparison (EpiCC)

Hypothesis: If epitopes in a vaccine closely match the epitopes in a circulating strain, the memory T cells induced by the vaccine are likely to recognize the epitopes in the proteins of the circulating strain.



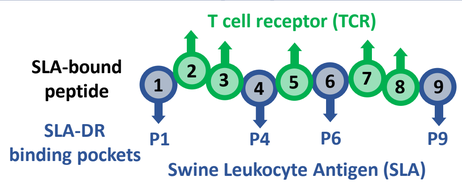
Application: Analyze existing or proposed vaccines for their potential to protect, based on a comparison between T cell epitopes in the vaccine and circulating strains of the same pathogen.

T cell epitope prediction



PigMatrix, built based on the pocket profile method, leverages pocket profiles already constructed for HLA-based epitope prediction in EpiMatrix to predict potential T cell epitopes for SLA class I and II alleles.

Shared T cell epitope content



JanusMatrix identifies potential cross-reactive T cell epitopes between the vaccine and circulating strains (predicted SLA ligands with Identical TCR-facing residues).

Data Input and Output

INPUT DATA

GISAID

930 HA sequences swine influenza A H1 strains in America 2019 to 2023.

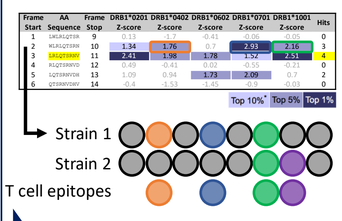
Phylogenetic classification

Frequency analysis

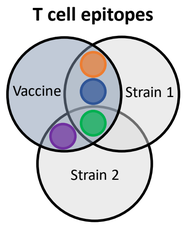
Non-redundant dataset (n=148) of HA sequences with phylogenetic classification and frequency annotation.

IN SILICO ANALYSIS

PigMatrix: CD4 and/or CD8 T cell epitope prediction



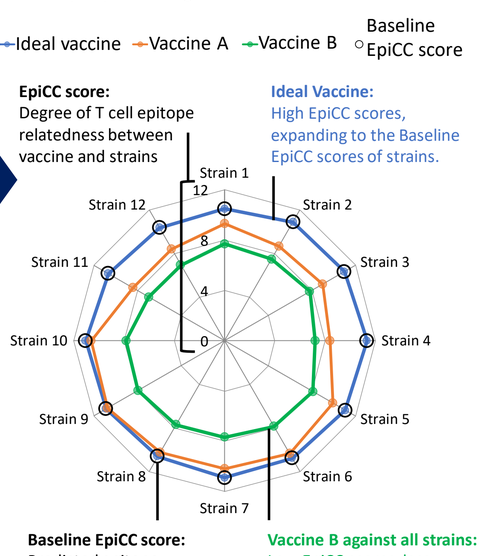
EpiCC: T cell Epitope Content Comparison



EpiCC score
Strain 1 > Strain 2
Vaccine shares more T cell epitope content with Strain 1.

EPICC OUTPUT

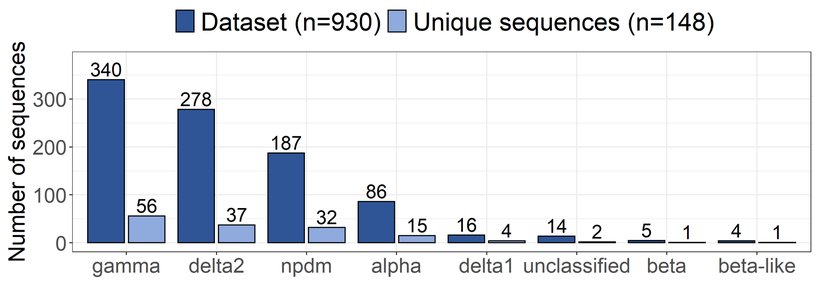
1. Number of shared T cell epitopes between pairs of sequences -> **Clustering**
2. EpiCC scores and radar plot -> **Mean EpiCC score calculation by cluster and vaccine selection**
3. T cell epitope coverage.



- HA sequences of American swine influenza A H1 strains from 2019 to 2023 from GISAID.
- A non-redundant dataset generated based on phylogenetic classification and frequency.
- Putative SLA class II T cell epitope content in the input proteins identified using PigMatrix.
- Assessed the relatedness of T cell epitopes contained in HA protein sequences using EpiCC.
- For this example, vaccine candidates selected based on mean EpiCC scores from the most frequent phylogenetic clusters.
- Each strain was evaluated as a potential vaccine.
- EpiCC generated a score for each vaccine-field isolate comparison.
- Higher EpiCC scores represent greater relatedness and higher T cell epitope coverage.

