Virology 494 (2016) 178–189

Contents lists available at ScienceDirect

Virology

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

Recoding structural glycoprotein E2 in classical swine fever virus (CSFV) produces complete virus attenuation in swine and protects infected animals against disease

Lauro Velazquez-Salinas [a], Guillermo R. Risatti [b], Lauren G. Holinka [a], Vivian O'Donnell [a,b], Jolene Carlson [a,d], Marialexia Alfano [a], Luis. L. Rodriguez [a], Consuelo Carrillo [c], Douglas P. Gladue [a], Manuel V. Borca [a,*]

[a] Plum Island Animal Disease Center, ARS, USDA, Greenport, NY 11944, USA
[b] Department of Pathobiology and Veterinary Science, University of Connecticut, Storrs, CT 06269, USA
[c] Plum Island Animal Disease Center, DHS, Greenport, NY 11944, USA
[d] Biosecurity Research Institute and Department of Diagnostic Medicine and Pathobiology, College of Veterinary Medicine, Kansas State University, Manhattan, KS 66506, USA

ARTICLE INFO

Article history:
Received 28 January 2016
Returned to author for revisions
1 April 2016
Accepted 4 April 2016

Keywords:
Codon usage
Codon de-optimization
Protection
Vaccine development
Classical swine fever virus

ABSTRACT

Controlling classical swine fever (CSF) mainly involves vaccination with live attenuated vaccines (LAV). Experimental CSFV LAVs has been lately developed through reverse genetics using several different approaches. Here we present that codon de-optimization in the major CSFV structural glycoprotein E2 coding region, causes virus attenuation in swine. Four different mutated constructs (pCSFm1-pCSFm4) were designed using various mutational approaches based on the genetic background of the highly virulent strain Brescia (BICv). Three of these constructs produced infectious viruses (CSFm2v, CSFm3v, and CSFm4v). Animals infected with CSFm2v presented a reduced and extended viremia but did not display any CSF-related clinical signs. Animals that were infected with CSFm2v were protected against challenge with virulent parental BICv. This is the first report describing the development of an attenuated CSFV experimental vaccine by codon usage de-optimization, and one of the few examples of virus attenuation using this methodology that is assessed in a natural host.

Published by Elsevier Inc.

1. Introduction

Classical swine fever (CSF) is a highly contagious disease of swine. The etiological agent, CSF virus (CSFV), is an enveloped virus with a positive-sense, single-stranded RNA genome, classified as a member of the genus *Pestivirus* within the family *Flaviviridae* (Becher et al., 2003). The 12.5 kb CSFV genome contains a single open reading frame that encodes a 3898-amino-acid polyprotein and ultimately yields 11–12 final cleavage products (NH_2-Npro-C-E^{rns}-E1-E2-p7-NS2-NS3-NS4A-NS4B-NS5A-NS5B-COOH) through co- and post-translational processing of the polyprotein by cellular and viral proteases (Leifer et al., 2013). Structural components of the CSFV virion include the Core (C) protein and glycoproteins E^{rns}, E1 and E2. E1 and E2 are type I transmembrane proteins with an N-terminal ectodomain and a C-terminal hydrophobic anchor; E^{rns} loosely associates with the viral envelope (Thiel et al., 1991; Weiland et al., 1990; Weiland et al., 1999). E2 is the most immunogenic of the CSFV glycoproteins (Weiland et al., 1999; Konig et al., 1995; van Gennip et al., 2000), inducing neutralizing antibodies that provide protection against lethal CSFV challenge. Several studies have demonstrated the presence of genetic determinants of virulence within the E2 gene. Using reverse genetics, it has been demonstrated that amino acid substitutions introduced at specific positions within the E2 gene lead to partial or complete attenuation of virulent viruses in swine (Risatti et al., 2005a, 2005b, 2006, 2007a, 2007b; Van Gennip et al., 2004). Most of these attenuated viruses have been shown to

* Correspondence to: Plum Island Animal Disease Center, USDA/ARS/NAA, P.O. Box 848, Greenport, NY 11944-0848, USA.
E-mail addresses: lauro.velazquez@ars.usda.gov (L. Velazquez-Salinas), guillermo.risatti@uconn.edu (G.R. Risatti), lauren.holinka@ars.usda.gov (L.G. Holinka), vivian.odonnell@ars.usda.gov (V. O'Donnell), jolene.carlson@ars.usda.gov (J. Carlson), marialexia.alfano@ars.usda.gov (M. Alfano), luis.rodriguez@ars.usda.gov (Luis.L. Rodriguez), consuelo.carrillo@aphis.usda.gov (C. Carrillo), douglas.gladue@ars.usda.gov (D.P. Gladue), manuel.borca@ars.usda.gov (M.V. Borca).

http://dx.doi.org/10.1016/j.virol.2016.04.007
0042-6822/Published by Elsevier Inc.

induce protection against the virulent parental strains, opening the possibility to use this approach to develop experimental vaccine strains.

