2021

Construction of an African Swine Fever Virus phagemid library through non-homologues random recombination

Author:

N.S. (Niels) van Heusden

6856241

Infection and Immunity Master Programme, Utrecht University

Supervised by:

Dr. Byron Martina

Artemis One Health Research Foundation, Delft

Abstract

African swine fever (ASF) is a haemorrhagic disease that affects swine in ongoing epidemics around the world. The etiologic agent of ASF, the African Swine Fever virus (ASFV), is a large multi-layered icosahedral dsDNA virus that encodes over 170 proteins, many of which have yet unknown functions. Partly because of its complexity, commercially available vaccines against ASF remain unavailable. Experimental attenuated-, subunit-, vector-, and DNA-vaccines have been developed previously, however, all fail to induce protective immunity outside of experimental conditions. Therefore, to aid the development of a multi-epitope chimeric vaccine against ASF, we attempted to construct a random multi-epitope library that can subsequently be screened for antigenic ASFV epitopes. To develop a randomized library of multi-epitope ASFV chimeras, gene fragment libraries were constructed using non-homologues recombination (NRR). Subsequently, we attempted to clone ASFV chimeras into phagemid vectors, to identify antigenic B-cell epitopes using phage-display.

Although this study did not successfully recombine ASFV-derived epitope fragments, we offer valuable insight into the development of randomized DNA libraries. Our results demonstrate that NRR can be used for the construction of randomly recombined dsDNA libraries. Additionally, we identified conformational B-cell epitopes on the surface of ASFV chimeras using *in silico* prediction tools. In conclusion, further research should optimize these methods to generate a diverse randomized multi-epitope chimeric library. Any antigenic multi-epitope chimera that arises from such study could contribute to the development of a multi-epitope-based ASFV vaccine.

Laymen summary

De Afrikaanse varkenspest (AVP) is een zeer besmettelijke, dodelijke virusziekte die wereldwijd een bedreiging vormt voor miljoenen gestalde varkens. Momenteel wordt de ziekte bestreden door besmette bedrijven te ruimen, wat tot veel sociaaleconomische schade leidt. Een effectieve oplossing voor AVP is de ontwikkeling van een effectief vaccin, maar dit blijft lastig omdat het onduidelijk is welke AVP epitopen immunogeen zijn (in staat zijn tot het induceren van een immuunreactie). Met als doel om nieuwe immunogene AVP epitopen te ontdekken, gebruikt dit onderzoek 'Non-Homologous Random Recombination (NRR)'. Deze methode isoleert en recombineert alle mogelijke AVP epitopen tot een bibliotheek die bestaat uit miljoenen combinaties, waarna gezocht kan worden naar de meest immunogene combinatie. Hierdoor worden nieuwe immunogene epitopen ontdekt aan de hand van een willekeurige benadering, in tegenstelling tot reguliere methodes die afhankelijk zijn van voorkennis over de pathogeen. Alhoewel wij er niet in slagen om deze bibliotheek te creëren, bewijzen we wel dat NRR in staat is om deze bibliotheek te produceren. Verder suggereert dit onderzoek dat NRR gebruikt kan worden voor de ontdekking van immunogene epitopen op basis van DNA sequenties.

Table of contents

Abstract 1
Laymen summary 1
Introduction 3
African swine fever virus 3
ASFV vaccines 3
Experimental vaccine development 4
CD2v 4
P22 5
P72 5
Materials and methods 6
Reagents 6
Gene constructs and vectors 6
Terminator hairpins 7
Q5 Site-directed mutagenesis 7
Gene amplification and isolation 7
Non-homologues random recombination of ASFV genes 7
TA cloning for sequencing 8
Construction of library pSEX81 phagemids 8
Sequence analysis and 3D modelling 8
Results 9
P72 site-directed mutagenesis 9
Fragmentation and recombination of ASFV genes 10
Sequencing to confirm recombination 11
Mapping back to reference sequences 12
In silico prediction of B-cell epitopes on ASFV library members 13
Discussion 14
References 17
Supplementary materials 20

Introduction

African swine fever virus

African swine fever (ASF) is a haemorrhagic disease that affects wild and domestic swine. The causative agent of ASF is the African swine fever virus (ASFV). In sub-Saharan Africa, where ASFV is endemic, ASFV is maintained through a sylvatic cycle between soft ticks and wild swine. Over the last decade, ASFV has reached Europe and Asia where it infects the domestic swine population resulting in devastating, continuous epidemics. Within this "domestic cycle", ASFV is transferred through direct contact with infected swine, carcasses or other pork products. Dependent on strain virulence and host susceptibility, symptoms can range from acute to chronic clinical forms of ASF. Whereas chronically infected individuals often remain asymptomatic, acute forms of ASF are characterized by haemorrhagic lesions resulting in high mortality (up to 100%) within ten days.

ASFV is the only member of the *Asfarviridae* family. *Asfarviridae* are a group of nucleoplasmic large DNA viruses (NCLDV), a phylum that primarily comprises structurally and genetically complex viruses. The ASFV genome consists of a single double-stranded DNA molecule of approximately 193 kilo base pairs, encoding a total of 170 proteins [1]. As of 2021, 24 different genotypes have been classified based on the 3' sequence of the B646L gene.

The virion consists of a concentric, multi-layered icosahedral structure with an overall diameter of about 200 nm (***Figure 1***) [2]. The central nucleoid is wrapped by a core shell, the inner envelope, capsid, and outer envelope. The nucleoid contains the genome and many DNA-binding nucleoproteins responsible for viral replication. The core shell, or inner capsid, consist of the processing products of two polyproteins that are crucial for capsid assembly; pp220 and pp62. The lipid bilayer of the inner envelope acts as a scaffold to support the core shell and the capsid, and contains proteins crucial for virion assembly, maturation, and entry. The capsid is mainly characterized by P72, the major structural ASFV protein, and provides the virion its icosahedral structure. Upon budding from the plasma membrane, the virion acquires the outer envelope containing CD2v, a protein that determines pathogenicity but is non-essential for infectivity [3–5].

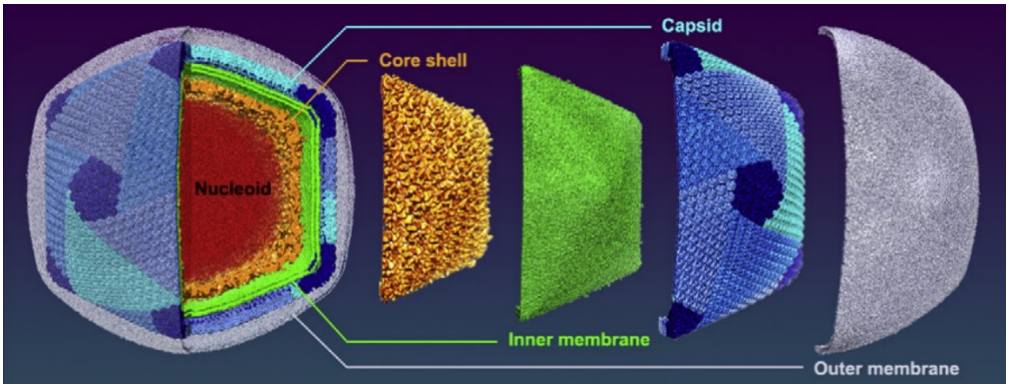


Figure 1: The multi-layered icosahedral structure of ASFV. The four layers of the virion are depicted in different colours. The inner envelope is also known as the inner membrane. Figure was taken from [6].

ASFV vaccines

Although the virus is harmless to humans, outbreaks within the domestic cycle have a high socioeconomic impact. At present, commercially available vaccines against ASF remain unavailable. Development of an effective vaccine is hampered by multiple factors. Firstly, due to its complex multi-layered structure, the virion is stable and extremely resistant against physical and chemical stress.

Secondly, the lack of defined correlates of protection challenge the assessment of the effectiveness of experimental vaccines. Current research indicates that protection against ASFV requires an adequate adaptive immune response, including both antibody and T-cell mediated responses [7].

Experimental ASFV vaccines have taken multiple approaches, mainly either attenuated vaccines, vectored or subunit vaccines. Inactivated virus vaccines and live attenuated vaccines (LAVs), whether naturally or recombinantly attenuated, conferred solid protective immunity to small test groups under experimental conditions [8–10]. However, protective effects were generally limited to strains within the same genotype and failed against heterologous virus challenge. Notably, the use of LAVs for ASFV protection is inherently linked to multiple disadvantages. Previous anti-ASFV LAVs have reverted to virulence and caused significant mortality up to 50% [11,12]. Moreover, LAVs do not allow for the development of a discriminatory test, further hindering the effective monitoring of disease control. For these reasons, subunit-, vector-based, or DNA-vaccines offer an alternative as they pose less biosafety risks and can be used as vaccines for Differentiating Infected from Vaccinated Animals (DIVA). As of yet, these vaccines have shown significant immunogenicity after injection of antigen cocktails, however, only a limited number of studies have tested against viral challenge. While most studies only yield limited successes [8,13], others have accomplished complete protection against ASFV after immunization with a cocktail of eight antigens [14]. Throughout most studies, experimental subunit-, vector-, or DNA-vaccines vaccines have targeted structural ASFV proteins such as P30, P54, P72, CD2v, and pp62.

Experimental vaccine development

Subunit-, vector-, and DNA-vaccines critically depend on ASFV components that elicit protective immunity. Therefore, researchers have shown an increased interest in the identification of highly immunogenic ASFV epitopes. Commonly, novel immunogenic epitopes have been identified either through computational analysis [15], using genomic or proteomic data, or by experimentally screening large libraries of biomolecules for binding to targets of interest. For example, phage-displayed gene fragment libraries have been used for the identification of novel B-cell epitopes [16]. The success of this technique is determined by the diversity of the starting library, a more diverse starting library contains more potential binding material, which yields more antigenic epitopes. Traditionally, diverse starting libraries have been developed through random mutagenesis and homologous recombination. However, since the genome of ASFV is extremely large and contains many non-homologues stretches, these methods would yield starting libraries with restricted diversity.