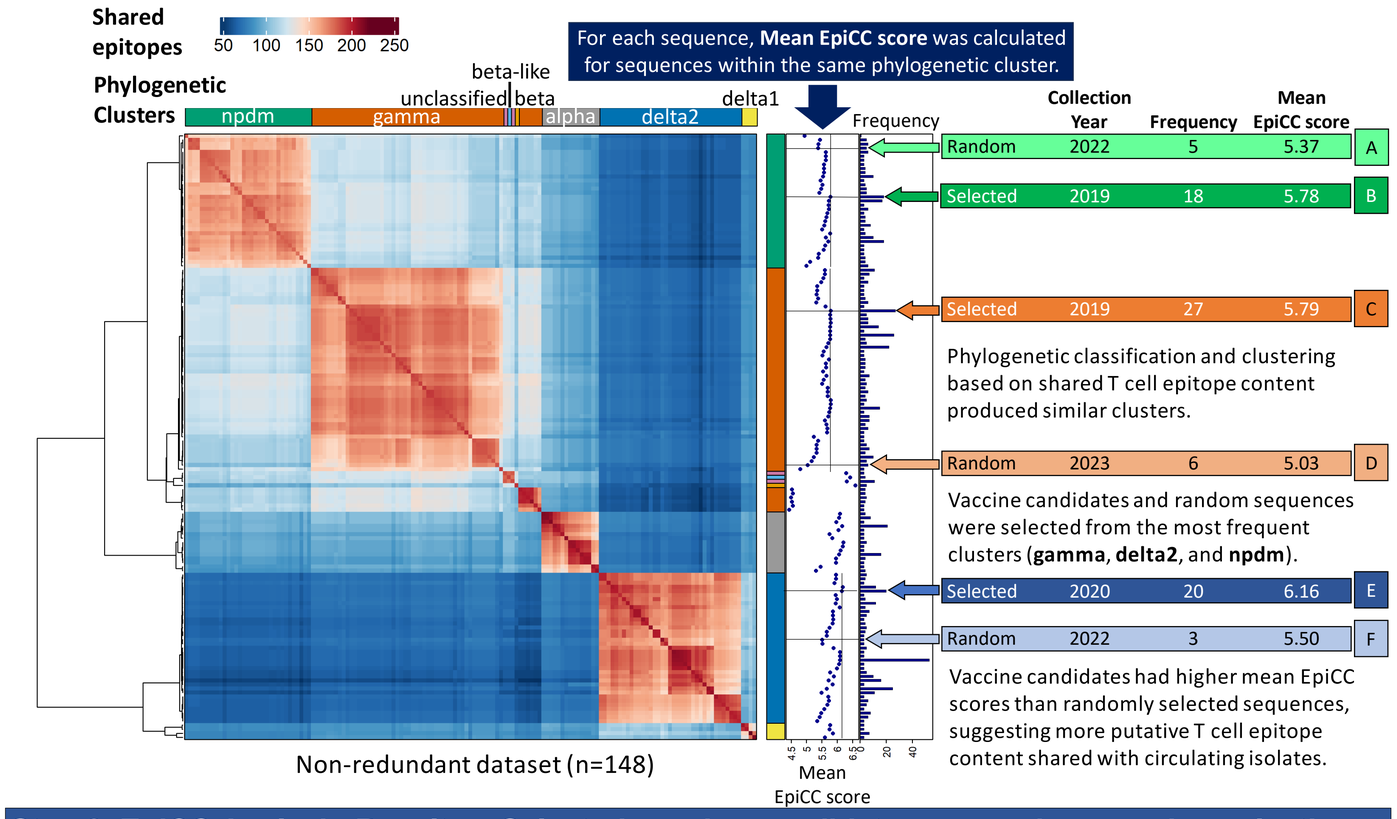
How to pick a Vaccine Strain

Step 1. Describe input data – phylogeny and frequency | Non-redundant dataset (n=148)



- **Gamma**, **delta2**, and **npdm** were the most frequently isolated HA gene sequences.
- These sequences represented 86.56% of the 930 HA sequences of swine influenza A H1 strains circulating in America from 2019 to 2023.
- The non-redundant dataset included the most recently collected unique sequences from each phylogenetic cluster.

Step 2. Process of vaccine candidate selection based on EpiCC scores and frequency