An alternative approach to produce live attenuated viruses consists of the systematic introduction of synonymous mutations by manipulating the codon usage bias or the pair codon usage. Codon usage bias refers to the phenomenon where specific codons are used more often than other synonymous codons during translation of genes, the extent of which varies within and among species. This variation is typically manifested by a balance between mutational and transitional pressure (Behura and Severson, 2013). In general the overall extent of codon usage bias in RNA viruses is low; the strong correlation between base and dinucleotide composition and codon usage bias suggests that mutational pressure rather than transitional pressure might be the most important factor shaping codon usage bias in viral RNA genomes (Jenkins and Holmes, 2003). Despite low codon usage in RNA viruses, the experimental de-optimization of codon usage bias or pair codon usage carried out in different viral genomes such as influenza A virus (Baker et al., 2015; Nogales et al., 2014; Luan et al., 2009), arenavirus (Cheng et al., 2015), respiratory syncytial virus (Meng et al., 2014; Le Nouen et al., 2014), porcine reproductive and respiratory syndrome virus (Ni et al., 2014), poliovirus (Mueller et al., 2006; Burns et al., 2006) and vesicular stomatitis virus (Wang et al., 2015), has produced not only viral attenuation in cell culture and animals, but also protection when animals have been challenged with the parental strain. Although the exact mechanism of action remains unknown, the use of this approach might represent a good platform for the development of live attenuated vaccines.

Here we present the development of several different constructs using mutational approaches based on the genetic background of the highly virulent CSFV strain Brescia (BICv). Most of the mutant constructs led to the production of infectious viruses and, interestingly, one of the viruses was completely attenuated in swine. Virus attenuation correlated with a decreased ability to replicate in primary swine macrophage cell cultures, the main target cell type during the infection *in vivo*. Animals infected with the attenuated strain were protected against the presentation of clinical CSF and infection after the challenge with the virulent parental BICv.

2. Materials and methods

2.1. Viruses, antibodies and cells

Swine kidney cells (SK6) (Terpstra et al., 1990), free of BVDV, were cultured in Dulbecco's minimal essential media (DMEM) (Gibco, Grand Island, NY) with 10% fetal calf serum (FCS) (Atlas Biologicals, Fort Collins, CO). CSFV strain Brescia was propagated in SK6 cells and was used for the construction of an infectious cDNA clone (Risatti et al., 2005). Growth kinetics was assessed using either SK6 cells (Risatti et al., 2005) or primary swine macrophage cell cultures prepared as described by Zsak et al. Zsak et al., (1996). Titration of CSFV from clinical samples was performed using SK6 cells in 96-well plates (Costar, Cambridge, MA). After 4 days in culture, viral infectivity was assessed using an immunoperoxidase assay utilizing the CSFmonoclonal antibody WH303 (mAb WH303) (Edwards et al., 1991) and the Vectastain ABC kit (Vector Laboratories, Burlingame, CA) (Risatti et al., 2005). Titers were calculated according to the method of Reed and Muench (Reed and Muench, 1938) and expressed as $TCID_{50}$/ml. As performed, test sensitivity was $\geq \log_{10}$ 1.8 $TCID_{50}$/ml.

2.2. Codon usage de-optimization strategy

Codon usage de-optimization was carried out *in silico* using the web server Optimizer (http://genomes.urv.es/OPTIMIZER/) (Puigbo et al., 2007). The procedure was conducted along the 373 codons of the glycoprotein E2 coding region of CSFV (nucleotides 2435-3553 from CSFV isolate Brescia; GenBank Accession number: AY578687.1). The reference host codon usage dataset was obtained from the Codon Usage Database (Nakamura et al., 2000). It included 1,168,059 codons contained in 2953 coding regions, which represent the most abundant codons for each amino acid in these regions in the species *Sus scrofa*. Two different algorithms were used to design the mutant viruses. First, the mutant CSFm1 was designed using the one amino acid-one codon algorithm, which is based on the codon adaptation index (CAI = 1) theory (Pesole et al., 1988). Basically, all the E2 codons that encode the same amino acid were substituted by the less commonly used synonymous codon in the reference host dataset (Puigbo et al., 2007). Second, the mutant CSFm2 was designed using the guided random method that is based on the Monte Carlo algorithm. Essentially based on the reference dataset, a new pool of codons was created for each amino acid of E2 that was equal to the number found in the target sequence, and that was in accord with their low occurrence in the reference dataset, so that the de-optimized sequence contained codons in the opposite proportions in which they are found in the adjusted reference dataset (Jayaraj et al., 2005). Mutant CSFm3 was de-optimized just in the first 123 codons of the E2 coding region (nucleotides 2435-2803), using (as with CSFm1) the one amino acid-one codon algorithm (Pesole et al., 1988).