To overcome this problem, this study set out to develop a highly diverse ASFV starting library using non-homologues random recombination (NRR). NRR was developed for the production of a library without homology or ordering requirements [17,18]. Herein, target DNA was randomly digested and subsequently randomly recombined to produce gene chimeras. Importantly, the amount of recombination events could be controlled by introducing oligonucleotide hairpin-loops, which also allowed for bidirectional cloning of library chimeras. With the aim of verifying the potential of NRR as method to develop an ASFV epitope library, this research performed NRR using three ASFV genes; CD2v, P22, and P72. All three genes have been mentioned in literature as potent ASFV immunogens [19].

CD2v

CD2v is an ASFV protein located on the outer envelope and resembles the cell adhesion molecule CD2 that is found on the surface of human T-cells. CD2v is thought to enhance viral infectivity by binding to host-cell Adaptor protein-1 during entry [5]. Interestingly, truncated versions of CD2v are found in attenuated ASFV strains. *In silico* research has predicted B- and T-cell epitopes on the surface of CD2v. Four potential T-cell epitopes are located on CD2v, two within the cytoplasmic domain and two on the

external domain [20]. The four major immunogenic B-cell epitopes were located on the external domain of CD2v [21].

P22

P22 is a structural protein that is important for virion stability and is located on the inner envelope. Although P22 interacts with several host proteins, it is not associated with viral replication or virulence in swine [22]. Nonetheless, computational analysis predicted conserved B-cell epitopes on the P22 protein, which supported the earlier observation that infected swine serum contains P22-specific antibodies [23,24]. Additionally, the existence of B-cell epitopes on P22 is corroborated by the finding that ASFV-infected host cells display P22 on the plasma membrane [25].

P72

P72, the major capsid protein, constitutes about 33% of the total virion mass and is used as a diagnostic marker for infection. High-resolution cryo-EM showed that P72 assembles as a trimer with a propeller-like cap that protrudes from the membrane [26]. Unsurprisingly, the P72-cap contains many predicted B-cell epitopes [23,27]. The existence of B-cell epitopes on P72 domains could be explained by the presence of 'naked' ASFV particles – virions that lack the outer envelope. Since P72 constitutes the majority of the capsid, and thus the antigenic surface of 'naked' particles, external domains of P72 could contain B-cell epitopes. Nevertheless, candidate vaccine formulations targeting P72 have failed at inducing protection against ASFV infection [28].

In this report, we describe the use of NRR for the recombination of three ASFV genes. Whilst this study fails to recombine ASFV genes, the results show that NRR can be used to randomly recombine dsDNA under slightly different experimental conditions. Ultimately, this study lays the groundwork for future development of a diverse ASFV-gene fragment library.

Materials and methods

The entire protocol, from gene fragmentation to cloning for sequencing, is displayed below in a schematic overview. Herein, only the three ASFV genes are subjected to NRR, however, we also performed NRR on ART#130 (a circular plasmid containing the P72 gene) to assess different starting materials.

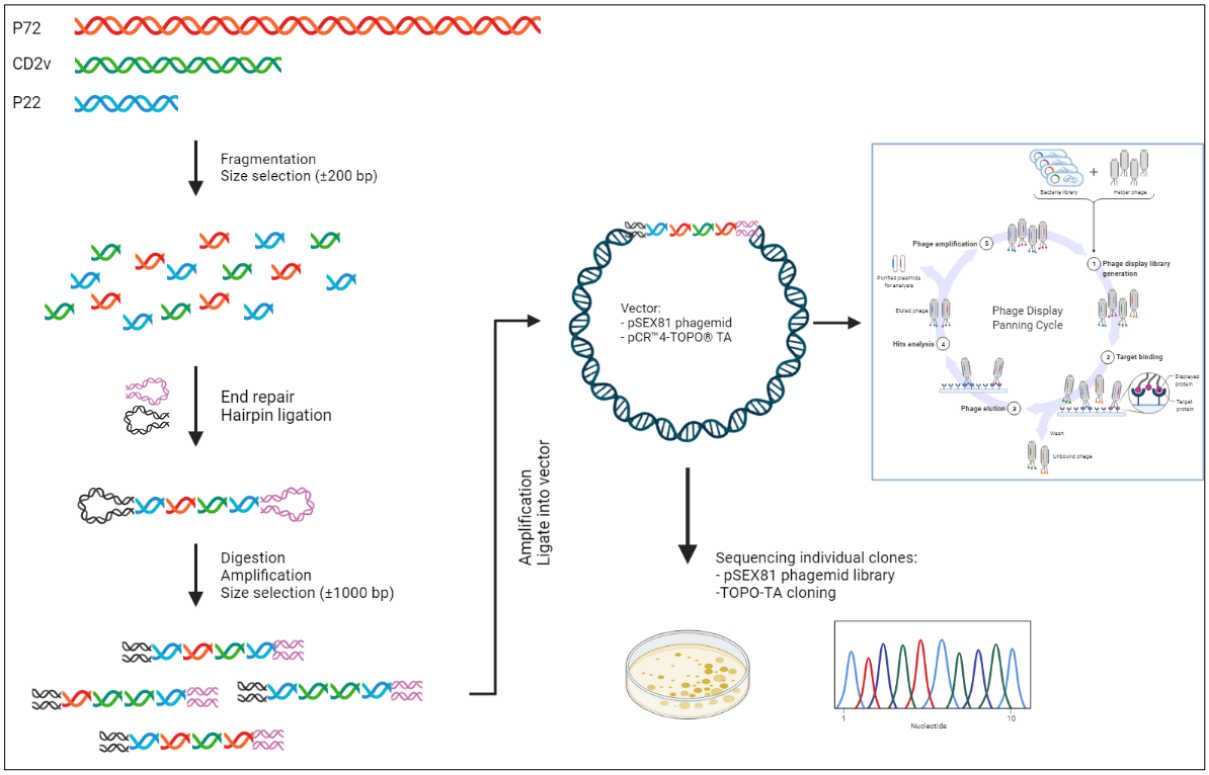


Figure 2: Schematic overview of non-homologues random recombination (NRR) for the development of a random ASFV chimeric library. Target dsDNA sequences are randomly digested after which 200-bp long fragments are extracted and blunt-ended. To randomly recombine four dsDNA fragments, a ligation mixture is prepared containing one hairpin oligonucleotide per two dsDNA fragments. Newly formed chimeric sequences are PCR amplified and separated to select for 1000-bp long chimeric sequences. Finally, ASFV chimeras are bidirectionally inserted into library vectors and used for sequence analysis or for phage display panning.

Reagents

Restriction enzymes, CutSmart® buffer, T4 DNA ligase, T4 DNA ligase reaction buffer, were purchased at New England biolabs. Recombinant *Taq* DNA polymerase, PCR buffer without $MgCl_2$, and $MgCl_2$ was ordered from ThermoFisher. PFU Ultra II Fusion HS DNA polymerase was purchased from Agilent. dNTP mix was ordered from Roche.

Gene constructs and vectors

Based on nucleotide sequences of circulating strains of ASFV, three genes were synthesized by BASEclear (Leiden, The Netherlands). CD2v (GenBank accession number: MK757459.1), P22 (GenBank accession number: KU641719.1), and P72 (GenBank accession number: MK554698.1) sequences were extracted from GenBank and optimized for expression in *Escherichia coli* using VectorBuilder. Additionally, the constructs contained KpnI and EcoRI restriction sites at the 5' and 3' ends, respectively, to allow for the purification of the ASFV genes. Constructs were delivered in pUC-SP vectors and cloned into DH5α competent cells (NEB). Single colonies were cultured in 300 mL LB medium containing Ampicillin (100 µg/mL) after which plasmid DNA was purified using the Nucleobond® Xtra Midi Plus kits (MACHERY-NAGEL), following the protocol provided by the

manufacturer. Subsequently, Nucleobond® Finalizers were used to concentrate the eluate into 200 µL stock volume.

Surface expression phagemid vector pSEX81 was used as recipient plasmid for recombined ASFV chimeras (**Supplementary figure 4)**. The vector was diluted and transferred into XL2-Blue Ultracompetent cells (Agilent Technologies) following the protocol provided by the manufacturer. Throughout this thesis, the vectors described here will be referred to as described by **Supplementary Table 2**.

Terminator hairpins

Two oligonucleotide hairpins **(Supplementary Table 1)**, Hp1 and Hp2, were modified from Bittker et al., [18] to use for the bidirectional ligation into phagemid vector pSEX81. The oligonucleotides were designed in such a way that it folds into a hairpin, the secondary structure of which was confirmed using folding tools **(Supplementary Figure 5)**. Both hairpins contain a SphI-HF site (italics) for the removal of hairpin ends and SalI-HF half-sites (**bold**) for digestion of hairpin dimers. Hp1 contains a HindIII restriction site and Hp2 an AvrII restriction site (underlined) to use for library insertion. Of note, none of the restriction sites used here were present on either CD2v, P22, or P72 genes. For PCR amplification of chimeric genes prior to pSEX81 insertion, hairpin-specific primers (Primer_Hp1 and Primer_Hp2) were used **(Supplementary Table 1)**. Finally, pSEX81 specific primers, pSEX81_Fw_Out and pSEX81_Rv_Out, were designed for sequence analysis of chimeric genes within the pSEX81 phagemid.

Q5 Site-directed mutagenesis

ART#130 contained multiple KpnI and EcoRI restriction sites, hindering PCR-based gene amplification. To create unique KpnI and EcoRI cleavage sites within the P72 construct, we substituted two nucleotides within the P72 construct (**Supplementary Figure 2**). The substitution was performed using the Q5 Site-Directed Mutagenesis Kit (New England Biolabs) and mutagenesis primers **(Supplementary Table 1)** were designed using NEBaseChanger v1.3.0. After selected colonies were brought into culture, ART#130.1 plasmids were extracted from DH5α cells using the NucleoSpin Plasmid EasyPure miniprep kit (MACHERY-NAGEL), according to the manufacturer's protocol.

Gene amplification and isolation

NRR required a high quantity of starting material, therefore, ASFV genes were amplified from vectors ART#128, ART#129, and ART#130.1 and isolated. To this end, we first PCR amplified the gene of interest using pUC-SP specific primers. To separate the gene of interest from vector specific sequences, PCR products were digested overnight using KpnI and EcoRI restriction enzymes (**Supplementary Figure 3**). Subsequently, the digestion products were separated on a 1% agarose gel and target bands were cut out and extracted from the gel.

