

ORIGINAL ARTICLE

Open Access



# Multi-epitope vaccine design of African swine fever virus considering T cell and B cell immunogenicity

Ting-Yu Chen<sup>1</sup>, Yann-Jen Ho<sup>2,3</sup>, Fang-Yu Ko<sup>2,4</sup>, Pei-Yin Wu<sup>5</sup>, Chia-Jung Chang<sup>5\*</sup> and Shinn-Ying Ho<sup>4,6,7,8\*</sup> 

## Abstract

T and B cell activation are equally important in triggering and orchestrating adaptive host responses to design multi-epitope African swine fever virus (ASFV) vaccines. However, few design methods have considered the trade-off between T and B cell immunogenicity when identifying promising ASFV epitopes. This work proposed a novel Pareto front-based ASFV screening method PFAS to identify promising epitopes for designing multi-epitope vaccines utilizing five ASFV Georgia 2007/1 sequences. To accurately predict T cell immunogenicity, four scoring methods were used to estimate the T cell activation in the four stages, including proteasomal cleavage probability, transporter associated with antigen processing transport efficiency, class I binding affinity of the major histocompatibility complex, and CD8+ cytotoxic T cell immunogenicity. PFAS ranked promising epitopes using a Pareto front method considering T and B cell immunogenicity. The coefficient of determination between the Pareto ranks of multi-epitope vaccines and survival days of swine vaccinations was  $R^2=0.95$ . Consequently, PFAS scored complete epitope profiles and identified 72 promising top-ranked epitopes, including 46 CD2v epitopes, two p30 epitopes, 10 p72 epitopes, and 14 pp220 epitopes. PFAS is the first method of using the Pareto front approach to identify promising epitopes that considers the objectives of maximizing both T and B cell immunogenicity. The top-ranked promising epitopes can be cost-effectively validated in vitro. The Pareto front approach can be adaptively applied to various epitope predictors for bacterial, viral and cancer vaccine developments. The MATLAB code of the Pareto front method was available at <https://github.com/NYCU-ICLAB/PFAS>.

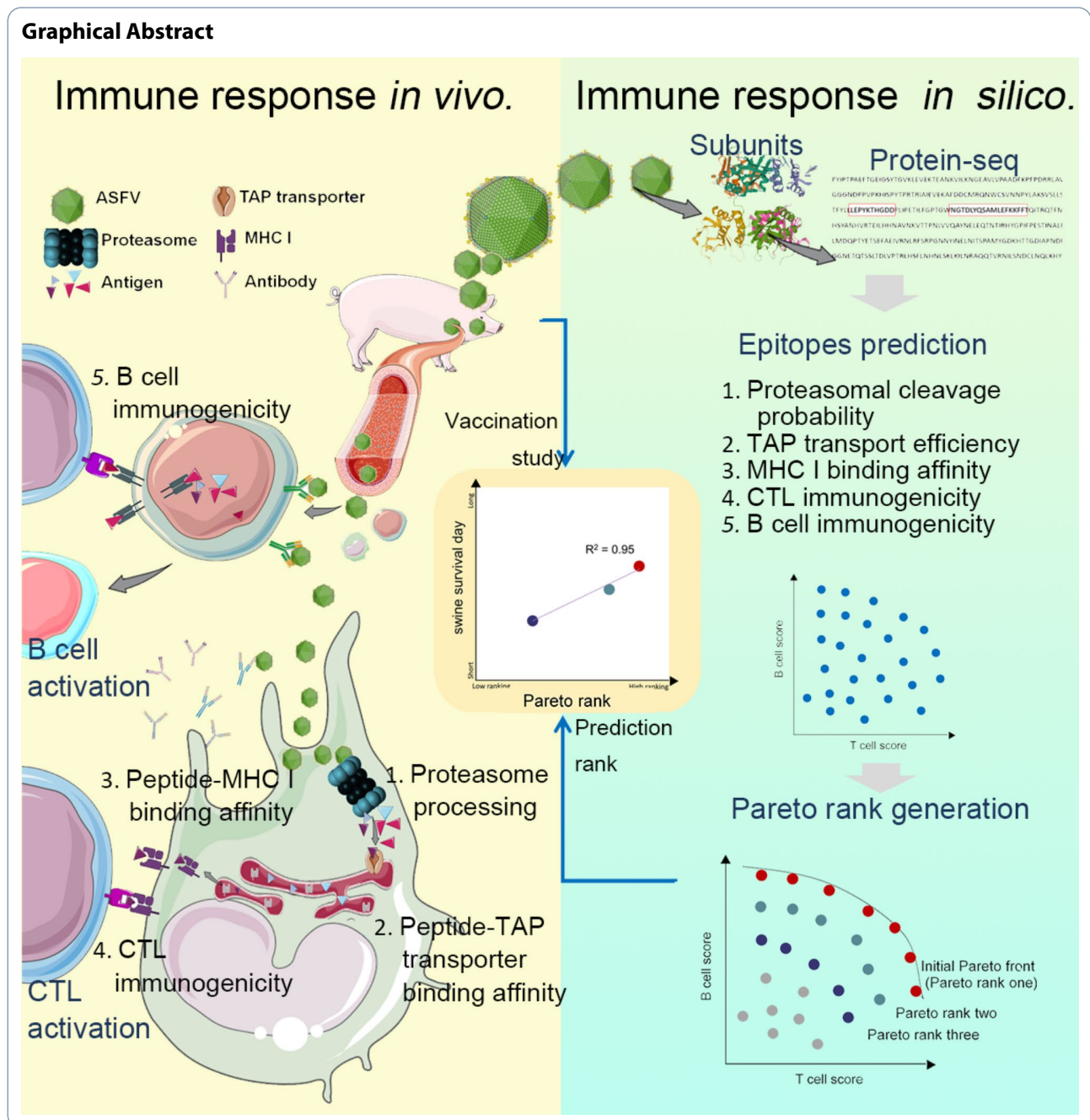
## Key points

- Proposing a Pareto front-based method for designing swine multi-epitope vaccine.
  - The method maximizes T and B cell immunogenicity while ranking promising epitopes.
  - Higher the epitope Pareto ranks leads to longer vaccination survival ( $R^2=0.95$ ).

**Keywords** African swine fever virus, Immunogenicity, Pareto front, Promising epitope, T and B cell activation, Vaccine design

\*Correspondence:  
Chia-Jung Chang  
frank.chang@reber.com.tw  
Shinn-Ying Ho  
syho@nycu.edu.tw

Full list of author information is available at the end of the article



## Introduction

African swine fever virus (ASFV) causes a lethal hemorrhagic disease and has become an epidemic swine viral disease in Asia. Simultaneous activation of T and B cells results in a better immune response and more immunological memory against ASFV than activation of one of the cell types alone (Bosch-Camos et al. 2020; Teklue et al. 2020). The development of subunit vaccines, especially multi-epitope vaccines, is more challenging than that of live-attenuated virus vaccines. However, live attenuated vaccines contain attenuated forms of pathogens that can

infect individuals with weakened immune systems and may revert to more virulent strains. As a result, these vaccines are only appropriate for endemic areas. This highlights the need for a shift towards the development of multi-epitope vaccines, which have significant applications in non-endemic regions and are driving increased research and development efforts. (Teklue et al. 2020). However, there is currently no effective multi-epitope vaccine for the prevention of ASFV (Blome et al. 2020).

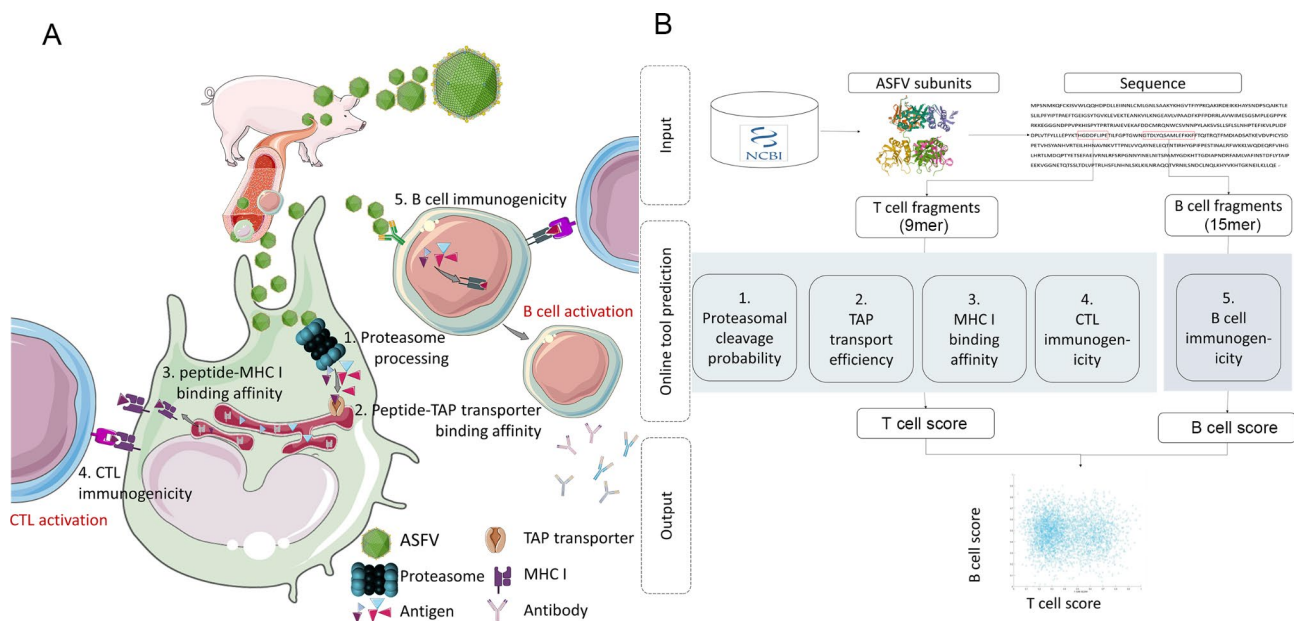
The selection of protein candidates for designing a multi-epitope vaccine should consider several factors,