With mutant CSFm4, we wanted to determine the potential effect of modifying synonymous sites which have remained conserved during the evolution of natural populations of CSFV in the E2 coding region. For this propose, we conducted a sequence alignment considering full-length sequences available in the GenBank database corresponding to all CSFV genotypes. Twelve codons were found highly conserved among all different genotypes, and based on low frequencies of codon usage found in the reference dataset for the respective amino acids, synonymous mutations were manually introduced at the corresponding nucleotide positions. Nucleotide sequences for each of the four CSFm constructs, compared with the parental virus, is provided in Fig. 1.

2.3. Development of mutant viruses

A full-length infectious cDNA clone (IC) of the BICv virus (pBIC) (Risatti et al., 2005) was used as a DNA template in which mutations were introduced to alter codon usage as described earlier. Areas of CSFV glycoprotein E2 containing the desired nucleotide changes were commercially synthesized (Epoch, Missouri City, TX) and further directionally cloned into pBIC using the Infusion (In-Fusion cloning, Takara-Clontech, Mountain View, CA) methodology as described by the manufacturer using the appropriate primers in the 5′ and 3′ ends of the cloned construct (CSFm1: 5′ CTGGGGCA-CAAGGCCGGCTAGCGTGATAAGAAGATCATCG 3′ and 5′ GCCTAACTG-TAGACCCGCGGCTAATTGTTCCGTT AAT 3′; CSFm2: 5′ CTGGGGCA-CAAGGCCGGCTAGCTTGCAAAGAAGATCATC G 3′ and 5′ GCCTAACTG-TAGACCCGCGGCCAACTGCTCCGTTAAGACAA 3′; CSFm3: 5′ CTGGGGCACAAGGCCGGCTAGCGTGTAAAGAAGATCATCGTTA 3′ and 5′ GCCTAACTGTAGACCCGCGGCGAGTTGTTCTGTTAGAACT 3′; CSFm4: 5′ CTGGGGCACAAGGCCGGCTAGCCTGCAAGGAAGATCACA 3′ and 5′ GCCTAA CTGTAGACCCGCGGCGAGTTGTTCTGTTA 3′) respectively. Fidelity of the produced constructs was assessed by full-genome sequencing. The resulting vector was linerized using BssHII and *In vitro* RNA synthesis, SK6 cells transfections and rescue of the mutant viruses were performed exactly as previously described (Risatti et al., 2005).