Non-homologues random recombination of ASFV genes

Purified ASFV genes were fragmented using the NEBNext® dsDNA Fragmentase® kit (NEB), according to the provided protocol. In short, 5 µg of each target gene was incubated with Fragmentase mixture, including 1 µl of 200 mM $MgCl_2$, for one hour at 37 ^{o}C. The complete mixture was then loaded onto a 3% agarose gel and cutouts were made around 200 bp (±50 bp). Gel-cutouts were eluted in 30 µL elution buffer using the QIAquick® Gel Extraction Kit (Qiagen), according to manufacturer's protocol. Next, purified gene fragments from all three ASFV genes were pooled and blunt-ended using the NEBNext® End Repair Module (NEB), all according to provided instructions. End-repaired fragments were then purified using the QIAquick® PCR Purification Kit (Qiagen) after which the DNA concentration was measured on a Qubit 3 Fluorometer using the Qubit dsDNA BR Assay Kit

(ThermoFisher). To calculate the amount of moles of end-repaired dsDNA fragments, the following formula was used in which we assumed the length of dsDNA to be 200 bp:

$$Moles\ of\ dsDNA\ (mol) = \frac{Mass\ of\ dsDNA\ (g)}{Length\ of\ dsDNA\ (bp) * 617.96\ (\frac{\frac{g}{mol}}{bp})} + 36.04\ (\frac{g}{mol}).$$

Subsequently, a ligase reaction was set up to recombine four 200 bp fragments into one 1000 bp long chimeric ASFV gene. A typical reaction, in which 8.089 pmol was recovered after end-repair, required 4.04 pmol of hairpin oligonucleotides (mixture of 2.02 pmol of both Hp1 and Hp2). The ligation was performed using 2.5 µL T4 ligase and 1X T4 Ligase buffer, in a total volume of 50 µL, for 18 hours in 16 ^{o}C. Finally, the ligation products were cleaned using the QIAquick® PCR Purification Kit. Next, to remove hairpin loops and hairpin-dimers from the recombined ASFV genes, the ligation products were incubated overnight using SphI-HF and SalI-HF restriction enzymes. Following another QIAquick® PCR purification step, the recombined genes were then PCR amplified using Hp1_Primer and Hp2_Primer. PCR amplification was performed under PFU polymerase optimized conditions, with the annealing temperature set to 52 ^{o}C. The PCR product was loaded onto a 1% agarose gel and cutouts were made around the 1000 bp (± 100 bp) mark. Finally, the final gel cutouts were cleaned and stored until further usage, and will be referred to ASFV chimeras / library sequences. From here, ASFV chimeras were used for the construction of library plasmid and for TA cloning for sequencing.

TA cloning for sequencing

To sequence the recombined chimeras, we cloned the ASFV vector library into the pCR™4-TOPO® TA vector using the TOPO® TA Cloning® Kit for Sequencing (ThermoFisher). In short, $\frac{1}{30}^{th}$ of the ASFV library sequences were PCR amplified with Taq polymerase using Hp1 and Hp2 primers. Amplicons were subsequently inserted into the pCR™4-TOPO® TA vector and transformed into TOP10 cells. Cells were plated onto LB agar plates containing kanamycin (100 µg/µL) after which eight random colonies were picked. Plasmid DNA was miniprepped from eight ASFV and four ART#130 library clones.

Construction of library pSEX81 phagemids

To prepare the phagemid library, ASFV library chimeras were inserted into pSEX81 using double digestion and ligation. First, pSEX81 and ASFV library chimeras were digested by HindIII and AvrII restriction enzymes for 18 hours. Subsequently, digestion products were ligated into pSEX81 using an equimolar insert:vector ratio. Ligation products were then transformed into XL2-blue cells following the protocol provided by the manufacturer. A fraction of the culture was plated onto LB-ampicillin agar plates and incubated for 18 hours in 37 ^{o}C.

Sequence analysis and 3D modelling

Sanger sequencing was performed by MACROgen (Amsterdam, The Netherlands). Sequence analysis of ASFV chimeras was performed using standard M13 primers and mutagenesis sequencing using pUC-SP primers. Raw sequencing data was imported into CodonCode Aligner 9.0.2 and assembled to generate a consensus sequence. The consensus sequence was then aligned against either the ART#130.1 sequence or CD2v/P22/P72 genes using Blastn (standard algorithm parameters were applied). Finally, alignment results were displayed using SnapGene 5.3.1.

To predict structural B-cell epitopes from hypothetical PIII:ASFV fusion proteins, ASFV chimeras were bidirectionally cloned into pSEX81 vectors in SnapGene, using HindIII and AvrII restriction sites . From here, the PIII:ASFVcol5/8 fusion amino acid sequence was used to predict the protein structure using the transform restrained Rosetta (trRosetta) tool [29]. The top PDB template (minimal estimated TM-score: 0.717) was subsequently submitted into the DiscoTope 2.0 server for the prediction of structural B-cell epitopes, with a threshold set to -3.7 (suitable for most predictions) [30]. Amino acid residues that

were recognized as conformational B-cell epitope were visualized on the PIII:ASFV protein structure using PyMOL 2.4.1.

Results

P72 site-directed mutagenesis

Three ASFV gene constructs were flanked by EcoRI and KpnI restriction sites to allow for PCR-based amplification. However, both restriction sites in ART#130 were not unique, which reduces cloning efficiency (**Supplementary Figure 2**). To correct this, we used site-directed mutagenesis to induce a silent mutation in the excess, off-target, restriction sites. Following each subsequent mutagenesis round, isolated colonies were sequenced to confirm the desired mutation. Both EcoRI (A>G) and KpnI (T>A) sites were successfully mutated, without altering the corresponding amino acid **(Figure 3)**. Of note, because colony #2 contained a frame-shift mutation following EcoRI mutagenesis, we selected colony #1 for subsequent KpnI mutagenesis (**Figure 3B**). After both site-corrections, plasmid DNA from colony #5 was purified, resulting in corrected ART#130, ART#130.1 (**Supplementary Figure 3**).

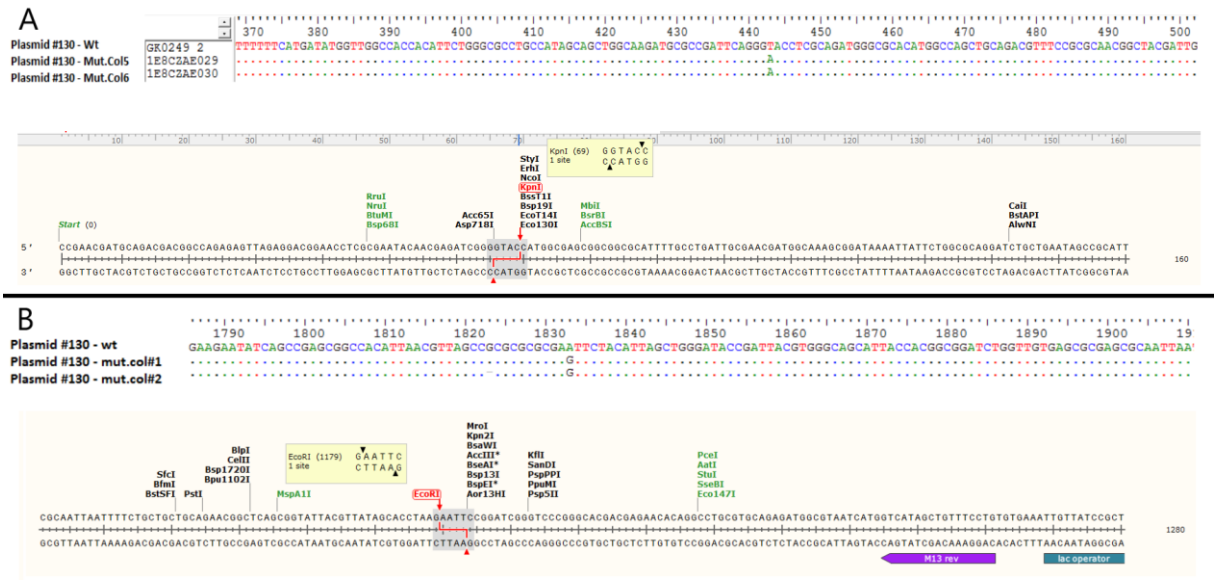


Figure 3: Q5 mutagenesis of ART#130 The ART#130 plasmid was mutated to create unique EcoRI and KpnI restriction sites, resulting in plasmid ART#130.1. **(A)** Mutagenesis to create a unique KpnI site yielded two successfully mutated colonies, MutCol5 and MutCol6. **(B)** Mutagenesis to create a unique EcoRI site yielded one successfully mutated colony, mut.col#1.

Fragmentation and recombination of ASFV genes

Originally, NRR was developed to produce a library of topologically rearranged functional enzymes with increased kinetic activity [18]. Here, we aimed to produce a library of non-homologously recombined ASFV genes, containing novel B-cell epitopes as a result of those rearrangements. To this end, CD2v, P22, and P72 genes were fragmented, end-repaired, and re-ligated into ASFV library chimeras of 1000 bp. Additionally, NRR was performed on ART#130 to assess the efficiency of the technique for different starting materials. Comparing ART#130 and ASFV gene fragmentation, we found that longer DNA molecules require less fragmentase incubation to obtain similarly sized fragments. Whereas ASFV genes required one hour **(Figure 4A)**, ART#130 required 20 minutes of fragmentation to yield 200 bp fragments (**Supplementary Figure 1**). Following gel extraction and blunt-end repair, fragments were recombined with hairpin oligonucleotides in a 2:1 molar ratio (fragments:hairpins). To isolate recombined ASFV chimeras, ligation products were amplified, separated on gel, and extracted based on length (**red rectangle, Figure 4B)**. Together, these results suggest that a random library can be produced through NRR of ASFV gene fragments.