including the conservation, abundance, extracellular localization, and cross-protection against various viral genotypes (Adamczyk-Poplawska et al. 2011; Alejo et al. 2018; Kessler et al. 2018). CD2v is a hemagglutinin and the main antigen protein involved in regulating immune responses and cell adhesion (Burmakina et al. 2019; Gaudreault and Richt 2019; Jia et al. 2017). p30 is a structural protein involved in the attachment and internalization of ASFV (Gomez-Puertas et al. 1996). p54 is the only membranous structural protein in the inner viral envelope associated with viral attachment (Gomez-Puertas et al. 1998). p72 is a major structural protein (approximately 31–33% of the entire virus) and an important antigenic protein owing to its high conservation and thermostable nature (Liu et al. 2019; Yu et al. 1996). pp220 is the largest multi-precursor protein (Lokhandwala et al. 2019). For immunogenicity and protection, CD2v, p30, and p72 show both antigenicity and immunogenicity, while p54 has antigenicity but low immunogenicity. pp220 displays immunogenicity and presents many peptides to CD8+ cytotoxic T cells (CTLs), triggering a strong antibody response (Bosch-Camos et al. 2020; Lokhandwala et al. 2019). In swine vaccine protection experiments, CD2v did not confer protection, while p30, p54, and p72 showed partial protection. Selecting proteins targeting different aspects of the immune response can provide effective protection. In addition, the binding of attachment proteins (p30, p54) to dominant B and T cell

epitopes (CD2v, p72) may enhance viral neutralization and clearance (Bosch-Camos et al. 2020; Jancovich et al. 2018).

Bioinformatic methods using a machine learning approach serve as an effective strategy to identify vaccine candidates for human (Guo et al. 2022; Hajjalibeigi et al. 2021; Kibria et al. 2022) and swine pathogens, including ASFV (Gao et al. 2021), influenza A virus (Baratelli et al. 2020; Fan et al. 2018), and porcine circovirus type 2 (Bandrick et al. 2020). Machine learning methods used to identify epitopes in the design of multi-epitope vaccines consider the biological presentation and activation of ASFV epitopes. Figure 1 shows the presentation and activation of ASFV epitopes, and the correspondence between biological and computational processes. In viral infections, CTLs (mainly involved in viral T cell immunity) support cell-mediated immunity against intracellular viruses, while B cells trigger humoral immunity and produce memory cells for future infection (Clem 2011).

The objective of vaccine design is to induce immune responses in both T and B cells. Existing computational methods for identifying ASFV epitopes screen potential epitopes by separately considering T and B cell immunogenicity (Bosch-Camos et al. 2021; Lopera-Madrid et al. 2017; Ros-Lucas et al. 2020). The Pareto front is a popular approach used for obtaining a set of non-dominated solutions to a bi-objective problem. Pareto-optimal methods have been used to dock proteins and peptides



**Fig. 1** Presentation and activation of ASFV epitopes, and the correspondence between biological and computational processes. **(A)** Illustration of the biological processing and presentation of the ASFV epitope. When infected by ASFV, four processing stages activate T cells in APCs and epitope presentation activates B cells in helper T cells. **(B)** The computational procedure of the Pareto front method. The input is a set of protein sequences. The outputs are prediction scores of T and B cell immunogenicity. The T cell score is the mean score of four stages: (1) proteasomal cleavage probability, (2) peptide-TAP transporter binding affinity, (3) peptide-MHC I binding affinity, and (4) CTL immunogenicity. CTL: Cytotoxic T lymphocyte. APC: Antigen-presenting cell. TAP: the transporter associated with antigen processing. Images of figure were partly taken from 'Smart Servier Medical Art' (<https://smart.servier.com/>)

(Masoudi-Sobhanzadeh et al. 2021) and improve amino acid and protein production in *Yarrowia lipolytica* (Jach et al. 2020). A study related to epitope-based vaccine design against human immunodeficiency virus used the Pareto front to simultaneously optimize cleavage and immunogenicity (Dorigatti and Schubert 2020). However, few studies have used the Pareto front method to simultaneously accommodate T and B cell immunogenicity.

This work proposes for the first time a novel Pareto front-based screening method PFAS to identify promising epitopes with high T and B cell immunogenicity for designing ASFV recombinant multi-epitope vaccines. First, PFAS used experimental T and B cell epitopes from the Immune Epitope Database (IEDB) to verify the state-of-the-art computational methods and their parameter settings. Next, PFAS used the Pareto front technique to deal with T and B cell prediction scores as bi-objective ranks and identify the top-ranked epitopes. PFAS scored whole epitope profiles and identified 72 promising epitopes. Based on the three combinations of epitopes in pp220, p30, p72, and p54 for a vaccination study against ASFV, the determination coefficient of determination between the Pareto ranks of recombinant multi-epitope vaccines and swine survival was  $R^2=0.95$ . The identified epitopes can be cost-effectively validated *in vitro* to design epitope-based ASFV vaccines.

Immunoinformatics-based reverse vaccinology approaches hold great promise for reducing the time and cost of vaccine development. Currently, several approaches can be used to optimize and validate multi-epitope vaccines. These approaches include allergenicity assessment, protein structure verification, docking conformation, and *in silico* immune simulations of the vaccine structure (Bappy et al. 2021; Gul et al. 2022). The number of potential epitopes screened can be dynamically adjusted using PFAS for subsequent validation in multi-epitope combination simulations. PFAS can be seamlessly integrated into immunoinformatics-based vaccine development pipelines to generate high-potential epitope combinations for biological experimental confirmation.

## Materials and methods

### Collection of ASFV protein sequences

The most lethal ASFV type, Georgia 2007/1 (GenBank: FR682468), was used as the target virus for screening. The protein sequences of Georgia 2007/1 were obtained from the NCBI, including CD2v (EP402R), p30 (CP204L), p54 (E183L), p72 (B646L), and pp220 (CP2475L). All sequences were cut into fragments using a sliding window. Finally, two datasets of ASFV proteins consisting of 9mer (Figure S1A) and 15mer (Figure S1B) fragments served as candidates for CTL and B cell epitopes, respectively.

### Collection of validation datasets

To obtain the best parameter settings for PFAS, this work established two datasets from the IEDB, consisting of experimentally validated CTL and B cell epitopes of swine. The CTL epitopes ( $n=243$ ) were annotated as *Sus scrofa*, infectious diseases, and Swine Leukocyte Antigen (SLA) class II. After simultaneously removing duplicate and uncertain sequences belonging to both positive and negative groups, the dataset contained 125 swine 9mer CTL fragments, including 37 epitopes and 88 non-epitopes.

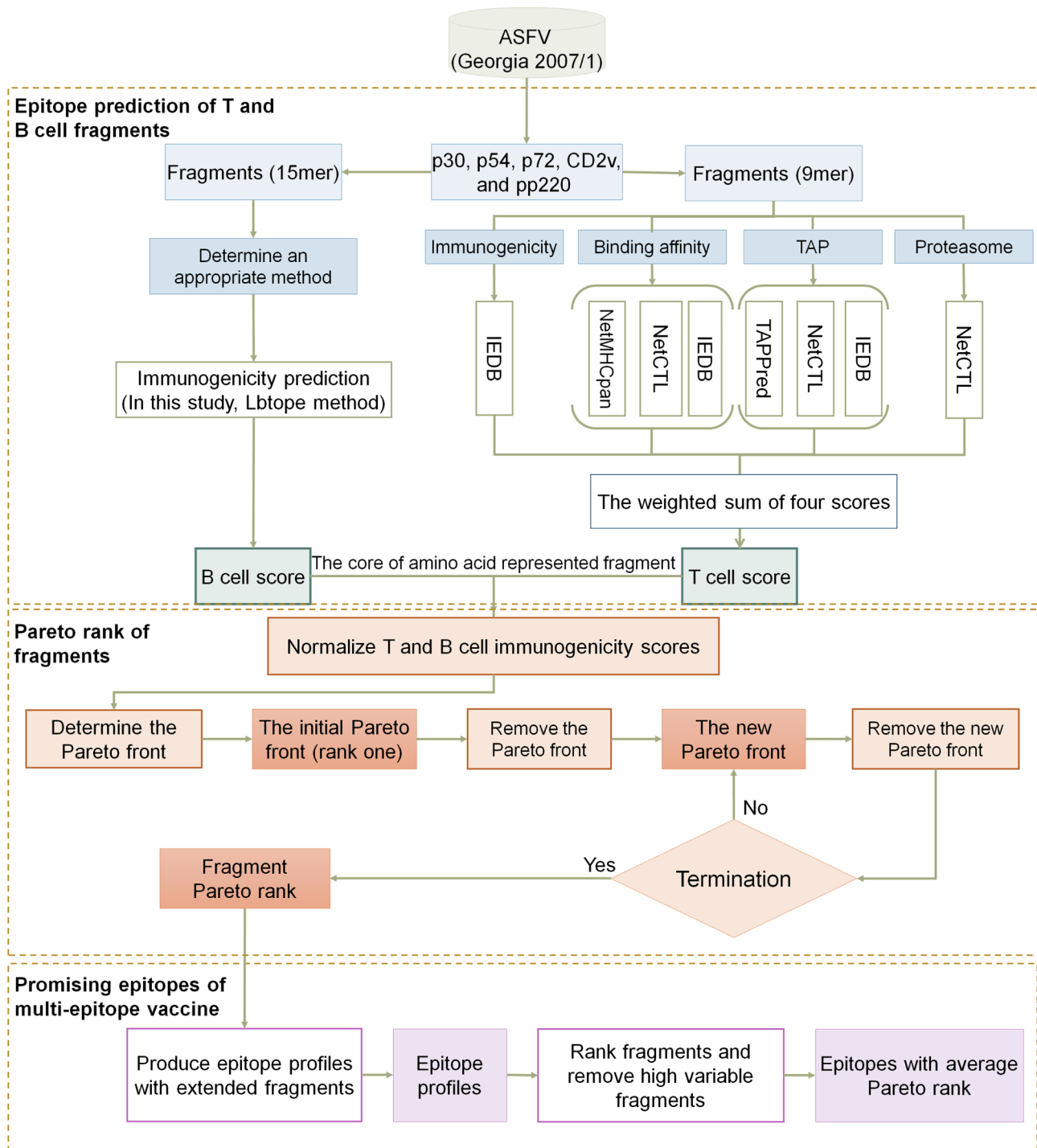
Similarly, 1,700 validated swine B cell epitopes (BCEs) which were annotated as *Sus scrofa* and infectious disease were retrieved. Because IgG production is part of the secondary humoral immune response to an antigen, we extracted 1,389 IgG epitopes. Among them, the 15mer epitope was the largest in the dataset, followed by the 12mer epitope. Therefore, we established two datasets: (1) 650 B cell 15mer epitopes, including 116 positive and 534 negative epitopes, and (2) 293 swine B cell 12mer epitopes, including 35 positive and 258 negative epitopes.