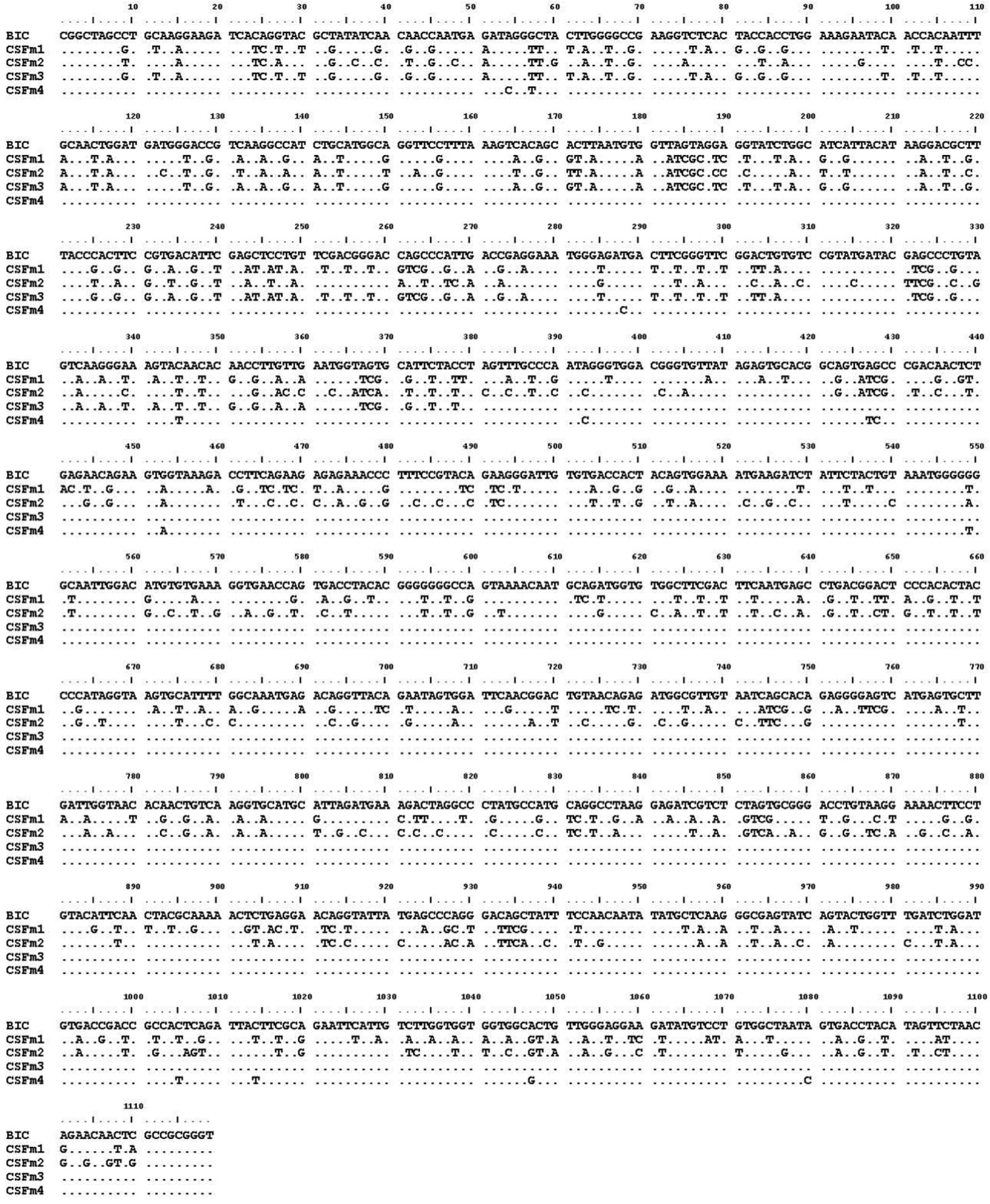

Fig. 1. Complete nucleotide sequence of CSFV E2 glycoprotein in CSFm1–CSFm4 constructs. All substituted nucleotides in the development of all mutated constructs are presented compared with the native sequence of E2 BICv.

2.4. Bioinformatic evaluation of mutant viruses

Different statistical methods were used to analyze *in silico* de-optimization of codon usage in the mutant viruses.

Relative synonymous codon usage (RSCU) and Codon adaptation index (CAI) calculations: RSCU is an index to assign a numerical value to individual codons that indicates how biased each of these codons is with reference to a random expectation codon usage, so that RSCU values larger than 1.0 indicate that one synonymous codon is favored over the rest; RSCU values less than 1.0 indicate an unfavored codon; and RSCU values of 1.0 indicate no preference (Sharp et al., 1986). CAI estimates the deviation of RSCU of a given protein coding sequence with respect to a reference dataset or relative adaptiveness. Values range from 0 to 1, with higher values indicating a higher proportion of the most abundant codons (Sharp and Li, 1987). Both analyses were conducted using the web server CAIcal (http://genomes.urv.es/CAIcal/) (Puigbo et al., 2007).

Codon pair bias (CPB): CPB for an entire reading frame has been defined as the arithmetic mean of the individual codon pair scores (CPS). The CPS of each codon pair is the natural log of the ratio of the observed over the expected number of occurrences over all coding regions in the genome. A positive CPB score indicates a prevalent use of overrepresented codon pairs, while a negative CPB score represents a predominant use of underrepresented codon pairs (Coleman et al., 2008). Following a methodology previously described (Tulloch et al., 2014), a reference codon pair usage table was developed based on the analysis of 25,883 *Sus scrofa* mRNA sequences (obtained from ftp://ftp.ensembl.org/pub/release-82/fasta/sus_scrofa/cds/). Based on the use of that table, we obtained CPB scores of all mRNA sequences used in this study from *Sus scrofa* as well as CPB scores from the E2 coding region of BICv and CSFm1-4v. All calculations were conducted using the program Composition Scan in the SSE software version 1.2 (Simmonds, 2012).

Additionally nucleotide and dinucleotide frequencies were obtained using the program Composition Scan in the SSE package (Simmonds, 2012).

2.5. DNA sequencing and analysis

Full-length clones of BICv and mutant constructs and derived *in vitro*-rescued viruses were completely sequenced. CSFV-specific primers were used to partially PCR amplify cDNA obtained from virus RNA by reverse transcription. PCR products were sequenced using the dideoxynucleotide chain-termination method (Sanger et al., 1977). Sequencing reactions were prepared with the Dye Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA). Reaction products were sequenced on a PRISM 3730xl automated DNA sequencer (Applied Biosystems). Sequence data was assembled with the Phrap software program (http://www.phrap.org), with confirmatory assemblies performed using CAP3 (Huang and Madan, 1999). The final DNA consensus sequence represented an average five-fold redundancy at each base position. Sequence comparisons were conducted using BioEdit software (http://www.mbio.ncsu.edu/BioEdit/bioedit.html).