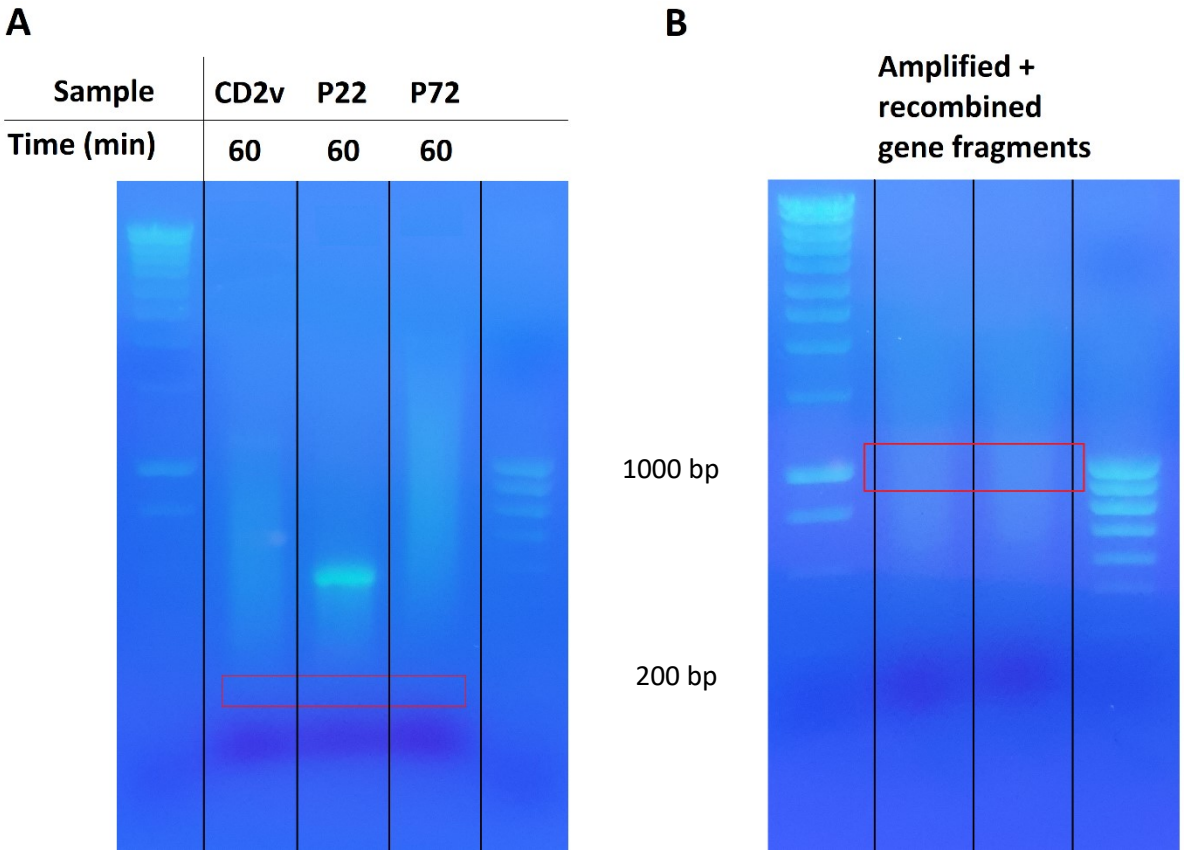


Figure 4: Isolation and recombination of ASFV gene fragments (A) Agarose electrophoresis of ASFV gene fragments after 1 hour of fragmentation. Lane 1 and 5 contain DNA ladders. Fragments of ± 200 bp were gel purified (red rectangle). **(B)** Agarose gel electrophoresis of recombined ASFV gene fragments amplified using hairpin specific primers. Lane 1 and 4 contain DNA ladders. Fragments of ± 1000 bp were gel purified (red rectangle) to obtain ASFV library chimeras.

Sequencing to confirm recombination

Following the development of ASFV and ART#130 library chimeras, we attempted to produce a phagemid library by inserting ASFV chimeras into pSEX81 phagemids. However, due to time constraints and experimental errors, we were unable to do so. Parallel to these attempts, we used conventional TA cloning for the insertion of ASFV and ART#130 library chimeras into the pCR-4 vector. Subsequently, eight ASFV library members and four ART#130 library members were sequenced of which two are displayed in **Figure 5**. ART#130 colony#3 showed two segments that aligned with the ART#130 sequence (green, **Figure 5A**). In addition to hairpin 1 and hairpin 2 alignments (purple and yellow), we observed two 34 bp hairpin dimers (red), resulting in multiple HindIII and AvrII restriction sites. The ASFV library member contained one segment that aligned with the CD2v gene (green, **Figure 5B**), which was flanked by hairpins 1 and 2.

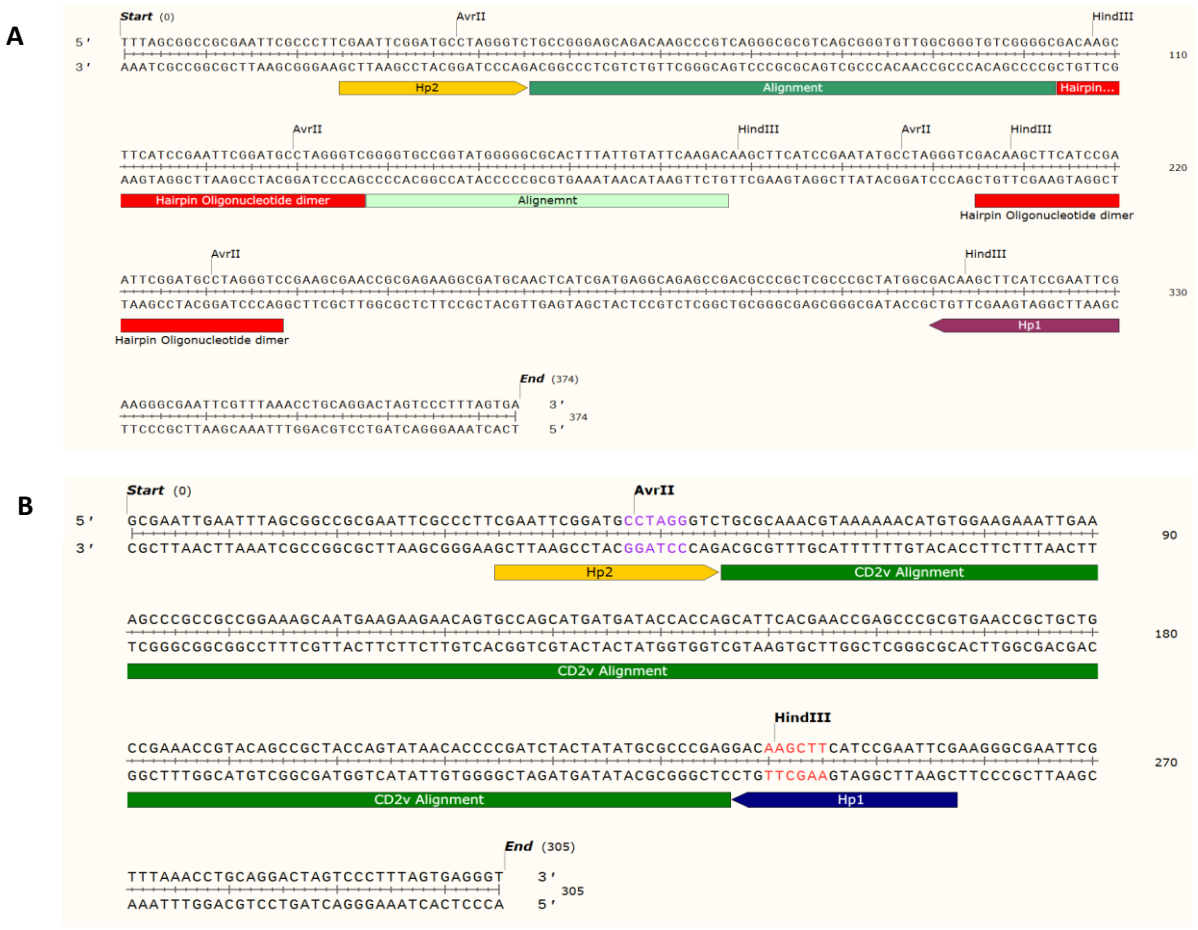


Figure 5: DNA sequences of NRR clones One ART#130 and one ASFV library chimera is displayed after TA sequencing using M13 primers. Hairpin sequences are displayed as Hp1 and Hp2, the direction of the arrow indicates the forward or reverse orientation. **(A)** The ART#130 library member (372 bp) contains two sequences that align with the ART#130 vector. Multiple restriction sites are formed. Hairpin dimers are also formed and ligated into the library chimera. **(B)** The ASFV library member (305 bp) contains one sequence that aligns with CD2v.

Mapping back to reference sequences

To evaluate library diversity, library members were aligned with either ART#130, or CD2v, P22, and P72 genes. The four members of the ART#130 library all contained sequences that originated from ART#130. Two out of the four library members, colonies 3 and 4, showed successful recombination of ART#130 fragments (green and blue) whereas colonies 1 and 8 did not contain recombined ART#130 sequences (black and orange) **(Figure 6A&B)**. Eight members of ASFV library were sequenced. Although no library members contained recombined sequences, all members did contain sequences that could be aligned with either CD2v, P22, or P72 **(Figure 6C&D)**. As opposed to the ART#130 library, all ASFV library members were flanked by both hairpins, without the inclusion of hairpin dimers. Altogether, these results implicate that NRR can be used for the recombination of multiple ASFV genes.