### Proposed method PFAS

Figure 2 shows a flowchart of the proposed method PFAS. Five protein sequences from Georgia 2007/1 were obtained and cut into 9mer and 15mer fragments. The CTL epitope predictor estimates T cell activation of 9mer fragments in the four stages and averages the four scores to obtain a T cell immunogenicity score. Similarly, the BCE predictor estimates B cell activation of 15mer fragments to obtain a B cell immunogenicity score. After normalizing these two scores into the range of [0, 1], the two fragments were superimposed by the central amino acid. Consequently, the fragments were extended, and thus conserved sequences were obtained. The Pareto front method produced ranks of the conserved fragments. The top-ranked fragments were considered as promising epitopes.

### Calculation of T and B cell scores

Good CTL epitopes are involved in viral processing and antigen presentation, with major histocompatibility complex (MHC) I molecules playing a major role. First, pathogen debris is degraded by proteasomal degradation in the cytosol of productively infected cells. NetCTL is based on the NetChop method and predicts the probability of proteasomal cleavage (Larsen et al. 2007). Second, peptides are transported to the endoplasmic reticulum (ER) by a transporter associated with antigen processing (TAP). To predict TAP transport efficiency, NetCTL and MHC I Processing in the IEDB use the stabilized matrix method, and TAPPred is based on a support vector machine (SVM) with 33 physical features of amino acids (Bhasin and Raghava 2004). Third, an antigen



**Fig. 2** The flowchart of PFAS. This flowchart includes three main parts: epitope prediction of T and B cell fragments, Pareto rank of fragments, and promising epitopes of multi-epitope vaccines

is loaded onto MHC I and appears on the cell surface through vesicles. NetMHCpan (Reynisson et al. 2020), MHC I Processing, and NetCTL are the most widely used ANN-based methods to predict MHC I binding affinity using the BLOSUM50 matrix. Finally, the epitope stimulates CTL activation and differentiation. MHC

I immunogenicity in the IEDB (Calis et al. 2013) is based on an immunogenicity score model to predict immunogenicity. In general, these four predictive roles are equally important.

To identify promising T cell epitopes (TCEs), five web predictors were used, including NetCTL (<https://services.>

[healthtech.dtu.dk/service.php?NetCTL-1.2](http://healthtech.dtu.dk/service.php?NetCTL-1.2)), IEDB MHC I Processing (<http://tools.iedb.org/processing/>), TAPPred (<https://webs.iitd.edu.in/raghava/tappred/index.html>), NetMHCpan (<https://services.healthtech.dtu.dk/service.php?NetMHCpan-4.0>), and IEDB MHC I Immunogenicity. NetCTL was used to predict proteasome processing, TAP transport efficiency, and MHC I binding affinity. To examine conserved epitope candidates that cover multiple MHC loci, including A1, A2, A3, A24, A26, B7, B8, B27, B39, B44, B58 and B62, we used sequences as inputs and applied ensemble learning with 12 supertype models. After averaging all predictive values in the 12 models, we obtained three estimated values: binding affinity, proteasome cleavage, and the TAP score. The IEDB MHC I Processing tool was used to estimate TAP transport efficiency and MHC I binding affinity. We used all 45 SLA I alleles (including 12 SLA1, 16 SLA2, 12 SLA3, and 5 SLA6) and set nine as the peptide length for each allele to obtain a file with the average predictive values, including the TAP and MHC scores in all sequence fragments. PFAS used TAPPred to predict the peptide-TAP transporter binding affinity based on SVM with validated sequences and obtained the prediction score. NetMHCpan was used to predict the binding affinity of peptide-MHC I. To obtain effective epitopes, we considered all 75 SLA alleles (including 23 SLA1, 26 SLA2, 21 SLA3, and five SLA6) and set nine as the peptide length for each allele. The binding affinity scores were estimated with mean scores for all fragments. This work used IEDB MHC I Immunogenicity to predict CTL immunogenicity considering all CTL active factors and obtained scores of all sequence fragments.

BCEs can induce the differentiation of naïve and memory B cells into plasma cells, including antigen processing, peptide-MHC II presentation, and cytokine promotion. In studies on BCE presentation, LBtope (Singh et al. 2013), iBCE-EL (Manavalan et al. 2018), IgPred (Gupta et al. 2013), and ABCpred (Saha and Raghava 2006) are sequence-based predictors. LBtope uses the sparse matrix and amino acid property profile features and is an SVM-based Weka Classifier using 38,197 IEDB experimental epitopes. iBCE-EL is based on ensemble learning using amino acid composition characteristics and proportions of 5,550 experimentally validated BCEs. IgPred uses 14,725 BCEs in different types of specific epitopes using physicochemical properties (PCPs) features and is based on Weka Classifiers. ABCpred is based on PCP features and the neural network method with a balanced BCE database. Among the aforementioned predictors, LBtope uses the largest dataset with ensemble learning.

To estimate the B cell immunogenicity score of 15mer and 12mer fragments, five online predictors (LBtope\_Variable, LBtope\_Confirm, iBCE-EL, IgPred, and ABCpred) were utilized and validated. Epitope

probabilities and IgG scores were determined using the iBCE-EL and IgPred prediction tools, respectively. LBtope is based on multiple peptides from prediction models using two variable-length epitope models. The LBtope\_Variable model was trained using 38,197 peptides. The LBtope\_Confirm model was reported in at least two studies and contained 2,837 peptides. By submitting multiple fragments, the probability of epitopes was obtained along with the physical property score. As ABCpred exclusively accepts an even number of epitope lengths and continuous amino acid sequences as submissions, PFAS used only one 12mer dataset with parameters containing a threshold of zero and an overlapping filter to obtain the predicted scores.

#### Immunogenicity prediction of T and B cell fragments

The CTL activation prediction has four important stages: proteasomal cleavage probability, TAP transport efficiency, MHC I binding affinity, and CTL immunogenicity. These predictions help identify potential TCE candidates. PFAS combined all the prediction values obtained from the online prediction tools in the four stages. The probability of proteasomal cleavage was estimated using NetCTL1.2. The TAP transport efficiency score is the mean score of NetCTL1.2, IEDB MHC I Processing, and TAPPred values. The peptide-MHCI binding affinity score is the mean score of NetCTL1.2, IEDB MHC I Processing, and NetMHCpan predictive values. The CTL immunogenicity score is obtained using IEDB MHC I Immunogenicity. After combining and normalizing the scores of each category using a combination of weights, PFAS compiled four stage score for the TCE prediction.

For the BCE prediction, the best predictor was evaluated and used to obtain B cell immunogenicity scores. After compiling the results of the prediction values from the web tools, the output values were normalized into the range of [0, 1] and B cell immunogenicity scores were compiled.

#### Pareto rank of fragments

The Pareto front is the set of all efficient solutions to bi-objective problems. In this study, a fragment *Frag* belonging to the Pareto front means that no other fragment has both larger T and B cell scores than *Frag*. The T and B cell scores of all fragments which were represented by their central amino acids were used as inputs of the Pareto front method to determine the Pareto rank of fragments. The Pareto front method iteratively removes the Pareto fronts, and Pareto rank of the fragments was the serial number of the removed front. For instance, the segments belonging to the initial Pareto front have a rank one. After removing the Pareto front, the fragments belonging to the new Pareto front have a rank two, and so on.

### Promising epitopes of the multi-epitope vaccine

This work extended the fragments to a length of 16–20 amino acids and obtained the epitope profiles with the average Pareto rank of the extended fragments. The average rank was defined as the sum of the Pareto ranks divided by the total number of fragments included in the extended fragments. Moreover, to select conserved epitopes, PFAS estimated protein variability using the Protein Variability Server (PVS) (Garcia-Boronat et al. 2008). PVS contains three methods, the Shannon entropy, the Simpson diversity index, and the Wu–Kabat variability coefficient method, which can be used as indicators of variability. In this study, the Shannon entropy greater than two was considered as the variability point. Accordingly, PFAS removed the variable fragments that contained highly variable sequences. Finally, in the 16mer to 20mer epitope profiles, PFAS ranked conserved fragments according to the average Pareto rank, and the top-ranked promising epitopes were provided to the biological decision makers for in vitro validation.