Full-length virus genome sequencing for all recovered viruses was obtained by Next Generation Sequencing. Full-length sequencing of the virus genome was basically performed as described elsewhere (Krug et al., 2015). Viral RNA was obtained from the supernatant of infected swine kidney tissue culture cells (SK6 cells) using the TRIzol method (Life Technologies, Grand Island, NY, USA). RNA concentration was determined using the Qubit® RNA high-sensitivity (HS) assay kit (Life Technologies) and read on a Qubit 2 fluorometer (Life Technologies). One microgram of viral RNA was enzymatically fragmented to obtain blunt-end fragments in a length range of 100–300 bp using the enzyme Rnase III (Life Technologies) and incubated at 37 °C in a Peltier DNA Engine Tetrad 2 thermal cycler. Ion adapters were hybridized to fragmented RNA, hybridized fragmented RNA was reverse transcribed to cDNA and cDNA was ligated to ion library barcodes and amplified using the Ion total RNA-seq v2 kit (Life Technologies). 1ul of amplified cDNA was loaded onto Agilent DNA 1000 chip (Agilent, Santa Clara, CA, USA) and quantified using 2100 bioanalyzer (Agilent). Using the 2100 expert software, a smear analysis was performed to quantify the molar concentration (nM) of the cDNA libraries in order to determine the library dilution required for template preparation. After dilution of cDNA library to a concentration of 20 pM was performed, the cDNA library was clonally amplified onto Ion Sphere particles (ISPs), generating template-positive ISPs using the Ion PGM template OneTouch 2 200 kit (Life Technologies) with the Ion OneTouch 2 instrument (Life Technologies). Before proceeding to enrichment, quality assessment of nonenriched template-positive ISPs was performed using the Ion Sphere quality control (QC) assay kit (Life Technologies) and a Qubit 2 fluorometer. The template-positive ISPs were then enriched using the Ion PGM template OneTouch 2 200 kit (Life Technologies) and Ion OneTouch enrichment system (ES) instrument (Life Technologies) to eliminate template-negative ISPs and to denature cDNA on template-positive ISPs. Using the Ion PGM 200 sequencing v2 kit (Life Technologies), enriched template ISPs were prepared for sequencing, loaded onto either an Ion 314 or Ion 316 chip, v2 (Life Technologies), and run on the Ion PGM sequencer (Life Technologies). The sequences obtained were then trimmed using Galaxy (https://usegalaxy.org/) NGS QC and manipulation tools. Sequences were aligned and analyzed using Sequencher 5.2.2 (Genecodes) and CLC Genomics Workbench (CLCBio) software.

2.6. Animal studies

Animals were randomly allocated into the corresponding groups following each of the protocols. All animals were acclimated to their environment for seven days prior to beginning each experiment. Animal experiments were performed in accordance with protocols approved by Plum Island Animal Disease Center Institutional Animal Care and Use Committee.

Virulence of CSFm mutant viruses relative to BICv was initially assessed in 10–12 week old, forty-pound commercial breed pigs inoculated intranasally (IN) with 10^5 $TCID_{50}$ of each virus. Animals were randomly allocated into 4 groups ($n=5$) and were inoculated with a mutant virus or BICv. Clinical signs (anorexia, depression, purple skin discoloration, staggering gait, diarrhea and cough) and changes in body temperature were recorded daily throughout the experiment and scored as previously described (Risatti et al., 2007). Blood was collected at times post-infection from the anterior vena cava in EDTA-containing tubes (Vacutainer) and quantification of viremia by virus titration was as previously described.

For infection-challenge studies, 15 pigs were randomly allocated into three groups ($n=5$). Pigs in groups 1 and 2 were IN inoculated with CSFm2v, and pigs in group 3 were mock-infected. At 3 days post-inoculation (DPI) (group 1) or 28 DPI (group 2), animals were challenged with BICv along with animals in group 3. Clinical signs and body temperature were recorded daily throughout the experiment as described above. Blood samples were collected at times post-challenge for quantification of viremia as described earlier.

2.7. Quantitative real-time PCR (qPCR)

qPCR was performed as previously described (Borca et al., 2008). Briefly, swine macrophage primary cell cultures were infected at MOI=1, and total cellular RNA was extracted at 36 h post-infection. Contaminant genomic DNA was removed by DNase

Table 1
RSCU values of native CSFvE2 as well as the CSFm1-4v related to Sus scrofa codon usage.