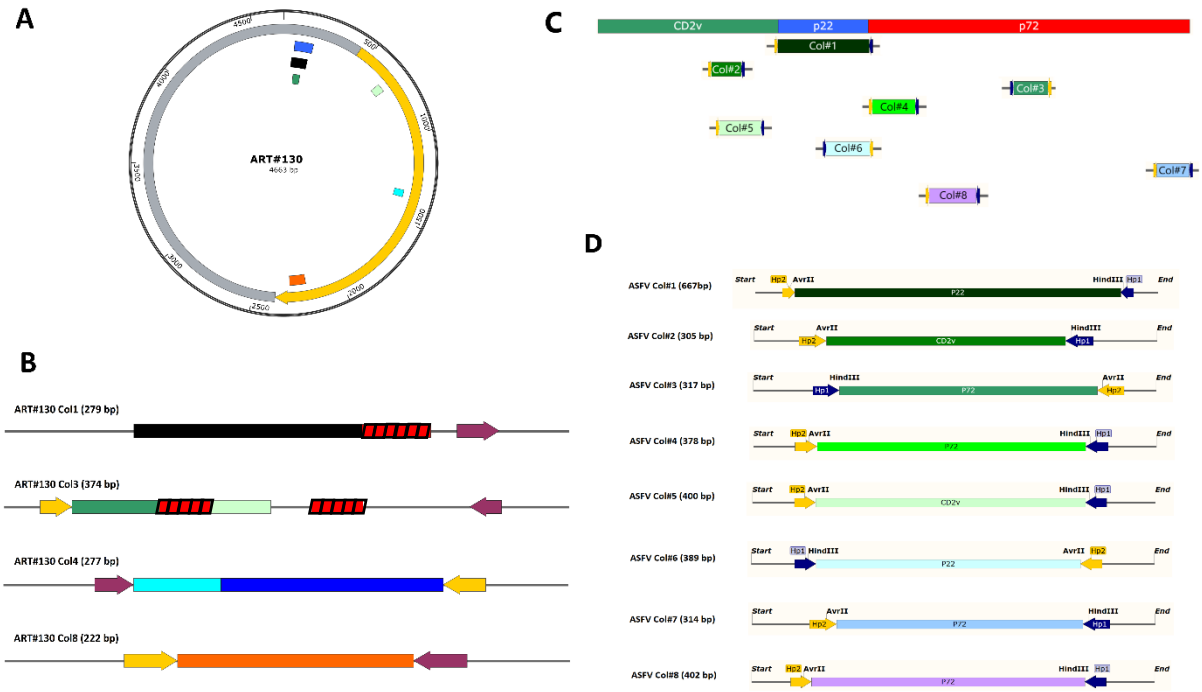


Figure 6: Sequence diversity of ART#130 and ASFV library members. Four ART#130 and eight ASFV chimeric library members were sequenced and aligned to ART#130 and CD2v, P22, P72, respectively. Two out of four ART#130 library members showed successful fragment recombination. The eight ASFV library members contained no fragment recombination. **(A)** The ART#130 sequence map (4663 bp) including the P72 gene (yellow) and the pUC-SP vector (grey) backbone. The colour-coded features indicate the location of an aligned ART#130 library member. **(B)** The feature view of four ART#130 chimeric library members (average sequence length: 288 bp). Coloured features correspond to the alignments in A. The orientation and location of hairpin 1 (yellow) and hairpin 2 (purple) sequences are indicated by the arrows. The striped red rectangles display hairpin dimer sequences. **(C)** The sequence map of CD2v (green), P22 (blue), and P72 (red) genes. The colour-coded features indicate the location of an aligned ASFV library member. **(D)** The feature view of four ART#130 chimeric library members (average sequence length: 397 bp). Coloured features correspond to the alignments in C. The orientation and location of hairpin 1 (dark blue) and hairpin 2 (yellow) sequences are indicated by the arrows.

In silico prediction of B-cell epitopes on ASFV library members

Since we were unable to generate pSEX81 colonies experimentally, we fused ASFV library chimeras to the PIII surface protein through in silico insertion into pSEX81 **(Supplementary Figure 4B)**. Based on the resulting ORF, the structure of the PIII fusion protein and corresponding conformational B-cell epitopes were predicted, using online prediction tools. The majority of ASFV library members (5/8) contained a STOP codon almost immediately after the fragment started. Library members 'Col4', 'Col5', and 'Col8' displayed longer peptides: 17, 55, and 20 amino acids, respectively (**Figure 7A**). The fusion proteins displayed by Col5 and Col8 library members both contain conformational B-cell epitopes on the ASFV residues (**Figure 7B&C**). Altogether, the results indicate that the ASFV library has the potential to generate B-cell epitopes.

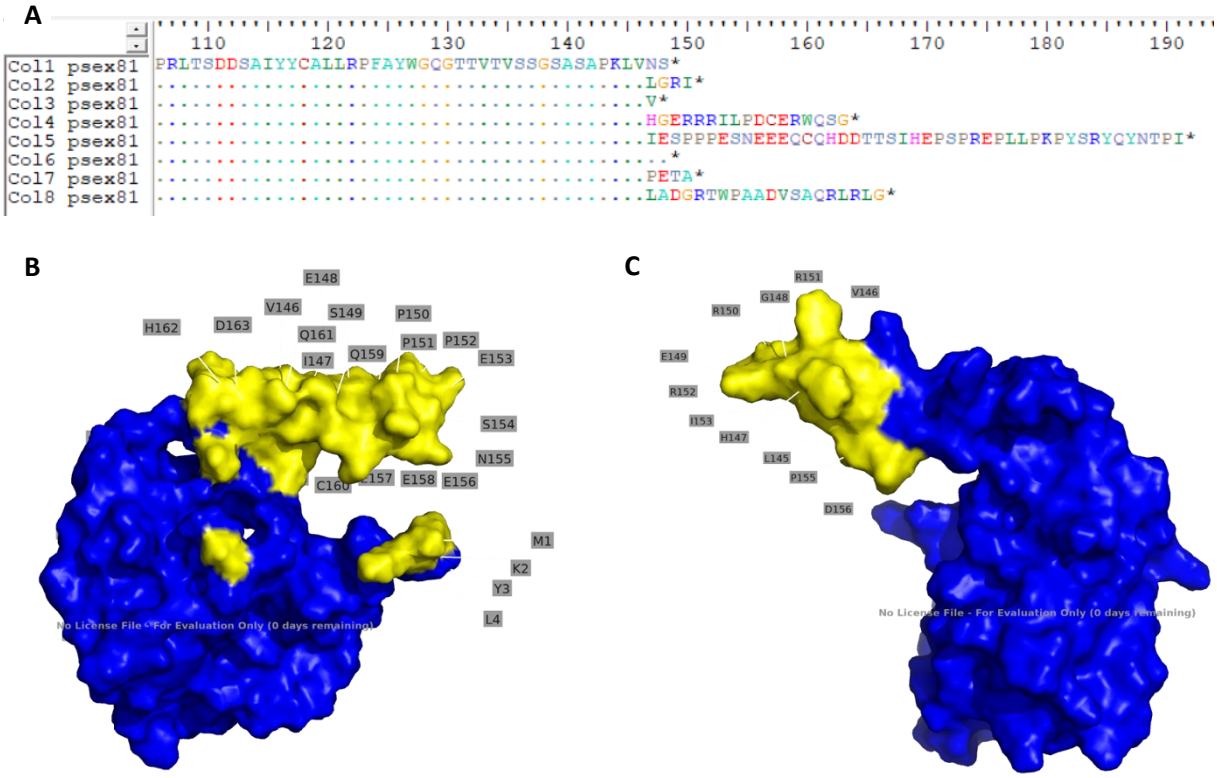


Figure 7: Location of conformational B-cell epitopes on PIII:ASFV fusion proteins. (A) Amino acid sequences of eight PIII:ASFV fusion proteins. From amino acid position 147 onwards, the peptide sequence of ASFV chimeras that are fused to the PIII protein, are displayed where a (*) indicates a STOP-codon. **(B) (C)** Surface view of the predicted conformational B-cell epitopes (yellow) of Col5 (B) and Col8 (C).

Discussion

The management of ASF disease relies on the development of an efficacious and safe ASFV vaccine. Although many studies have identified immunogenic antigens, it is yet unclear which are involved in protection. Multi-epitope-based chimeric vaccines are potential potent inducers of an immune response. To identify the specific epitopes that should be included, random libraries of multi-epitope chimeras can be screened against anti-ASFV antibodies, using phage-display. The development of an ASFV-derived phage-display library requires random recombination of DNA sequences. Here we aimed to establish a proof of principle that NRR can be utilized for the development of an ASFV-derived multi-epitope phagemid library, which may be suitable for development of a diagnostic test or prophylactic vaccine. Although a more rational approach is preferred over random fragmentation and recombination, NRR is appropriate when dealing with veterinary viruses, for which the immune system of the host (animal) has not extensively been investigated.

The results of the present study suggest that NRR is a suitable technique for the generation of dsDNA libraries. Although we did not successfully recombine ASFV-derived epitope fragments in this study, we offer valuable insight into the development of randomized DNA libraries. As expected, we observed that random fragmentation and recombination of ASFV epitopes results in many out-of-frame products, severely restricting library diversity. Altogether, considerably more work is required to produce and diversify epitope libraries using NRR.

Although we successfully induced recombination events in the ART#130 library, mainly three experimental challenges require further improvement. First, the method used here for the fragmentation of different templates was not optimal for the generation of 200-bp fragments. Despite multiple optimization experiments, we found that fragmentation efficiency was determined by the length of the dsDNA substrate. Fragmentation of smaller sequences, such as the P22 gene, was only partially successful. To avoid this problem, future research should synthesize, or fuse, a single chimeric CD2v/P22/P72 gene that can be subjected to fragmentation. Second, the re-ligation of 200-bp fragments did not result in any recombination event within the ASFV library, despite the similar recombination mixtures for both ART#130 and ASFV libraries. Although only 1000-bp chimeras were isolated in the subsequent step, which should have selected for library chimeras with four fragments, none of the colonies contained chimeras that underwent recombination. Third, 34-bp long hairpin heteroduplexes were found in two of the four ART#130 library members. Sequence analysis suggested that dimerization occurred based on homology in the 5'-CGAATTCG-3' sequence during PCR amplification, following the hairpin-loop digestion step. Consequently, multiple AvrII and HindIII restriction sites were introduced to library members, restricting the insert from subsequent phagemid insertion.

NRR is used to create sequence diversity from genes without sequence homology, in a random manner. However, the diversity of NRR-derived libraries is inherently restricted by large numbers of non-functional members as result of (1) out-of-frame recombination events and (2) PCR-based amplification. Because of incorrect hairpin orientations (success rate: 0.25) and out-of-frame gene fragments (success rate: $(1/18)^n$), a library with n recombination events contains a small fraction (success rate: $2.83*10^{-6}$ if n=4) of protein chimeras that is analogous to the original ASFV protein, and thus relevant for library screening. Of note, our hairpin orientation success rate is significantly higher than 0.25, due to yet unknown reasons. To improve library efficiency, ORF-selection systems have been developed to enrich randomized screening libraries [31]. These systems select for in-frame library members based on enzymatic activity and are specifically available for the M13 phage display system [32], however, further research should identify an appropriate ORF-selection system for our ASFV library. Due to unspecified reasons, PCR amplification could lead to disproportional amplification of specific

library templates, depressing the library from lesser amplified sequences thus reducing diversity. Generally, the aim of PCR based amplification is to amplify library members in the proportion that corresponds to the proportion in the original pool. Despite the use of a single primer pair, it remains difficult to optimize and to verify that every template is amplified equally.