## Results

### Estimation of T and B cell immunogenicity scores

After T cell online prediction, the T cell score was the weighted sum of four scores in the four stages: proteasomal cleavage probability, TAP transport efficiency, MHC I binding affinity, and CTL immunogenicity scores. For each stage, the scores were normalized and averaged. To validate the weights of four scores, the experimentally validated swine 9mer CTL epitopes from the IEDB were used. Figure S2A shows a good performance with an area under the receiver operating characteristic curve (AUC) of 0.71, and the largest AUC was reached when using the top 30% epitopes (Figure S3). The set of four equal weights 1/4, 1/4, 1/4, and 1/4 in determining the T cell score is the most stable one. This result is consistent to the previous hypothesis that the importance of the four stages in the swine immunogenicity prediction is equal.

In B cell immunogenicity prediction, previous studies have revealed that the use of sequence-based predictors, such as LBtope (Singh et al. 2013), iBCE-EL (Manavalan et al. 2018), IgPred (Gupta et al. 2013), and ABCpred (Saha and Raghava 2006) is an efficient approach to identifying BCEs. Owing to different aims of the training dataset and machine learning approaches, the prediction results were different from these methods. Therefore, PFAS used LBtope, the largest experimental dataset with ensemble learning, as a prediction model to identify BCEs. Experimentally validated swine 15mer and 12mer BCEs were used to validate this hypothesis, the validation pipeline consistent with that used for TCE prediction. Figure S2B shows performance of the four methods: iBCE-EL, IgPred, LBtope with a variable dataset, and LBtope with a confirmed dataset. LBtope with the

variable dataset achieved an AUC of 0.86. When using the top-ranked 30% epitopes, LBtope with a confirmed dataset was better than the other methods (Figure S4–5). These results were in good agreement with the hypothesis, showing that LBtope is an appropriate method for predicting ASFV BCEs.

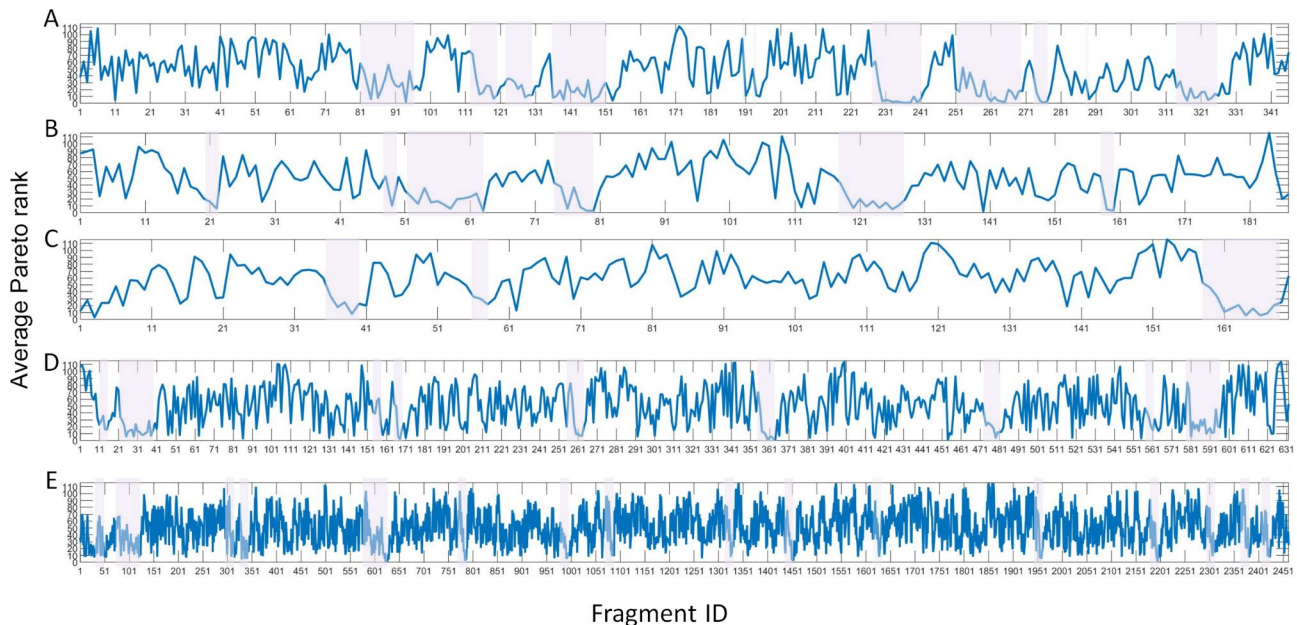
### Epitopes identification using the Pareto front method

Given that both T and B cell activation are equally important for mobilizing adaptive immunity, we applied the Pareto front method to identify potential epitopes. To rank and identify fragments simultaneously, the Pareto front method iteratively determined 116 Pareto ranks (Figure S6). T and B cell scores in the bi-objective problem were converted into Pareto ranks to identify epitope candidates.

Figure 3 shows 15mer epitope profiles for the five ASFV proteins. A higher average Pareto rank indicates a more promising epitope. Table S1 shows the results of the selected fragments in the five proteins, including 346 fragments of CD2v, 187 fragments of p30, 170 fragments of p54, 632 fragments of p72, and 2462 fragments of pp220. In short, Table 1 shows the Pareto ranks of the top three fronts. The best protein is CD2v, which has the most selected and continuous fragments in the rank one Pareto front.

### Evaluation of screening efficiency

To evaluate the screening efficiency of PFAS, two validation datasets were used consisting of 30 experimentally validated and 34 predicted epitopes of T or B cells annotated in previous studies and the IEDB database (Bosch-Camos et al. 2021; Ivanov et al. 2011; Ros-Lucas et al. 2020). PFAS selected the top 30% fragments as promising epitopes. Table 2 lists the public epitopes with the Pareto ranks. PFAS identified 17 epitopes from 30 experimental ones and 24 epitopes from 34 predicted ones. Figure 4 revealed scatter points of experimental, predicted and PFAS selected epitopes. Since animal studies can prove the actual antigenicity and immunogenicity of epitopes, the top-ranked epitopes may be superior to the published epitopes. Three recombinant multi-epitope vaccines with synthesized epitope groups were used to determine the determination coefficient between the predicted ranks of PFAS and swine immunization (Ivanov et al. 2011). The combinations 1, 2, and 3 consisting of four pp220 epitopes, six p30 and p72 epitopes, and two p54 epitopes, respectively. The Pareto ranks of peptides in each combination were determined using the Pareto front method. Figure 5 indicates a significant coefficient of determination with  $R^2=0.95$  between the mean Pareto ranks of the recombinant vaccine and swine survival days in a vaccination study. The higher the combination Pareto rank, the longer the survival days of the pig (Table S2). These



**Fig. 3** Epitope profiles of the five ASFV proteins. Each 15mer fragment has an average Pareto rank represented by the central amino acid. The higher Pareto rank indicates greater epitope potential. The light pink background indicates the epitope hotspot. **(A)** CD2v. **(B)** p30. **(C)** p54. **(D)** p72. **(E)** pp220. ID, identification

results reveal that PFAS is an efficient approach to epitope identification.

#### Identification of promising epitopes for multi-epitope vaccines

Clustering epitopes into hotspots (high-ranked epitopes) would be an effective method to obtain vaccine candidates in the multi-epitope vaccine design. The results shown in Table 1 are consistent to previous studies in which highly-ranked sequences were continuous. Accordingly, PFAS extended the fragments to sequences of 16–20 amino acids and produced epitope profiles with the average Pareto ranks of extended fragments (Figure S7–11). The higher the front rank, the greater potential of the epitope. Enrichment of potential epitopes is considered as an epitope hotspot.

Additionally, epitope variability is important for biologists in identifying vaccine candidates. To obtain conserved epitopes, PFAS calculated the variability of the five proteins using PVS (Table S3). A total of 45 sites were identified, including two CD2v sites and 43 p54 sites, which can be regarded as sites with high variability because their Shannon entropy was greater than two (Figure S12). Similarly, if a fragment contained highly variable sites, it was regarded as a highly variable region. After removing 69 highly variable fragments (Table S4), 3728 fragments were obtained, including 341 CD2v fragments, 187 p30 fragments, 106 p54 fragments, 632 p72 fragments, and 2462 pp220 fragments. After conservation verification, the mean Pareto rank of the extended

fragments for each protein was determined. CD2v had the highest average rank among the five proteins, and p30, p72, p220, and p54 ranked second, third, fourth, and fifth, respectively. Furthermore, we estimated all conserved fragments (Table S5). Biological decision makers can flexibly choose the appropriate epitope length and sample size for experimental validation in vitro (Tables S6–10). For example, Table 3 shows 72 promising epitopes with an average Pareto rank of four, including 26 16mer epitopes, 16 17mer epitopes, 12 18mer epitopes, 10 19mer epitopes, and 8 20mer epitopes. In addition, these epitopes came from four proteins, including 46 CD2v epitopes, two p30 epitopes, 10 p72 epitopes, and 14 pp220 epitopes.

#### Discussion

Since ASFV is a complex and lethal multi-antigen virus, it originated in Africa but has recently caused an emerging epidemic in Asia. With advances in computational biology and machine learning in the field of immunology, computational epitope prediction provides a new opportunity to improve ASFV multi-epitope vaccines. Several studies have demonstrated that ASFV enhances or modulates the host immune response through multiple proteins. Therefore, a recombinant multi-epitope vaccine has a potential to be an excellent ASFV vaccine.