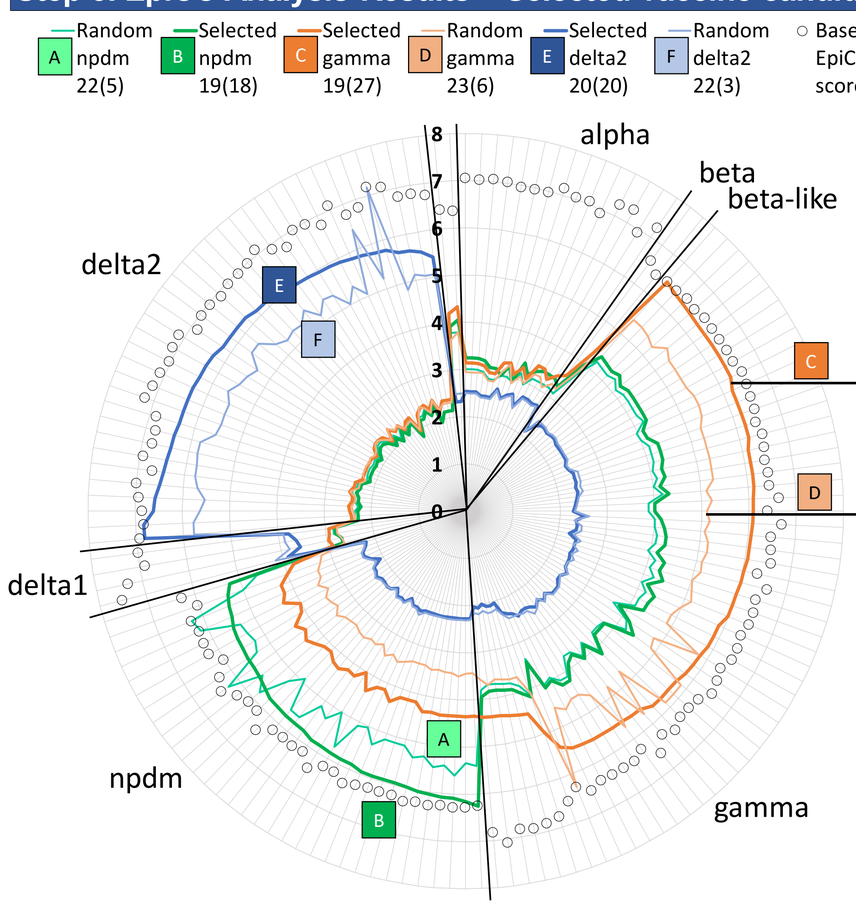


Phylogenetic classification and clustering based on shared T cell epitope content produced similar clusters.

Vaccine candidates and random sequences were selected from the most frequent clusters (**gamma**, **delta2**, and **npdm**).

Vaccine candidates had higher mean EpiCC scores than randomly selected sequences, suggesting more putative T cell epitope content shared with circulating isolates.

Step 3. EpiCC Analysis Results – Selected vaccine candidates vs. random vaccine selection



- The radar plot shows EpiCC scores for vaccine candidates vs. circulating strains.
- Each axis corresponds to the sequence of one strain.
- Baseline EpiCC scores (i.e., T cell epitope density) for each sequence are shown as open circles.

EXAMPLE OF VACCINE SELECTED

- Highest mean EpiCC score for gamma sequences.
- More related T cell epitopes between candidate and strains.
- EpiCC scores were more variable for gamma sequences

RANDOM VACCINE SELECTION

- Lower EpiCC scores for gamma sequences compared to the selected vaccine.
- Less related T cell epitopes with circulating strains.

INTERPRETATION OF RESULTS

Vaccine-to-strain comparisons were cluster-specific, which explains low EpiCC scores for alpha, beta, beta-like, and delta1 strains.

Vaccine candidates shared more T cell epitope content with strains than randomly selected vaccines.

Vaccine candidates may induce broader cross-reactive cell-mediated immune response.

Conclusions

- EpiCC was developed to help veterinarians, practitioners, producers, and farmers selecting the best-matched commercial vaccine for immunization against circulating field isolates and surveillance to identify variants that may represent a potential threat.
- EpiCC was used to identify potential swine influenza A H1 vaccine candidates based on HA, the key antigen. EpiCC can also be performed for complete proteomes.
- We selected vaccine candidates that had the highest EpiCC scores for sequences within the same phylogenetic cluster. Vaccine candidates cover more of the putative T cell epitope content predicted in the HA protein sequences of swine influenza A H1 strains.
- EpiCC may complement current serological methods for selecting the best-matched vaccine virus for immunization against circulating or emerging strains.
- Clinical data may help to refine EpiCC predictions and understand the relationship between shared epitope content and clinical outcomes.
- EpiCC can be applied to evaluate T cell epitope relatedness for other pathogens such as Porcine circovirus and Porcine reproductive and respiratory syndrome virus.

References

[1] Gutiérrez AH, et al. Development and validation of an epitope prediction tool for swine (PigMatrix) based on the pocket profile method. BMC Bioinformatics. 2015;16,290.
[2] Gutiérrez AH, et al. In vivo validation of predicted and conserved T cell epitopes in a swine influenza model. PLOS ONE.2016;11(7):e0159237.
[3] Gutiérrez, AH, et al. T cell epitope content comparison (EpiCC) of swine H1 influenza A virus hemagglutinin. Influenza Other Respi Viruses.2017;11:531- 542.
[4] Moise L, et al. New Immunoinformatics Tools for Swine: Designing Epitope-Driven Vaccines, Predicting Vaccine Efficacy, and Making Vaccines on Demand. Front Immunol. 2020;11:563362.
[5] Bandrick M, et al. 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. Vet Immunol Immunopathol. 2020;223:110034.
[6] Bandrick, M, et al. A bivalent porcine circovirus type 2 (PCV2), PCV2a-PCV2b, vaccine offers biologically superior protection compared to monovalent PCV2 vaccines. Vet Res. 2022;53,12.
[7] Foss, D, et al. T cell Epitope Content Comparison (EpiCC) analysis of porcine circovirus type 2 isolates from 2017-2021 confirms the global relevance of a bivalent vaccine approach. Vet Vac. 2023;2(2).