As mentioned above, a large fraction of library members will encode protein chimeras that are not homologous to native ASFV proteins, and irrelevant for library screening. *In silico* analysis of two ASFV library chimeras predicted the presence of conformational B-cell epitopes (**Figure *7***). Unfortunately, these epitopes were not derived from native ASFV protein fragments but arose from out-of-frame translation of CD2v and P72 genes (not shown). This suggests that non-analogous oligopeptides introduce false-positives, as they interfere with binding of properly displayed ASFV library chimeras. To control for the bias that is introduced by out-of-frame translation products, future studies should include native, non-shuffled ASFV proteins into the phage-display library.

As compared to full-length antigen vaccines, multi-epitope vaccine chimeras consist of highly immunogenic epitopes that are connected to through linker segments. The assumption is that the combined immunogenicity of specific epitopes, through the construction of multi-epitope chimeras, exceeds the immunogenicity elicited by full-length ASFV antigens. So far, however, only one proof-of-concept multi-epitope chimeric vaccine has been tested *in vivo* for immunogenicity [33]. Nevertheless, our research supports the early development of multi-epitope chimeric vaccines by eliminating the need for *in silico* prediction tools. Whereas a rational approach uses reverse vaccinology to construct multi-epitope chimeras, phage-display screening analyses all possible epitope-chimeras and identifies the most antigenic ones.

On the other hand, development and validation of diverse multi-epitope libraries can be labour intensive and time consuming. To accelerate the development of ASFV vaccines, *in silico* epitope prediction tools have been built by others. The most prominent example is iVAX, an online toolkit that facilitates the design of swine vaccines through prediction of T-cell epitopes based on binding to MHC class 1 and 2 [15,34]. Interestingly, iVAX allows for the modification of candidate antigens and elimination of regulatory T-cell epitopes, to increase immunogenicity. The humanized-version of iVAX has since been applied to provide T-cell epitopes for the development of a multi-epitope Q-Fever vaccine [35]. Reverse vaccinology also relies on structural biology, as many B-cell epitope prediction tools require structural knowledge of antigens. The recent release of AlphaFold 2, a machine-learning algorithm that predicts protein structures based on amino acid sequence, could provide a better analysis of pathogen surfaces [36,37]. Because the model predicts structures without the need for homologous templates, ASFV antigens with previously unknown structures can be screened for B-cell epitopes based on *in silico*-derived protein structure prediction. Quick analysis of pathogen surfaces can prove especially useful in the fight against emerging epidemics.

Taken together, these advancements raise questions regarding the viability of experimental epitope prediction methods. Currently, the main advantage of experimental screening is the inclusion of immunogenically relevant antibodies, since *in silico* methods solely predict epitopes that bind to universal antibodies. However, further development of immunoinformatic techniques might lead to the inclusion of data from highly antigenic antibodies, with the aim to validate predicted epitopes *in silico*. For now, the golden standard for epitope discovery remains to confirm sequences using experimental methods.

The scientific community has focussed on triggering humoral responses in the fight against ASFV, however, it is highly unlikely that antibodies alone will ever confer solid protection. Injection with the most immunogenic antigens did not achieve antibody-mediated protection against ASFV [24].

Additionally, antibody dependent enhancement has been observed as result of P72/CD2v injection, suggesting that the ASFV-specific antibody response might be detrimental [38]. Moreover, it is generally assumed that severe ASF symptoms are associated with immune-overactivation, which could be mediated by the overproduction of non-neutralizing antibodies [39]. Together, these findings suggest an immune enhancement mechanism that warrants further investigation, and may require the identification of less immunogenic ASFV antigens.

In summary, this research was unable to produce an ASFV-derived multi-epitope phagemid library that is ready for phage-display screening. Despite its exploratory nature, we demonstrated that NRR is a valid method for the creation of randomized multi-epitope libraries. Further research should optimize these methods for multiple ASFV genes, and implement functional ORF-selection systems to increase library efficiency. Ultimately, any antigenic multi-epitope chimera that arises from this study could contribute to the development of a multi-epitope-based ASFV vaccine.

References

1. Dixon, L. K., Chapman, D. A. G., Netherton, C. L. & Upton, C. African swine fever virus replication and genomics. *Virus Res.* **173**, 3–14 (2013).

2. Wang, N. *et al.* Architecture of African swine fever virus and implications for viral assembly. *Science (80-.).* **644**, 640–644 (2019).

3. Andrés, G., García-Escudero, R., Viñuela, E., Salas, M. L. & Rodríguez, J. M. African Swine Fever Virus Structural Protein pE120R Is Essential for Virus Transport from Assembly Sites to Plasma Membrane but Not for Infectivity. *J. Virol.* **75**, 6758–6768 (2001).

4. Karger, A. *et al.* An update on African swine fever virology. *Viruses* **11**, 1–14 (2019).

5. Pérez-Núñez, D. *et al.* CD2v interacts with Adaptor Protein AP-1 during African swine fever infection. *PLoS One* **10**, 1–19 (2015).

6. Xian, Y. & Xiao, C. The Structure of ASFV Advances the Fight against the Disease. *Trends Biochem. Sci.* **45**, 276–278 (2020).

7. Carlson, J. *et al.* Association of the host immune response with protection using a live attenuated African swine fever virus model. *Viruses* **8**, (2016).

8. Lokhandwala, S. *et al.* 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 (2019).

9. Chen, W. *et al.* A seven-gene-deleted African swine fever virus is safe and effective as a live attenuated vaccine in pigs. *Sci. China Life Sci.* **63**, 623–634 (2020).

10. Borca, M. V. *et al.* Asfv-g-Δi177l as an effective oral nasal vaccine against the eurasia strain of africa swine fever. *Viruses* **13**, 1–9 (2021).

11. Turlewicz-Podbielska, H., Kuriga, A., Niemyjski, R., Tarasiuk, G. & Pomorska-Mól, M. African Swine Fever Virus as a Difficult Opponent in the Fight for a Vaccine—Current Data. *Viruses* **13**, 1212 (2021).

12. Bøtner, A. *et al.* Appearance of acute PRRS-like symptoms in sow herds after vaccination with a modified live PRRS vaccine. *Vet. Rec.* **141**, 497–499 (1997).

13. Gaudreault, N. N. & Richt, J. A. Subunit Vaccine Approaches for ASFV. *Vaccines* **7**, (2019).

14. Goatley, L. C. *et al.* A pool of eight virally vectored African swine fever antigens protect pigs against fatal disease. *Vaccines* **8**, 1–25 (2020).

15. Moise, L. *et al.* New Immunoinformatics Tools for Swine: Designing Epitope-Driven Vaccines, Predicting Vaccine Efficacy, and Making Vaccines on Demand. *Front. Immunol.* **11**, (2020).

16. Pulzova, L. *et al.* Identification of B-cell epitopes of Borrelia burgdorferi outer surface protein C by screening a phage-displayed gene fragment library. *Microbiol. Immunol.* **60**, 669–677 (2016).

17. Bittker, J. A., Le, B. V. & Liu, D. R. Nucleic acid evolution and minimization by nonhomologous random recombination. *Nat. Biotechnol.* **20**, 1024–1029 (2002).

18. Bittker, J. A., Le, B. V., Liu, J. M. & Liu, D. R. Directed evolution of protein enzymes using nonhomologous random recombination. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 7011–7016 (2004).

19. Netherton, C. L. *et al.* Identification and immunogenicity of African swine fever virus antigens.

Front. Immunol. **10**, 1–21 (2019).

20. Burmakina, G. *et al.* Identification of T-cell epitopes in African swine fever virus CD2v and C-type lectin proteins. *J. Gen. Virol.* **100**, 259–265 (2019).

21. Mima, K. A., Katorkina, E. I., Katorkin, S. A. & Malogolovkin, A. S. In silico идентификация B- и T-клеточных эпитопов белка CD2v вируса африканской чумы свиней (African swine fever virus , Asfivirus , Asfarviridae). *Probl. Virol.* **65**, 103–112 (2020).

22. Vuono, E. A. *et al.* Evaluation of the Function of the ASFV KP177R Gene , Encoding for Structural Protein p22 , in the Process of Virus Replication and in Swine Virulence. *Viruses* **13**, (2021).

23. Ros-Lucas, A., Correa-Fiz, F., Bosch-Camós, L., Rodriguez, F. & Alonso-Padilla, J. Computational analysis of african swine fever virus protein space for the design of an epitope-based vaccine ensemble. *Pathogens* **9**, 1–19 (2020).

24. Neilan, J. G. *et al.* Neutralizing antibodies to African swine fever virus proteins p30, p54, and p72 are not sufficient for antibody-mediated protection. *Virology* **319**, 337–342 (2004).

25. Alejo, A., Matamoros, T., Guerra, M. & Andrés, G. A Proteomic Atlas of the African Swine Fever Virus Particle. *J. Virol.* **92**, 1–18 (2018).

26. Andrés, G., Charro, D., Matamoros, T., Dillard, R. S. & Abrescia, N. G. A. The cryo-EM structure of African swine fever virus unravels a unique architecture comprising two icosahedral protein capsids and two lipoprotein membranes. *J. Biol. Chem.* **295**, 1–12 (2020).

27. Liu, Q. *et al.* Structure of the African swine fever virus major capsid protein p72. *Cell Res.* **29**, 953–955 (2019).

28. Bosch-Camós, L., López, E. & Rodriguez, F. African swine fever vaccines: A promising work still in progress. *Porc. Heal. Manag.* **6**, 1–14 (2020).

29. Yang, J. *et al.* Improved protein structure prediction using predicted interresidue orientations. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 1496–1503 (2020).