In this work, we have analyzed five proteins: CD2v, p30, p54, p72, and pp220. Variation in MHC polymorphisms would induce different immune responses (Opriessnig et al. 2021), which plays an important role in identifying



**Table 1** Fragments of the top three fronts using PFAS

Front	Protein	ID	Fragment	T cell score	B cell score	Rank	
Rank 1	CD2v	236	KHVEEIESPPESNE	0.16	0.97	1	
	CD2v	237	HVEEIESPPESNEE	0.13	1.00	1	
	CD2v	238	VEEIESPPESNEEE	0.27	0.93	1	
	CD2v	240	EIESPPESNEEEQC	0.43	0.90	1	
	CD2v	276	YSRYQYNTPIYYMRP	1.00	0.73	1	
	p72	360	KLASQKDLVNEFPGL	0.85	0.88	1	
	p72	454	KLMSALKWPIEYMF	1.00	0.33	1	
	pp220	623	WKATVSAIELEYDVK	0.90	0.84	1	
	pp220	626	TVSAIELEYDVKRRF	0.41	0.92	1	
	pp220	2195	FRTQLEDTRREVNNL	0.56	0.90	1	
	Rank 2	CD2v	94	TYQVWVWNIINYTIK	0.93	0.72	2
CD2v		147	FVKYTNESILEYNWN	0.96	0.69	2	
CD2v		233	KRKKHVEEIESPPPE	0.38	0.90	2	
CD2v		235	KKHVEEIESPPESN	0.31	0.90	2	
CD2v		265	PSPREPLLKPYSTRY	0.71	0.85	2	
p30		140	LAQKTVQHIEQYGKA	0.99	0.52	2	
p72		168	GTKNAYRNLVYYCEY	0.99	0.64	2	
p72		363	SQKDLVNEFPGLFVR	0.75	0.84	2	
p72		364	QKDLVNEFPGLFVRQ	0.83	0.73	2	
pp220		785	SPLQIYKTLLEYLQH	0.79	0.79	2	
pp220		1453	QSSERFEQYGRVFSR	0.64	0.86	2	
Rank 3		CD2v	230	SLRKRKKHVEEIESP	0.71	0.81	3
		CD2v	234	RKKHVEEIESPPES	0.34	0.89	3
	CD2v	261	SIHEPSPRELLPKP	0.64	0.84	3	
	CD2v	264	EPSPRELLPKPYSR	0.24	0.90	3	
	CD2v	277	SRQYNTPIYYMRPS	0.77	0.77	3	
	p30	63	VKSARIYAGQGYTEH	0.87	0.68	3	
	p30	79	AQEEWNMILHVLFFEE	0.71	0.83	3	
	p30	80	QEEWNMILHVLFFEE	0.79	0.76	3	
	p30	160	VIRAHNFIQTIYGTP	0.95	0.59	3	
	p54	3	SEFFQPVYPRHYGEC	0.83	0.68	3	
	p72	86	LGKLTFGIPQYGDF	0.96	0.59	3	
	p72	555	SKFCSSYIPFHYGGN	0.93	0.66	3	
	pp220	1450	NNPQSSERFEQYGRV	0.79	0.74	3	
	pp220	1452	PQSSERFEQYGRVFS	0.36	0.85	3	

Scores are normalized into the range of [0, 1]. ID, identification of fragments

potential epitopes. Human epitopes have been widely used to build prediction models. However, very few swine epitopes and prediction models are available. To identify potential epitopes for swine vaccine, PFAS used state-of-the-art predictors with promising parameter setting to calculate T and B cell scores. Even when experimentally validated porcine epitopes were used for parameter validation, cross-species prediction models may still reduce prediction accuracy.

Although both the T and B cell immunogenicity are important, there must be a trade-off in identifying promising epitopes for conventional prediction methods. Therefore, the Pareto front method was proposed to cope with the bi-objective problem by converting both the T and B cell scores of a fragment into a single Pareto rank that vaccine designers can easily determine the number

of promising epitopes for biological experiments. The validation of ASFV recombinant multi-epitope vaccines reported suggests that the Pareto front method would be a potentially useful approach to identifying promising epitopes in the design of multi-epitope vaccines against ASFV.

Because ASFV has multiple antigens and complex immune interactions with the host immune system, ASFV recombinant multi-epitope vaccines require a multi-epitope combination. Therefore, we analyzed the protein features of the top-ranked epitopes in five potential proteins, and some results were consistent to those of the biological studies. Some studies have shown that CD2v exhibits serological specificity, participates in immune evasion, enhances viral replication, and damages lymphocyte functions (Sanna et al. 2017). In this work,

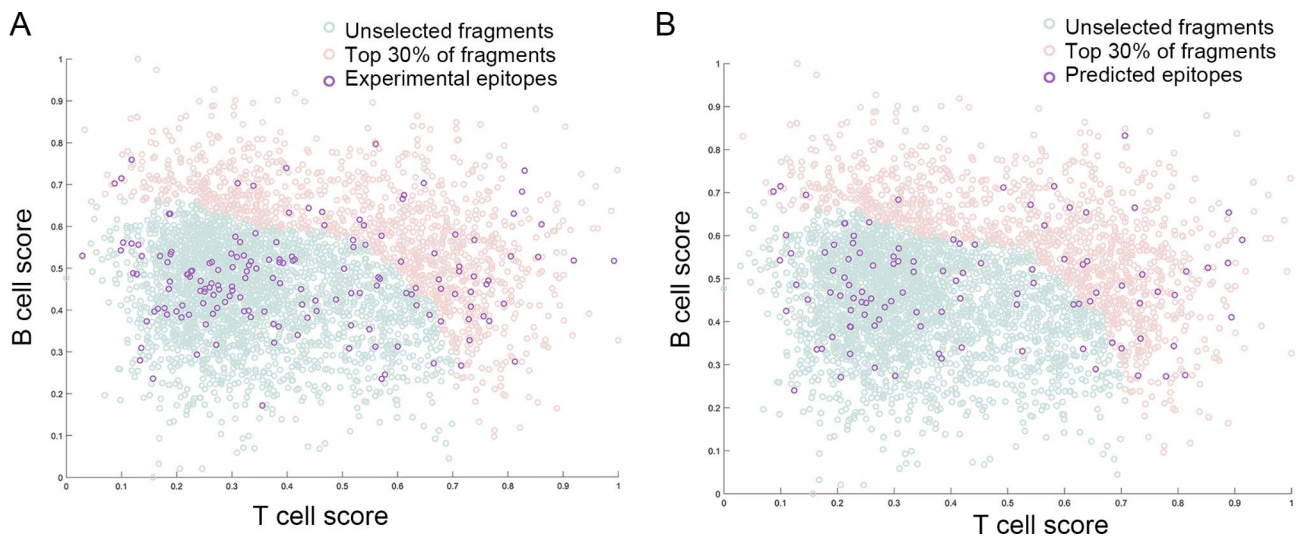
**Table 2** The Pareto rank of the epitopes in the top 30% fragments for the experimentally validated and predicted epitopes

Protein	Experimental epitope	Rank	Protein	Predicted epitope	Rank
CD2v	KPCPPPKPCPPPKC	21	CD2v	CTYLTLSSNYFYTFKLYIPL	–
CD2v	PPKPCPPPKPCPPPK	–	p30	SQWFHAGSLY	6
CD2v	YSPPKPLPSIPLPN	14	p30	AQEEWNMIL	3
CD2v	SPPKPLPSIPLLPNI	14	p54	YTHKDLENSL	–
CD2v	PPKPLPSIPLLPNIPPLSTQNISLI	–	p72	AAIEEEDIQFINPYQD	–
p30	EVIFKTD	21	p72	KPYVPVGFY	7
p30	TSSFETLFEQ	8	p72	GFEYNKVRPHTGTPTLGNKLT	20
p30	TVQHIEQYGKA	2	p72	QMGAGHQLQTFPRNGYDWDNQTPLE	4
p30	QHIEQYGKAPDFNKV	27	p72	NVRFDVNGNSL	18
P30	LKEEKEVRLMVIKLLKKNKL	–	p72	YCEYPGERLYENVRFDVNGNSLDEYSSDVTTL	15
p54	MDSEFFQPVYPRHYGECLS	3	p72	HKPHQSKPILTDENDTQRTCS	–
p54	FQPVYPRHYGECLSP	–	p72	FPENSHNIQTAGKQD	–
p54	QPVYPRHYGECLSPV	20	p72	HTNPKFLSQHFPENSHNIQTAGKQDITPITD	33
p54	PVYPRHYGECLSPVT	–	p72	RPSRRNIRF	4
p54	VYPRHYGECLSPVTT	–	p72	TWNISDQNPQHHRDWHK	22
p54	YPRHYGECLSPVTTT	–	p72	VHTNNNHDEKLMS	–
p54	PRHYGECLSPVTTPSFF	–	p72	SFQDRDTPALPDACSSISDI	16
p54	YGECLSPVTTPSFFS	–	p72	LLQNGSAVLRYSY	–
p54	GECLSPVTTPSFFST	23	pp220	NKALQKVG	–
p54	SRKKKAAAIEEEDI	–	pp220	SQVDLNQAINTFMYYYYVAQIY	26
p54	NKPVTDNPVTDRL	–	pp220	HNKQEFQSY	15
p72	YCEYPGERLYENVRFDVNGNSLDEYSSDVTTL	15	pp220	ITKTFVNNI	14
p72	LCNIHDLHKPHQSKPILTDENDTQRTCS	20	pp220	DNAPAGHY	30
p72	QKDLVNEFPGLFIQSRFIPGRPSRRNIRFKP	2	pp220	TPEEAAQRVY	27
p72	ACSSISDISPVYTPITLPIIKNISVTAHGINLIDK	4	pp220	VNDALSTRW	31
p72	LKPREEYQPS	6	pp220	MAAKIFIVL	9
pp220	YDSCSRLQIIDFYTDIVQKQYGGGEDCECTRV	19	pp220	EFYQKLFSS	21
pp220	PKGQTRTLGNSRERERI	–	pp220	ARTMNDFGM	–
pp220	GYMSRIFRGDNALNM	–	pp220	NRSNPGSFY	25
pp220	YMSRIFRGDNALNMG	29	pp220	IQNNRSMVMVFNQLIASYITRFY	30
			pp220	IPIYLKENY	24
			pp220	YMSRYNKEPLMPF	11
			pp220	RERERIFNL	18
			pp220	YINQALHEL	–