Amino acid	Codon	*Sus scrofa*	BICv	CSFm1v	CSFm2v	CSFm3v	CSFm4v
Phe	UUU	0.79	0.38	**2.00**	**1.25**	1.00	0.38
Phe	UUC	**1.21**	**1.63**	0.00	0.75	1.00	**1.63**
Leu	UUA	0.37	0.53	**6.00**	**1.76**	**2.82**	0.53
Leu	UUG	0.77	1.41	0.00	1.06	0.71	1.41
Leu	CUU	0.75	0.35	0.00	0.35	0.00	0.35
Leu	CUC	1.55	0.88	0.00	0.88	0.53	0.88
Leu	CUA	0.38	1.24	0.00	1.59	1.06	1.24
Leu	CUG	**3.07**	**1.59**	0.00	0.35	0.88	**1.59**
Ile	AUU	0.91	0.64	0.00	**1.29**	0.64	0.64
Ile	AUC	**1.67**	0.64	0.00	0.43	0.43	**1.29**
Ile	AUA	0.42	**1.71**	**3.00**	**1.29**	**1.93**	1.07
Val	GUU	0.57	0.63	0.00	1.50	0.50	0.63
Val	GUC	1.06	0.88	0.00	0.75	0.50	0.88
Val	GUA	0.34	0.63	**4.00**	**1.63**	1.38	0.75
Val	GUG	**2.03**	**1.88**	0.00	0.13	**1.63**	**1.75**
Ser	UCU	0.99	0.35	0.00	0.35	0.35	0.38
Ser	UCC	1.50	1.06	0.00	0.71	0.35	1.13
Ser	UCA	0.72	1.41	0.00	1.76	0.71	**1.50**
Ser	UCG	0.39	0.00	**6.00**	**2.12**	**2.82**	0.00
Ser	AGU	0.77	1.41	0.00	1.06	0.71	**1.50**
Ser	AGC	**1.62**	**1.76**	0.00	0.00	1.06	**1.50**
Pro	CCU	1.05	1.11	0.00	1.11	0.89	1.11
Pro	CCC	**2.11**	0.89	0.00	1.11	0.67	0.89
Pro	CCA	1.37	**1.33**	0.00	0.44	1.11	**1.33**
Pro	CCG	0.80	0.67	**4.00**	**1.33**	**1.33**	0.67
Thr	ACU	0.83	0.76	0.00	1.08	0.54	0.76
Thr	ACC	**1.63**	1.30	0.00	0.65	0.54	1.30
Thr	ACA	0.89	**1.41**	0.00	0.76	0.97	**1.41**
Thr	ACG	0.56	0.54	**4.00**	**1.51**	**1.95**	0.54
Ala	GCU	0.95	0.44	0.00	**1.33**	0.00	0.44
Ala	GCC	**2.17**	0.89	0.22	0.44	0.22	0.89
Ala	GCA	0.89	**2.22**	0.00	0.89	1.33	**2.00**
Ala	GCG	0.61	0.44	**3.78**	**1.33**	**2.44**	0.67
Tyr	UAU	0.73	0.76	**2.00**	0.76	**1.14**	0.95
Tyr	UAC	0.93	**1.24**	0.00	**1.24**	0.86	**1.05**
His	CAU	0.70	0.86	**2.00**	**1.71**	**1.43**	**1.14**
His	CAC	**1.57**	**1.14**	0.00	0.29	0.57	0.86
Gln	CAA	0.44	**1.43**	**2.00**	**1.43**	**1.43**	**1.43**
Gln	CAG	**1.92**	0.57	0.00	0.57	0.57	0.57
Asn	AAU	0.79	**1.14**	**2.00**	0.71	**1.43**	**1.14**
Asn	AAC	0.80	0.86	0.00	**1.29**	0.57	0.86
Lys	AAA	0.76	0.89	**2.00**	**1.11**	**1.44**	0.89
Lys	AAG	0.83	**1.11**	0.00	0.89	0.56	**1.11**
Asp	GAU	0.80	**1.24**	**2.00**	**1.43**	**1.52**	**1.14**
Asp	GAC	0.95	0.76	0.00	0.57	0.48	0.86
Glu	GAA	0.73	0.96	**2.00**	**1.04**	**1.22**	0.96
Glu	GAG	**1.22**	**1.04**	0.00	0.96	0.78	**1.04**
Cys	UGU	0.78	0.93	**2.00**	0.93	**1.20**	0.88
Cys	UGC	**1.40**	**1.07**	0.00	**1.07**	0.80	**1.13**
Arg	CGU	0.44	0.00	**6.00**	0.95	0.95	0.00
Arg	CGC	**1.82**	0.32	0.00	**1.58**	0.32	0.32
Arg	CGA	0.84	0.00	0.00	1.26	0.00	0.00
Arg	CGG	1.79	0.00	0.00	0.63	0.00	0.00
Arg	AGA	1.55	**2.84**	0.00	0.00	**2.84**	**2.84**
Arg	AGG	1.71	**2.84**	0.00	**1.58**	1.89	**2.84**
Gly	GGU	0.57	1.16	**4.00**	1.16	**2.19**	**1.42**
Gly	GGC	**1.60**	0.65	0.00	0.52	0.65	0.65
Gly	GGA	1.00	0.90	0.00	**1.29**	0.52	0.90
Gly	GGG	1.15	**1.29**	0.00	1.03	0.65	1.03

RSCU values are presented for different codons encoding for the 18 amino acid residues used in the different versions of E2. In Bold are represented the highest RSCU value for each amino acid.