30. Kringelum, J. V., Lundegaard, C., Lund, O. & Nielsen, M. Reliable B Cell Epitope Predictions: Impacts of Method Development and Improved Benchmarking. *PLoS Comput. Biol.* **8**, (2012).

31. Verma, V., Joshi, G., Gupta, A. & Chaudhary, V. K. An efficient ORF selection system for DNA fragment libraries based on split betalactamase complementation. *PLoS One* **15**, 1–22 (2020).

32. Hust, M. *et al.* Enrichment of open reading frames presented on bacteriophage M13 using Hyperphage. *Biotechniques* **41**, 335–342 (2006).

33. Agallou, M., Margaroni, M., Kotsakis, S. D. & Karagouni, E. A canine-directed chimeric multi-epitope vaccine induced protective immune responses in balb/c mice infected with leishmania infantum. *Vaccines* **8**, 1–35 (2020).

34. De Groot, A. S. *et al.* Better epitope discovery, precision immune engineering, and accelerated vaccine design using Immunoinformatics tools. *Front. Immunol.* **11**, 1–13 (2020).

35. Scholzen, A. *et al.* Promiscuous Coxiella burnetii CD4 Epitope Clusters Associated with Human Recall Responses Are Candidates for a Novel T-Cell Targeted Multi-Epitope Q Fever Vaccine. *Front. Immunol.* **10**, 1–22 (2019).

36. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* (2021). doi:10.1038/s41586-021-03819-2

37. Higgins, M. K. Can We AlphaFold Our Way Out of the Next Pandemic? *J. Mol. Biol.* 167093 (2021). doi:10.1016/j.jmb.2021.167093

38. Sunwoo, S. Y. *et al.* Dna-protein vaccination strategy does not protect from challenge with African swine fever virus Armenia 2007 strain. *Vaccines* **7**, (2019).

39. Blome, S., Gabriel, C. & Beer, M. Pathogenesis of African swine fever in domestic pigs and European wild boar. *Virus Res.* **173**, 122–130 (2013).

Supplementary materials

Supplementary Table 1. **Oligonucleotides.**

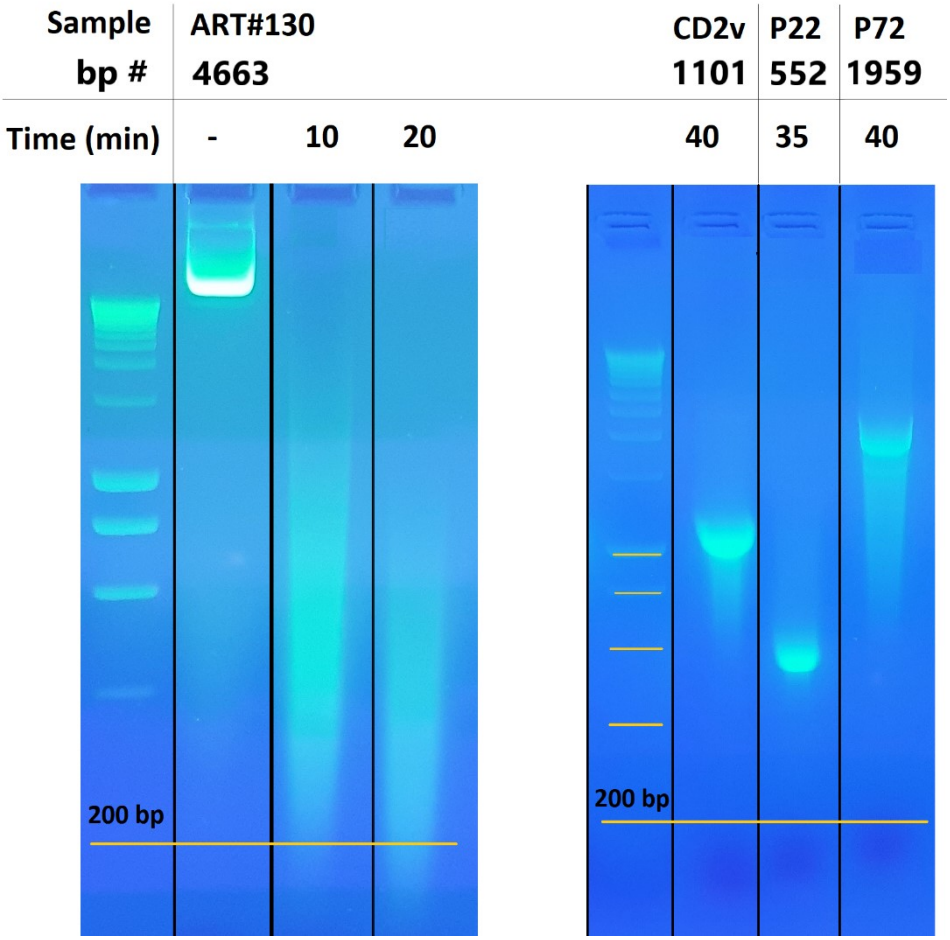
Name	Sequence 5' → 3'
MUT.ART#130-KpnI-Fw	CGATTCAGGGAACCTCGCAGATG
MUT.ART#130-KpnI-Rv	GCGCATCTTGCCAGCTGC
MUT.ART#130-EcoRI-Fw	GCGCGCGCGAGTTCTACATTA
MUT.ART#130-EcoRI-Rv	GGCTAACGTTAATGTGGCC
Hairpin 1 (Hp1)	**GAC**AAGCTTCATCCGAATTC*GCATGC*CCGGGCGCGCCCGG*GCATGC*GAATT CGGATGAAGCTT**GTC**
Hairpin 2 (Hp2)	**GAC**CCTAGGCATCCGAATTC*GCATGC*CCGGGCGCGCCCGG*GCATGC*GAATT CGGATGCCTAGG**GTC**
Primer_Hp1	CGAATTCGGATGAAGCTTGTC
Primer_Hp2	CGAATTCGGATGCCTAGGGTC
M13_Fw	GTAAAACGACGGCCAG
M13_Rv	CAGGAAACAGCTATGAC
pUC-SP_Fw_Out	GTTGGGTAACGCCAGGGTTTT
pUC-SP_Rv_Out	CGACTGGAAAGCGGGCAG
pSEX81_Fw_Out	ACTCGTCCCCAAAAGAACCG
pSEX81_Rv_Out	CATTCCACAGACAGCCCTCAT

Supplementary Table 2. **Plasmids.**

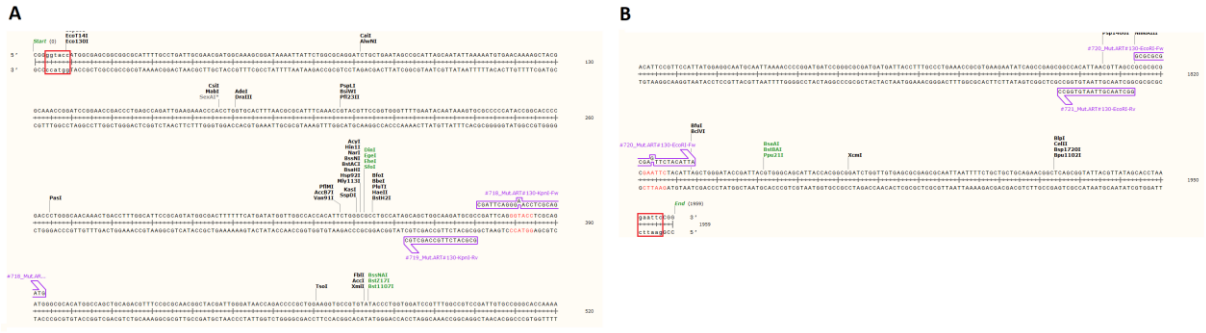
Name	Backbone	Insert	Resistance	Host cells
ART#128	pUC-SP	CD2v	Ampicillin	DH5α
ART#129	pUC-SP	P22	Ampicillin	DH5α
ART#130	pUC-SP	P72	Ampicillin	DH5α
ART#130.1	pUC-SP	P72 (Q5 - corrected)	Ampicillin	DH5α
pCR4_Library_ASFV	pCR™4-TOPO® TA	CD2v/P22/P72 chimeras	Kanamycin	TOP10
pCR4_Library_ART#130	pCR™4-TOPO® TA	CD2v/P22/P72 chimeras	Kanamycin	TOP10
pSEX81	pSEX81	-	Ampicillin	XL2-Blue