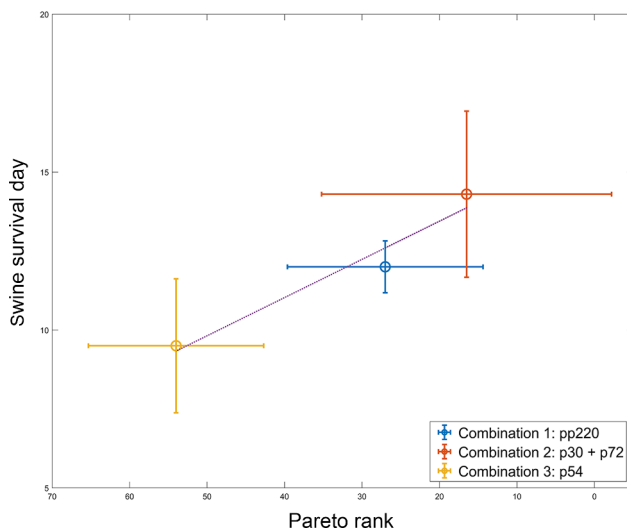
The rank of the evaluated epitope was the highest Pareto rank in the selected epitopes. There were 17 and 24 selected epitopes in the 30 experimentally validated and 34 predicted epitopes of T or B cells

we observed that CD2v had the highest average rank of extended fragments and the highest proportion ( $n=46$ ) of the top 72 epitopes. These results are consistent to the findings of existing studies revealing that CD2v plays an important role in activating adaptive host immune responses (Jia et al. 2017). CD2v may be a potential protein candidate in the ASFV vaccine due to its high ranking. p30, a phosphoprotein involved in ASFV entry, is synthesized in the early phase and continues to be synthesized during the late phase of viral infection. Although p30 is an antigenic and conserved structural protein, the immune response triggered by p30 alone is insufficient for antibody-mediated protection. However, combining it with other proteins, such as hemagglutinin, can increase humoral and cellular responses (Argilagué et al. 2012). In this work, p30 had the second highest average

rank, and p30 had two of the top 72 epitopes. It appears that p30 can be an important part of multi-epitope vaccine. p54 is important for the recruitment of envelope precursors to assembly factories and induces apoptosis during the early phase of infection (Hernaiz et al. 2004; Rodriguez et al. 1996). p54 is an antigenic structural protein that induces the production of specific antibodies. However, protein variability analysis demonstrated that there is a highly variable region in the C-terminus of the p54 protein. Accordingly, p54 had the lowest average rank among the five proteins in this work, and these show that the p54 epitope selection may depend on the target swine species and predominant MHC type. p72 is conserved and essential for viral icosahedron formation during viral infection (Cobbold and Wileman 1998); therefore, p72 has the characteristics of high antigenicity



**Fig. 4** Scatter points of experimental, predicted and selected epitopes in the top 30% of the fragments using PFAS. **(A)** Experimental epitopes. **(B)** Predicted epitopes



**Fig. 5** The correlation between Pareto ranks and swine survival days. Combinations 1, 2, and 3 contain four, six, and two epitopes, respectively. The vaccination experiments of combinations 1, 2, and 3 were performed in triplicate, quadruplicate, and triplicate, respectively. There was an  $R^2 = 0.95$  between mean ranks of the recombinant vaccines and swine survival days. Results are presented using mean  $\pm$  SD

and immunogenicity and is enriched and assembled in the ER during late-stage expression of infection. This study (Neilan et al. 2004) showed that p72 may produce high levels of p72-specific IgG antibodies, but there exists partial protection when using p72 epitopes alone. In this work, p72 had the third highest Pareto front and 10 p72 epitopes were selected from the top 72 epitopes, the data show that p72 may increase antibody production in multi-epitope vaccine. The ASFV polyprotein precursor pp220 is highly conserved in the viral genome, and pp220 is cleaved by proteases to produce the mature

virion proteins p150, p37, p14, and p34, which account for approximately 30% of the total viral protein mass and play an important role in the assembly process of the viral capsids and viral infection (Andres et al. 2002). In this work, pp220 had seven epitopes in the top three fronts, and pp220 had the second highest proportion ( $n=14$ ) among the top 72 promising epitopes. These results indicate that pp220 can be an important component of multi-epitope vaccine.

However, in the swine computational studies, the prediction models were trained mainly using human datasets and small amounts of animal data. In addition, the percentage of immune cell populations and the function of T cells differ between pigs and humans (Gerner et al. 2015; Rubic-Schneider et al. 2016), and variations in MHC polymorphisms induce different immune responses (Opriessnig et al. 2021). For computational prediction, identifying individual predictors is important to improve swine epitope prediction.

Although cross-species epitope prediction increases the uncertainty of the results, this study has demonstrated a relationship between Pareto rank and swine survival days based on the Pareto front approach. These findings support the hypothesis that accurate predictors with the Pareto front method may reduce vaccine development time and costs when applied to human vaccine development.

In this study, we use the Pareto front method to consider T and B cell immunogenicity simultaneously and ranked the epitopes with Pareto ranks. This procedure involved state-of-the-art computational methods and confirmed parameters. In addition, the evaluation of the experimental epitope ranks for the vaccination study had a significant coefficient of determination, demonstrating

**Table 3** The top 72 epitopes and their average ranks with the 16–20 amino acids

Mer type	Protein	Sequence	Average rank	Mer type	Protein	Sequence	Average rank
16mer	CD2v	KHVEEIESPPESNEE	1.00	17mer	CD2v	HVEEIESPPESNEEEQ	4.00
16mer	CD2v	HVEEIESPPESNEEE	1.00	17mer	p72	KLASQKDLVNEFPGLFV	4.00
16mer	CD2v	KKHVEEIESPPESNE	1.50	17mer	CD2v	SLRKRKKHVEEIESPPP	4.00
16mer	CD2v	YSRYQYNTPIIYMRPS	2.00	17mer	pp220	ATVSAIELEYDVKRRFY	4.00
16mer	p72	SQKDLVNEFPGLFVRQ	2.00	17mer	pp220	NPQSSERFEQYGRVFSR	4.00
16mer	pp220	TVSAIELEYDVKRRFY	2.50	17mer	CD2v	VEEIESPPESNEEEQC	4.00
16mer	CD2v	KRKKHVEEIESPPES	2.50	18mer	CD2v	KKHVEEIESPPESNEEE	1.25
16mer	CD2v	RKKHVEEIESPPESN	2.50	18mer	CD2v	RKKHVEEIESPPESNEE	1.75
16mer	CD2v	EPSPRELLPKPYSRY	2.50	18mer	CD2v	KRKKHVEEIESPPESNE	2.00
16mer	pp220	PQSSERFEQYGRVFSR	2.50	18mer	CD2v	RKRKKHVEEIESPPESN	2.75
16mer	CD2v	PYSRYQYNTPIIYMRP	2.50	18mer	CD2v	KHVEEIESPPESNEEEQ	3.25
16mer	p72	KLASQKDLVNEFPGLF	3.00	18mer	CD2v	HVEEIESPPESNEEEQC	3.25
16mer	pp220	SPLQIYKTLLEYLQHS	3.00	18mer	CD2v	LRKRKKHVEEIESPPES	3.50
16mer	p30	AQEEWNNMILHVLFEEE	3.00	18mer	p72	KLASQKDLVNEFPGLFVR	3.50
16mer	CD2v	RKRKKHVEEIESPPPE	3.00	18mer	CD2v	SLRKRKKHVEEIESPPPE	3.50
16mer	CD2v	HEPSPRELLPKPYSR	3.50	18mer	pp220	NNPQSSERFEQYGRVFSR	3.75
16mer	pp220	AFRTQLEDTRREVNNL	3.50	18mer	p72	LASQKDLVNEFPGLFVRQ	3.75
16mer	CD2v	EIESPPESNEEEQCQ	4.00	18mer	pp220	WKATVSAIELEYDVKRRF	4.00
16mer	pp220	WKATVSAIELEYDVKR	4.00	19mer	CD2v	RKKHVEEIESPPESNEEE	1.60
16mer	pp220	FRTQLEDTRREVNNLI	4.00	19mer	CD2v	KRKKHVEEIESPPESNEE	1.80
16mer	CD2v	SLRKRKKHVEEIESPP	4.00	19mer	CD2v	RKRKKHVEEIESPPESNE	2.40
16mer	pp220	PLQIYKTLLEYLQHSA	4.00	19mer	CD2v	KHVEEIESPPESNEEEQC	2.80
16mer	p30	KVIRAHNFIQTIYGTP	4.00	19mer	CD2v	KKHVEEIESPPESNEEEQ	3.00
16mer	p72	PGTKNAYRNLVYYCEY	4.00	19mer	CD2v	LRKRKKHVEEIESPPESN	3.20
16mer	p72	ASQKDLVNEFPGLFVR	4.00	19mer	p72	KLASQKDLVNEFPGLFVRQ	3.20
16mer	pp220	ATVSAIELEYDVKRRF	4.00	19mer	CD2v	SLRKRKKHVEEIESPPES	3.40
17mer	CD2v	KHVEEIESPPESNEEE	1.00	19mer	CD2v	HVEEIESPPESNEEEQCQ	4.00
17mer	CD2v	KKHVEEIESPPESNEE	1.33	19mer	pp220	WKATVSAIELEYDVKRRFY	4.00
17mer	CD2v	RKKHVEEIESPPESNE	2.00	20mer	CD2v	KRKKHVEEIESPPESNEEE	1.67
17mer	CD2v	KRKKHVEEIESPPESN	2.33	20mer	CD2v	RKRKKHVEEIESPPESNEE	2.17
17mer	CD2v	PYSRYQYNTPIIYMRPS	2.67	20mer	CD2v	KKHVEEIESPPESNEEEQC	2.67
17mer	CD2v	RKRKKHVEEIESPPES	3.00	20mer	CD2v	LRKRKKHVEEIESPPESNE	2.83
17mer	CD2v	HEPSPRELLPKPYSRY	3.00	20mer	CD2v	RKKHVEEIESPPESNEEEQ	3.00
17mer	pp220	SPLQIYKTLLEYLQHSA	3.33	20mer	CD2v	SLRKRKKHVEEIESPPESN	3.17
17mer	p72	ASQKDLVNEFPGLFVRQ	3.33	20mer	CD2v	KHVEEIESPPESNEEEQCQ	3.50
17mer	CD2v	LRKRKKHVEEIESPPPE	3.67	20mer	p72	IKLASQKDLVNEFPGLFVRQ	4.00

that the Pareto front method has effective screening efficiency. Finally, promising epitopes based on fragment extension and the peptide sequences with Pareto ranks were provided for biological experimental verification and confirmation. Overall, our study has proposed a computational prediction method based on the Pareto front method, provides Pareto rank of all fragments, promising epitopes, and may contribute to the development of recombinant multi-epitope vaccines for ASFV. The method may be used for human or cross-species promising epitope identification.