treatment and genomic DNA contamination of RNA stocks assessed by qPCR. cDNA was synthesized in 100 µl reactions containing 2500 ng of total RNA. For gene expression qPCR was performed using *Power* SYBR® Green with primer pairs described in (Borca et al., 2008). Melting curve analysis was performed to verify specificity of the amplification. Normalization of gene expression was performed with β-actin mRNA. Relative quantities (RQ) of mRNAs were estimated using the $2^{-\Delta\Delta Ct}$ method (described in ABI PRISM® 7700 Sequence Detection System User Bulletin #2 (PN 4303859)). The normalized mRNA expression level of a cellular gene in the infected cells was considered significant when it departed from its level in uninfected cells threefold, in either direction. Swine genes targeted included: AMCF-2, CD46, DAP12, ICAM, ICE, IL-1α, IL-1β, IL-1R, IL-2, IL-2α, IL-4, IL-5, IL-6, IL-6R, IL-7, IL-8, IL-8R, IFNα, IFNβ, IFNγ, IL-10, IL-12p35, IL-12p40, IL-12R, IL-13, IL-15, IL-16, IRF-1, IRF-3, IRF-6, IRF-7, IRF-9, MCP-1, MCP-2, OAS, NKG2D, NCP-1, NCP-2, NOS, PKR, RANTES, TGF-β1, TGF-β2, TGF-B3R, TLR-1, TLR-2, TLR-3, TLR-5 and TLR-10.

3. Results

3.1. Bioinformatic evaluation

To evaluate the effect produced by codon usage de-optimization along the glycoprotein E2 coding region of CSFV we employed two different algorithms (see Section 2) to produce mutant viruses in which codons that encode the same amino acid were substituted by the less commonly used synonymous codon in the reference host dataset (*Sus scrofa*). After the de-optimization process the total number of synonymous mutations varied among the mutants, with mutant CSFm1v having the highest number of synonymous mutations with 312, followed by CSFm2v with 298, and CSFm3v with 109. Finally, CSFm4 has just 12 synonymous mutations manually introduced in highly E2 conserved codons corresponding to nucleotide positions: 2482$^{(I)ATA\text{-}ATC}$, 2485$^{(G)GGG\text{-}GGT}$, 2716$^{(D)GAT\text{-}GAC}$, 2773$^{(Y)TAC\text{-}TAT}$, 2821$^{(I)ATA\text{-}ATC}$, 2855$^{(S)AGC\text{-}TGC}$, 2881$^{(V)GTG\text{-}GTA}$, 2977$^{(G)GG\text{-}GGT}$, 3433$^{(H)CAC\text{-}CAT}$, 3442$^{(Y)TAC\text{-}TAT}$, 3475$^{(A)GCA\text{-}GCG}$, and 3508$^{(I)ATA\text{-}ATC}$. The exact positions of all synonymous mutations introduced in the CSFm mutant constructs are shown in Fig. 1.

3.2. Effects on relative synonymous codon usage in CSFm1v to CSFm4v mutants

To evaluate the impact on synonymous codon usage after the de-optimization process we evaluated BICv and CSFm1v-CSFm4 E2 by

using RSCU and CAI indexes (see Section 2). RSCU index indicates how biased each of these codons is with reference to a random expectation codon usage. In general, the average RSCU of the coding regions analyzed from *Sus scrofa* was characterized for the high preference of usage of C/G ending codons (20 out of 23 with RSCU values > 1), with the CUG codon being one of the six codons to encode the amino acid leucine which has the highest RSCU (3.07), while there is a low codon usage pattern especially of codons containing the dinucleotides CpG (UCG, CCG, ACG, GCG, CGU, CGA) and UpA (UUA, AUA, GUA, UAU, UAC) in their composition. On the other hand, codon adaptation index (CAI), which measures the relative adaptiveness of a given protein coding gene sequence with respect to a reference set of genes, was calculated between the reference data set of genes from *Sus scrofa* and BICv (E2 coding region).

Results indicated that there is a 66% similarity in the RSCU pattern between *Sus scrofa* and native CSFV E2 of BICv (Table 1). BICv showed the same low preference for the usage of codons that contain CpG and UpA in their composition. However, after substituting the mutant viruses native codons by the less synonymous codons used by *Sus scrofa* CAI scores changed to different levels according to the de-optimization strategy employed to obtain each mutant, with CSFm1v being the one with the lowest CAI value (0.38), followed by CSFm3v (0.55), CSFm2v (0.56) and finally CSFm4v that remained at the same value as BICv (0.66). As expected, due to the used algorithm, the highest decrease of relative adaptiveness regarding *Sus scrofa* was found in the CSFm1 E2 sequence.

It has been previously shown that by altering codon pair bias (CPB) throughout the viral genome can cause differences in viral virulence in the natural host (Coleman et al., 2008). Although our de-optimization strategies were not based on altering CPB in these mutant viruses, we wanted to evaluate any potential effect that could occur on CPB by using the two algorithms employed in this report. Our analysis showed that there was a normal distribution of the *Sus scrofa* genes, with an average CPB score of 0.067 ± 0.035 (Fig. 2). The CPB score of BICv was calculated to be -0.023, which remained very similar to that for CSFm1v (-0.025) and CSFm4v (-0.028). Nonetheless, while the one amino acid-one codon algorithm does not have significant effects on the CPB value of CSFm1v, in the case of CSFm3v (developed using the same algorithm but just in 1/3 of E2 coding sequence) produced a positive CPB value of 0.007, which made this mutant closer to the average CPB scores from *Sus scrofa* genes. Conversely, the algorithm used to design CSFm2v had the highest negative impact in CPB (-0.086) among all CSFm mutants, by increasing the usage of unrepresented pair codons in *Sus scrofa* (Fig. 2).