NEBNext® dsDNA Fragmentase® optimization

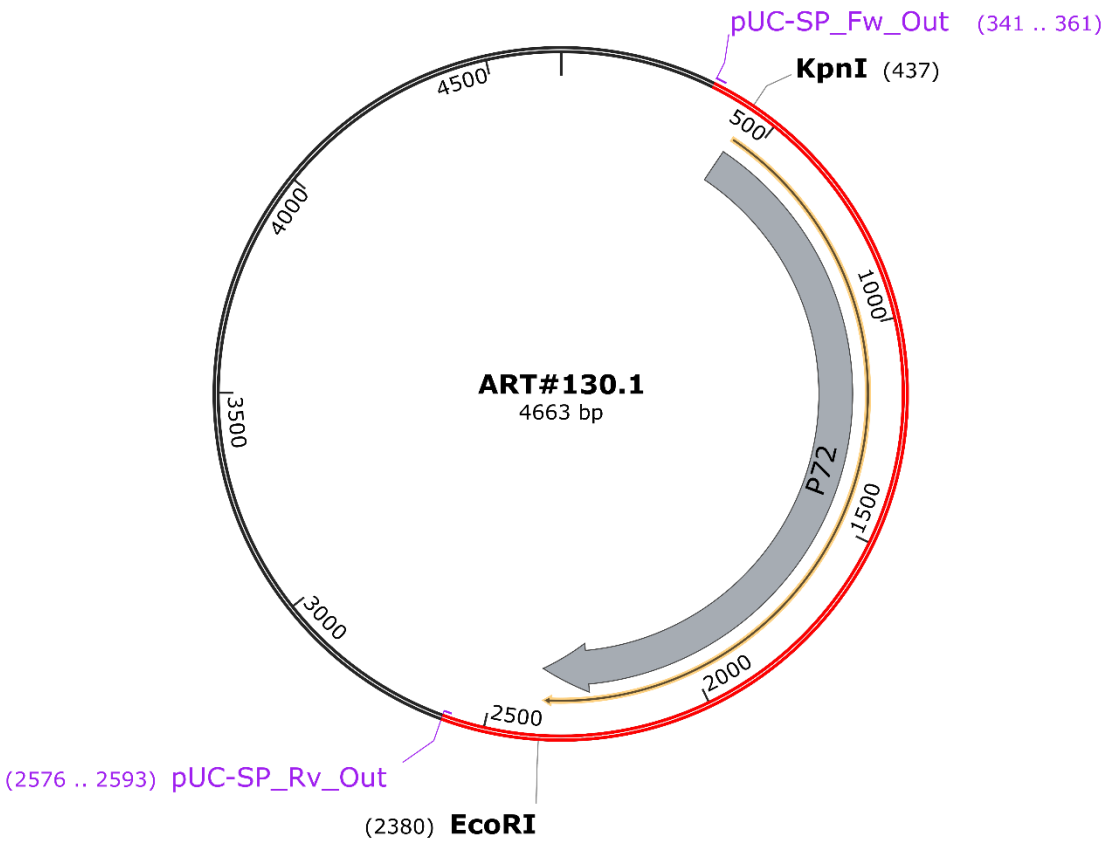
To prevent mutations that may arise as result of physical DNA shearing, we preferred enzymatical shearing using NEBNext® dsDNA Fragmentase® kit. To determine the optimal incubation parameters for our substrates, we incubated target sequences for different periods. Whereas, ART#130 (4663 bp) required 20 minutes of incubation to yield fragments of 200 bp, ASFV genes required more than 40 minutes. Altogether, the results suggest that, dependent on the length of the dsDNA input, different targets require different incubation parameters. Additionally, longer starting sequences (ART#130 and P72) were fragmented more efficiently than shorter ones (CD2v and P22).



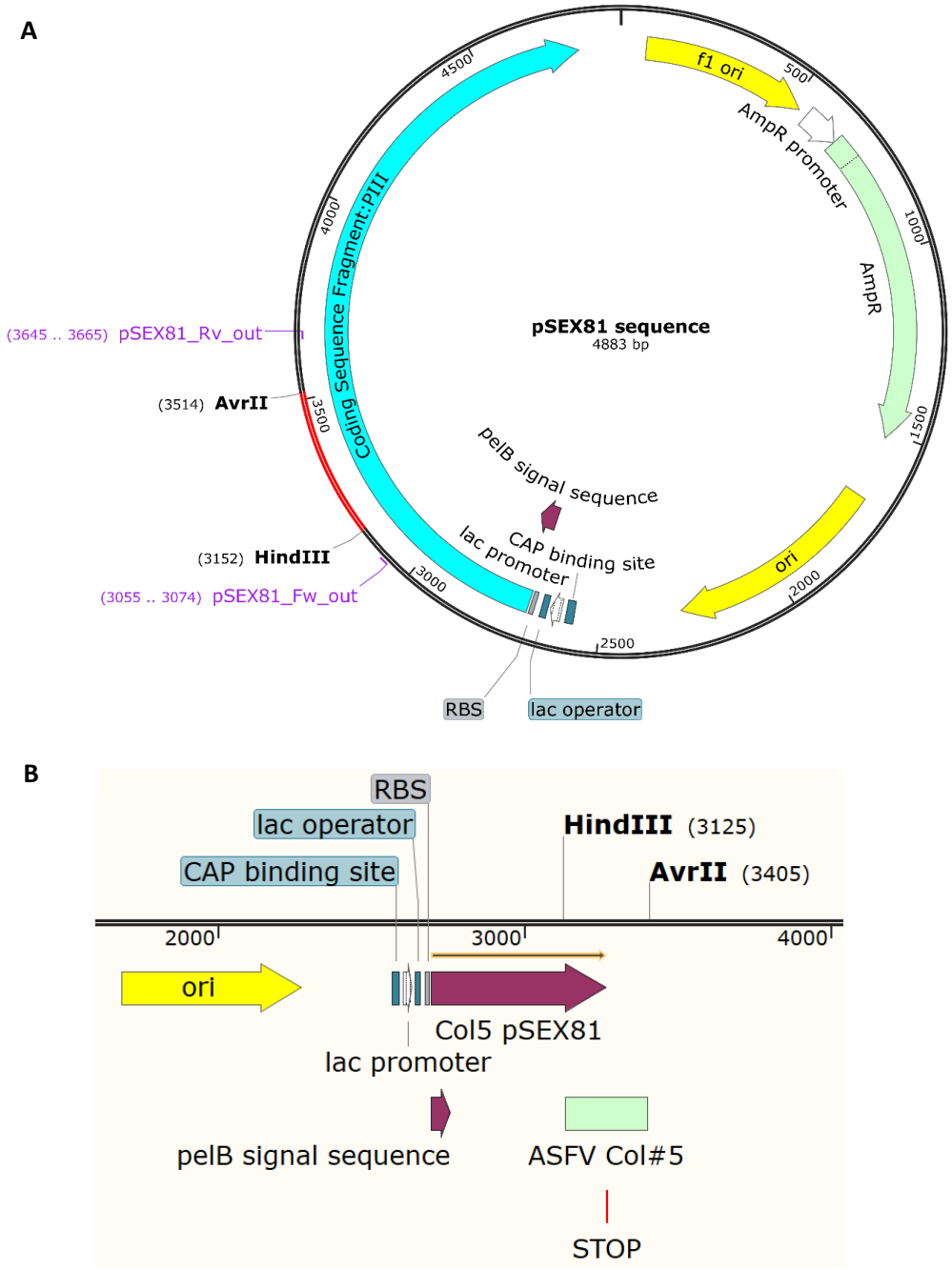
Supplementary Figure 1: Validation and optimalisation of the fragmentation protocol. Agarose electrophoresis of ART#130 and ASFV gene fragments. Identical amounts (5 µg) of dsDNA sample was fragmented using the NEBNext® dsDNA Fragmentase® kit. Non-fragmented ART#130 was used as negative fragmentation control.



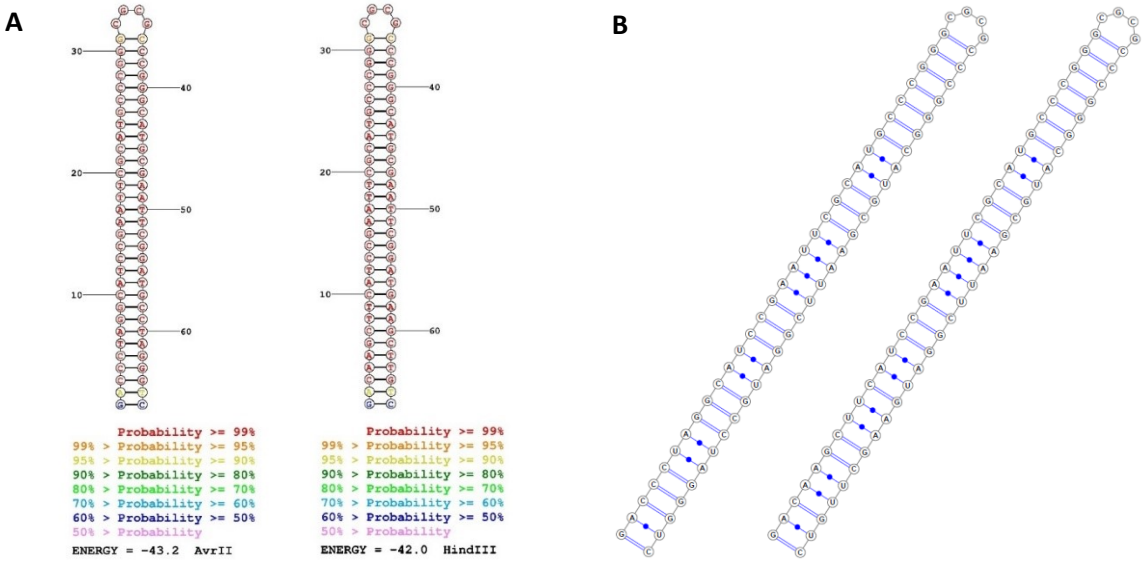
Supplementary Figure 2. **Restriction site mutagenesis in ART#130.** The 5' and 3' termini of the codon-optimized P72 sequence are displayed in (A) and (B), respectively. The red box displays the terminal restriction sites that should be maintained. The red text displays the restriction sites that should be disabled through site directed mutagenesis. The primers used for this technique are displayed in purple.



Supplementary Figure 3. **ART#130.** The ASFV gene constructs were synthesized and inserted into pUC-SP vectors. To purify P72 (grey) primers were used to amplify the sequence in red. Afterwards, restriction digestion was performed using EcoRI and KpnI enzymes after which the gene of interest was isolated on gel.



Supplementary Figure 4: Schematic map of the pSEX81 display vector. The phagemid vector pSEX81 is used for the expression of recombinant PIII fusion proteins on the surface of M13 phage. (A) The library members are cloned bidirectionally between HindIII and AvrII restriction sites (red), resulting in a continuous open reading frame Fragment:PIII (blue). To ensure the secretion of the fusion protein into the periplasmic space during virion assembly, the signal peptide sequence of bacterial pectate lyase (PelB) is fused to the Fragment:PIII protein. ColE1 (Yellow, ori) and the intergenic region of phage F1 (Yellow, F1 ori) are included in pSEX81 to allow for replication and packaging into phage particles. Finally, pSEX81-specific primers (purple) can be used for the analysis of the plasmid library. (B) An example of how ASFV fragments are inserted into pSEX81 to create PIII:ASFV fusion proteins. As can be seen, frame shift obstructs the translation of the entire ASFV protein fragment.



Supplementary Figure 5: **Secondary structure prediction of Hairpin oligonucleotides.** (A) The RNAstructure Web Server was used to predict the secondary structure of the hairpin oligonucleotides based on the lowest free energy structure . Default values were selected for the calculation. Both Hp1 (HindIII) and Hp2 (AvrII) fold towards a hairpin-like structure with high probability scores. (B) In accordance with the RNAstructure Web Server, IPknot web server displayed a similar secondary hairpin-like structure.