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13568-024-01749-6>.

Supplementary Tables

Supplementary Figures

### Acknowledgements

We would like to thank National Core Facility for Biopharmaceuticals (NCFB, 111-2740-B-492-001) and National Center for High-performance Computing (NCHC) of National Applied Research Laboratories (NARLabs) of Taiwan for providing computational resources and storage resources.

### Authors' contributions

TC, CC, SH, and PW conceived the study. TC and SH designed the experiments. TC, YH and FK performed the experiments and analyzed the data. CC and PW performed the formal analysis and technical assistance. TC drafted the original manuscript. SH and CC validated, supervised, and proof-read the manuscript. All authors read and approved the final manuscript.

### Funding

The work was supported by grants from National Science and Technology Council, Taiwan (110-2221-E-A49-099-MY3, 112-2740-B-400-005-), and was

financially supported by the "Center for Intelligent Drug Systems and Smart Bio-devices (IDS<sup>2</sup>B)" from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan.

#### Data availability

The data that download from prediction tools are available from the corresponding author upon request. The MATLAB codes of the Pareto front method are available at <https://github.com/NYCU-ICLAB/PFAS>. The validated epitopes and experimental results are available at supplementary information. The promising epitopes are available at supplementary information.

#### Declarations

#### Competing interests

Pei-Yin Wue and Chia-Jung Chang are employed by the Reber Genetics Co.

#### Ethical standards

This article does not contain any studies with animals performed by any of the authors.

#### Author details

<sup>1</sup>Institute of Molecular Medicine and Bioengineering, National Yang Ming Chiao Tung University, Hsinchu, Taiwan

<sup>2</sup>Department of Life Science, National Chung Hsing University, Taichung, Taiwan

<sup>3</sup>Department of Life Science, Genome and Systems Biology Degree Program, National Taiwan University, Taipei, Taiwan

<sup>4</sup>Institute of Bioinformatics and Systems Biology, National Yang Ming Chiao Tung University, Hsinchu, Taiwan

<sup>5</sup>Reber Genetics Co., Ltd. 13F, No. 160, Sec. 6, Minquan E. Rd., Neihu Dist, Taipei 114, Taiwan

<sup>6</sup>Department of Biological Science and Technology, National Yang Ming Chiao Tung University, Hsinchu, Taiwan

<sup>7</sup>Center for Intelligent Drug Systems and Smart Bio-devices (IDS<sup>2</sup>B), National Yang Ming Chiao Tung University, Hsinchu, Taiwan

<sup>8</sup>College of Health Sciences, Kaohsiung Medical University, Kaohsiung, Taiwan

Received: 19 December 2023 / Accepted: 13 August 2024

Published online: 31 August 2024

#### References

- Adamczyk-Poplawska M, Markowicz S, Jagustyn-Krynicka EK (2011) Proteomics for development of vaccine. *J Proteom* 74(12):2596–2616. <https://doi.org/10.1016/j.jprot.2011.01.019>
- Alejo A, Matamoros T, Guerra M, Andres G (2018) A Proteomic Atlas of the African Swine Fever Virus Particle. *J Virol*. <https://doi.org/10.1128/JVI.01293-18>
- Andres G, Alejo A, Salas J, Salas ML (2002) African swine fever virus polyproteins pp220 and pp62 assemble into the core shell. *J Virol* 76(24):12473–12482. <https://doi.org/10.1128/jvi.76.24.12473-12482.2002>
- Argilaguat JM, Perez-Martin E, Nofrarias M, Gallardo C, Accensi F, Lacasta A, Mora M, Ballester M, Galindo-Cardiel I, Lopez-Soria S, Escibano JM, Reche PA, Rodriguez F (2012) DNA vaccination partially protects against African swine fever virus lethal challenge in the absence of antibodies. *PLoS ONE* 7(9):e40942. <https://doi.org/10.1371/journal.pone.0040942>
- Bandrick M, Gutierrez AH, Desai P, Rincon G, Martin WD, Terry FE, De Groot AS, Foss DL (2020) 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* 223:110034. <https://doi.org/10.1016/j.vetimm.2020.110034>
- Bappy SS, Sultana S, Adhikari J, Mahmud S, Khan MA, Kibria KMK, Rahman MM, Shibly AZ (2021) Extensive immunoinformatics study for the prediction of novel peptide-based epitope vaccine with docking confirmation against envelope protein of Chikungunya virus: a computational biology approach. *J Biomol Struct Dyn* 39(4):1139–1154. <https://doi.org/10.1080/07391102.2020.1726815>
- Baratelli M, Morgan S, Hemmink JD, Reid E, Carr BV, Lefevre E, Montaner-Tarbes S, Charleston B, Fraile L, Tchilian E, Montoya M (2020) Identification of a newly conserved SLA-II Epitope in a structural protein of Swine Influenza Virus. *Front Immunol* 11:2083. <https://doi.org/10.3389/fimmu.2020.02083>
- Bhasin M, Raghava GP (2004) Analysis and prediction of affinity of TAP binding peptides using cascade SVM. *Protein Sci* 13(3):596–607. <https://doi.org/10.1110/ps.03373104>
- Blome S, Franke K, Beer M (2020) African swine fever - A review of current knowledge. *Virus Res* 287:198099. <https://doi.org/10.1016/j.virusres.2020.198099>
- Bosch-Camos L, Lopez E, Rodriguez F (2020) African swine fever vaccines: a promising work still in progress. *Porcine Health Manag* 6:17. <https://doi.org/10.1186/s40813-020-00154-2>
- Bosch-Camos L, Lopez E, Navas MJ, Pina-Pedrero S, Accensi F, Correa-Fiz F, Park C, Carrascal M, Dominguez J, Salas ML, Nikolin V, Collado J, Rodriguez F (2021) Identification of Promiscuous African Swine Fever Virus T-Cell Determinants Using a Multiple Technical Approach. *Vaccines (Basel)*. <https://doi.org/10.3390/vaccines9010029>
- Burmakina G, Malogolovkin A, Tulman ER, Xu W, Delhon G, Kolbasov D, Rock DL (2019) Identification of T-cell epitopes in African swine fever virus CD2v and C-type lectin proteins. *J Gen Virol* 100(2):259–265. <https://doi.org/10.1099/jgv.0.001195>
- Calis JJ, Maybeno M, Greenbaum JA, Weiskopf D, De Silva AD, Sette A, Kesmir C, Peters B (2013) Properties of MHC class I presented peptides that enhance immunogenicity. *PLoS Comput Biol* 9(10):e1003266. <https://doi.org/10.1371/journal.pcbi.1003266>
- Clem AS (2011) Fundamentals of vaccine immunology. *J Glob Infect Dis* 3(1):73–78. <https://doi.org/10.4103/0974-777X.77299>
- Cobbold C, Wileman T (1998) The major structural protein of African swine fever virus, p73, is packaged into large structures, indicative of viral capsid or matrix precursors, on the endoplasmic reticulum. *J Virol* 72(6):5215–5223. <https://doi.org/10.1128/JVI.72.6.5215-5223.1998>
- Dorigatti E, Schubert B (2020) Graph-theoretical formulation of the generalized epitope-based vaccine design problem. *PLoS Comput Biol* 16(10):e1008237. <https://doi.org/10.1371/journal.pcbi.1008237>
- Fan S, Wang Y, Wang X, Huang L, Zhang Y, Liu X, Zhu W (2018) Analysis of the affinity of influenza A virus protein epitopes for swine MHC I by a modified in vitro refolding method indicated cross-reactivity between swine and human MHC I specificities. *Immunogenetics* 70(10):671–680. <https://doi.org/10.1007/s00251-018-1070-6>
- Gao Z, Shao JJ, Zhang GL, Ge SD, Chang YY, Xiao L, Chang HY (2021) Development of an indirect ELISA to specifically detect antibodies against African swine fever virus: bioinformatics approaches. *Virology* 18(1):97. <https://doi.org/10.1186/s12985-021-01568-2>
- Garcia-Boronat M, Diez-Rivero CM, Reinherz EL, Reche PA (2008) PVS: a web server for protein sequence variability analysis tuned to facilitate conserved epitope discovery. *Nucleic Acids Res*. <https://doi.org/10.1093/nar/gkn211>
- Gaudreault NN, Richt JA (2019) Subunit Vaccine Approaches for African Swine Fever Virus. *Vaccines (Basel)*. <https://doi.org/10.3390/vaccines7020056>
- Gerner W, Talker SC, Koinig HC, Sedlak C, Mair KH, Saalmuller A (2015) Phenotypic and functional differentiation of porcine alpha T cells: current knowledge and available tools. *Mol Immunol* 66(1):3–13. <https://doi.org/10.1016/j.molimm.2014.10.025>
- Gomez-Puertas P, Rodriguez F, Oviedo JM, Ramiro-Ibanez F, Ruiz-Gonzalvo F, Alonso C, Escibano JM (1996) Neutralizing antibodies to different proteins of African swine fever virus inhibit both virus attachment and internalization. *J Virol* 70(8):5689–5694. <https://doi.org/10.1128/JVI.70.8.5689-5694.1996>
- Gomez-Puertas P, Rodriguez F, Oviedo JM, Brun A, Alonso C, Escibano JM (1998) The African swine fever virus proteins p54 and p30 are involved in two distinct steps of virus attachment and both contribute to the antibody-mediated protective immune response. *Virology* 243(2):461–471. <https://doi.org/10.1006/viro.1998.9068>
- Gul I, Hassan A, Muneeb JM, Akram T, Haq E, Shah RA, Ganai NA, Ahmad SM, Chikan NA, Shabir N (2022) A multi-epitope vaccine candidate against infectious bursal disease virus using immunoinformatics-based reverse vaccinology approach. *Front Vet Sci* 9:1116400. <https://doi.org/10.3389/fvets.2022.1116400>
- Guo F, Tang Y, Zhang W, Yuan H, Xiang J, Teng W, Lei A, Li R, Dai G (2022) DnaJ, a promising vaccine candidate against *Ureaplasma urealyticum* infection. *Appl Microbiol Biotechnol* 106(22):7643–7659. <https://doi.org/10.1007/s00253-022-12230-4>
- Gupta S, Ansari HR, Gautam A, Open Source Drug Discovery C, Raghava GP (2013) Identification of B-cell epitopes in an antigen for inducing specific class of antibodies. *Biol Direct* 8:27. <https://doi.org/10.1186/1745-6150-8-27>

- Hajjalibeigi A, Amani J, Gargari SLM (2021) Identification and evaluation of novel vaccine candidates against *Shigella flexneri* through reverse vaccinology approach. *Appl Microbiol Biotechnol* 105(3):1159–1173. <https://doi.org/10.1007/s00253-020-11054-4>
- Hernaiz B, Diaz-Gil G, Garcia-Gallo M, Ignacio Quetglas J, Rodriguez-Crespo I, Dixon L, Escribano JM, Alonso C (2004) The African swine fever virus dynein-binding protein p54 induces infected cell apoptosis. *FEBS Lett* 569(1–3):224–228. <https://doi.org/10.1016/j.febslet.2004.06.001>
- Ivanov V, Efremov EE, Novikov BV, Balyshchev VM, Tsibanov S, Kalinovsky T, Kolbasov DV, Niedzwiecki A, Rath M (2011) Vaccination with viral protein-mimicking peptides postpones mortality in domestic pigs infected by African swine fever virus. *Mol Med Rep* 4(3):395–401. <https://doi.org/10.3892/mmr.2011.454>
- Jach ME, Baj T, Juda M, Swider R, Mickowska B, Malm A (2020) Statistical evaluation of growth parameters in biofuel waste as a culture medium for improved production of single cell protein and amino acids by *Yarrowia Lipolytica*. *AMB Express* 10(1):35. <https://doi.org/10.1186/s13568-020-00968-x>
- Jancovich JK, Chapman D, Hansen DT, Robida MD, Loskutov A, Craciunescu F, Borovkov A, Kibler K, Goatley L, King K, Netherton CL, Taylor G, Jacobs B, Sykes K, Dixon LK (2018) Immunization of Pigs by DNA Prime and Recombinant Vaccinia Virus Boost To Identify and Rank African Swine Fever Virus Immunogenic and Protective Proteins. *J Virol*. <https://doi.org/10.1128/JVI.02219-17>
- Jia N, Ou Y, Pejsak Z, Zhang Y, Zhang J (2017) Roles of African swine fever virus structural proteins in viral infection. *J Vet Res* 61(2):135–143. <https://doi.org/10.1515/jvetres-2017-0017>
- Kessler C, Forth JH, Keil GM, Mettenleiter TC, Blome S, Karger A (2018) The intracellular proteome of African swine fever virus. *Sci Rep* 8(1):14714. <https://doi.org/10.1038/s41598-018-32985-z>
- Kibriá KMK, Faruque MO, Islam MSB, Ullah H, Mahmud S, Miah M, Saleh AA (2022) A conserved subunit vaccine designed against SARS-CoV-2 variants showed evidence in neutralizing the virus. *Appl Microbiol Biotechnol* 106(11):4091–4114. <https://doi.org/10.1007/s00253-022-11988-x>
- Larsen MV, Lundegaard C, Lamberth K, Buus S, Lund O, Nielsen M (2007) Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. *BMC Bioinformatics* 8:424. <https://doi.org/10.1186/1471-2105-8-424>
- Liu Q, Ma B, Qian N, Zhang F, Tan X, Lei J, Xiang Y (2019) Structure of the African swine fever virus major capsid protein p72. *Cell Res* 29(11):953–955. <https://doi.org/10.1038/s41422-019-0232-x>
- Lokhandwala S, Petrovan V, Popescu L, Sangewar N, Elijah C, Stoian A, Olcha M, Ennen L, Bray J, Bishop RP, Waghela SD, Sheahan M, Rowland RRR, Mwangi W (2019) Adenovirus-vectored African swine fever virus antigen cocktails are immunogenic but not protective against intranasal challenge with Georgia 2007/1 isolate. *Vet Microbiol* 235:10–20. <https://doi.org/10.1016/j.vetmic.2019.06.006>
- Lopera-Madrid J, Osorio JE, He Y, Xiang Z, Adams LG, Laughlin RC, Mwangi W, Subramanya S, Neilan J, Brake D, Burrage TG, Brown WC, Clavijo A, Bounpheng MA (2017) Safety and immunogenicity of mammalian cell derived and modified Vaccinia Ankara vectored African swine fever subunit antigens in swine. *Vet Immunol Immunopathol* 185:20–33. <https://doi.org/10.1016/j.vetimm.2017.01.004>
- Manavalan B, Govindaraj RG, Shin TH, Kim MO, Lee G (2018) iBCE-EL: a New Ensemble Learning Framework for Improved Linear B-Cell Epitope Prediction. *Front Immunol* 9:1695. <https://doi.org/10.3389/fimmu.2018.01695>
- Masoudi-Sobhanzadeh Y, Jafari B, Parvizpour S, Pourseif MM, Omid Y (2021) A novel multi-objective metaheuristic algorithm for protein-peptide docking and benchmarking on the LEADS-PEP dataset. *Comput Biol Med* 138:104896. <https://doi.org/10.1016/j.combiomed.2021.104896>
- Neilan JG, Zsak L, Lu Z, Burrage TG, Kutish GF, Rock DL (2004) Neutralizing antibodies to African swine fever virus proteins p30, p54, and p72 are not sufficient for antibody-mediated protection. *Virology* 319(2):337–342. <https://doi.org/10.1016/j.virol.2003.11.011>
- Opriessnig T, Mattei AA, Karuppanan AK, Halbur PG (2021) Future perspectives on swine viral vaccines: where are we headed? *Porcine Health Manag* 7(1):1. <https://doi.org/10.1186/s40813-020-00179-7>
- Reynisson B, Alvarez B, Paul S, Peters B, Nielsen M (2020) NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res* 48(W1):W449–W454. <https://doi.org/10.1093/nar/gkaa379>
- Rodriguez F, Ley V, Gomez-Puertas P, Garcia R, Rodriguez JF, Escribano JM (1996) The structural protein p54 is essential for African swine fever virus viability. *Virus Res* 40(2):161–167. [https://doi.org/10.1016/0168-1702\(95\)01268-0](https://doi.org/10.1016/0168-1702(95)01268-0)
- Ros-Lucas A, Correa-Fiz F, Bosch-Camos L, Rodriguez F, Alonso-Padilla J (2020) Computational Analysis of African Swine Fever Virus Protein Space for the Design of an Epitope-Based Vaccine Ensemble. *Pathogens*. <https://doi.org/10.3390/pathogens9121078>
- Rubic-Schneider T, Christen B, Brees D, Kammüller M (2016) Minipigs in Translational Immunology: a perspective. *Toxicol Pathol* 44(3):315–324. <https://doi.org/10.1177/0192623315621628>
- Saha S, Raghava GP (2006) Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins* 65(1):40–48. <https://doi.org/10.1002/prot.21078>
- Sanna G, Dei Giudici S, Bacciu D, Angioi PP, Giammaroli M, De Mia GM, Oggiano A (2017) Improved strategy for molecular characterization of African swine fever viruses from Sardinia, based on analysis of p30, CD2V and I73R/I329L variable regions. *Transbound Emerg Dis* 64(4):1280–1286. <https://doi.org/10.1111/tbed.12504>
- Singh H, Ansari HR, Raghava GP (2013) Improved method for linear B-cell epitope prediction using antigen's primary sequence. *PLoS ONE* 8(5):e62216. <https://doi.org/10.1371/journal.pone.0062216>
- Teklu T, Sun Y, Abid M, Luo Y, Qiu HJ (2020) Current status and evolving approaches to African swine fever vaccine development. *Transbound Emerg Dis* 67(2):529–542. <https://doi.org/10.1111/tbed.13364>
- Yu M, Morrissy CJ, Westbury HA (1996) Strong sequence conservation of African swine fever virus p72 protein provides the molecular basis for its antigenic stability. *Arch Virol* 141(9):1795–1802. <https://doi.org/10.1007/BF01718302>

## Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.