Increasing CpG and UpA dinucleotide frequencies in the genome of some viruses has been linked with viral attenuation

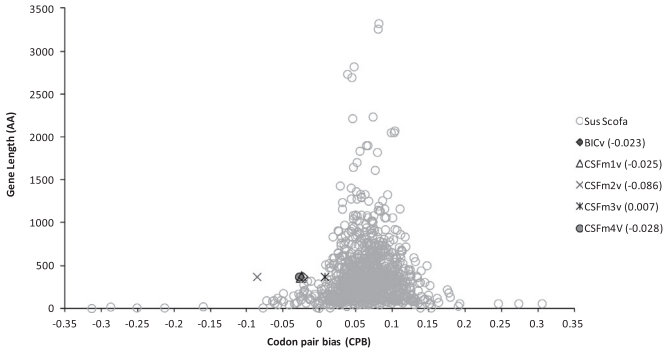

Fig. 2. Codon pair bias (CPB) scores of swine genes and CSFV E2 from BICv and CSFm1-4 mutants. Comparison was performed based on the CPB score of 25,883 *Sus scrofa* genes. Each mark represents the calculated CPB score of a gene plotted against its amino acid length. CPB scores of the CSFm1-4 mutants are indicated. Calculation of CPB values were performed as described in Section 2.

(Tulloch et al., 2014). Since our de-optimization strategy increases the usage in the mutant viruses of less commonly used synonymous codons in *Sus scrofa* associated with codons containing CpG and UpA dinucleotides in their composition, we expected an increase of the frequencies of these two dinucleotides in the genomic composition of the mutant viruses. Comparing odds ratios of observed/expected frequencies of these two dinucleotides between BICv E2 and those in the CSFm mutant viruses, we found a disruption at different levels in the frequency patterns associated with these two dinucleotides along the E2 coding region of the mutant viruses. The highest changes in the observed/expected odds ratios in these two dinucleotides were found in CSFm1v, followed by CSFm2v and CSFm3v, while CSFm4v remained very similar to BICv (Fig. 3).

It has been demonstrated that low levels of G/C content are associated with decreased gene expression in mammalian cells (Kudla et al., 2006). Comparison of G/C nucleotide compositions between BICv and mutant viruses indicated that the most important changes occurred in mutant CSFm1v E2, where total G/C composition decreased 9% in contrast with BICv E2, while no significant changes were detected in the other three CSFm E2 mutants (CSFm2v and CSFm3v have a decrease of 2 and 3%, respectively, while no change was detected for CSFm4v).

3.3. Production of CSFm viruses

To assess the importance of the introduced mutations in the process of virus replication and virulence a CSFm construct series of recombinant CSFVs containing mutations in E2 was designed using the cDNA infectious clone (IC) of the Brescia strain (BICv) as a template. A total of 4 cDNA IC constructs containing the distinctive areas of E2 mutated as described above were developed, pCSFm1-pCSFm4. Infectious RNA was *in vitro* transcribed from each mutated full-length cDNA and used to transfect SK6 cells. Infectious virus was rescued by day 4 post-transfection from cells transfected with viral RNA derived from constructs pCSFm2, pCSFm3, and pCSFm4. In contrast, after three independent transfection procedures, pCSFm1 viral RNA did not produce infectious virus. Whole genome sequencing was performed by next generation sequencing (described in the Section 2) for each of the mutant viruses, and no changes were detected at the nucleotide level in the viral stock of viruses pCSFm2v-pCSFVm4v when compared to the initial designed construct.

3.4. Replication of the CSFm mutant viruses in vitro

The importance of the introduced mutations in the process of virus replication was primarily assessed in SK6 cells. *In vitro* growth characteristics of the CSFm2v, CSFm3v, and CSFm4v viruses relative to parental BICv were evaluated in a multistep growth curve. SK6 cell cultures were infected at an MOI of 0.01 $TCID_{50}$ per cell. Viruses were adsorbed for 1 h (time zero), and samples were collected at 72 h post-infection and titrated in SK6 cell cultures using immunoperoxidase staining, calculation of titers was previously described by Risatti et al. (2005). Results indicate that all three mutant CSFm viruses exhibited an almost indistinguishable replication kinetics compared to that of the parental BICv (Fig. 4).

3.5. Virulence of CSFm mutant viruses in vivo

To examine whether the genomic changes introduced into E2 of CSFm2v, CSFm3v, and CSFm4v viruses may affect virulence, four different groups of pigs (n=5) were intranasally (IN) inoculated with 10^5 $TCID_{50}$ of each of the CSFm mutant viruses (CSFm2v, CSFm3v, and CSFm4v) and monitored for clinical disease, evaluated relative to parental BICv (Table 2, Fig. 5A and B). All animals

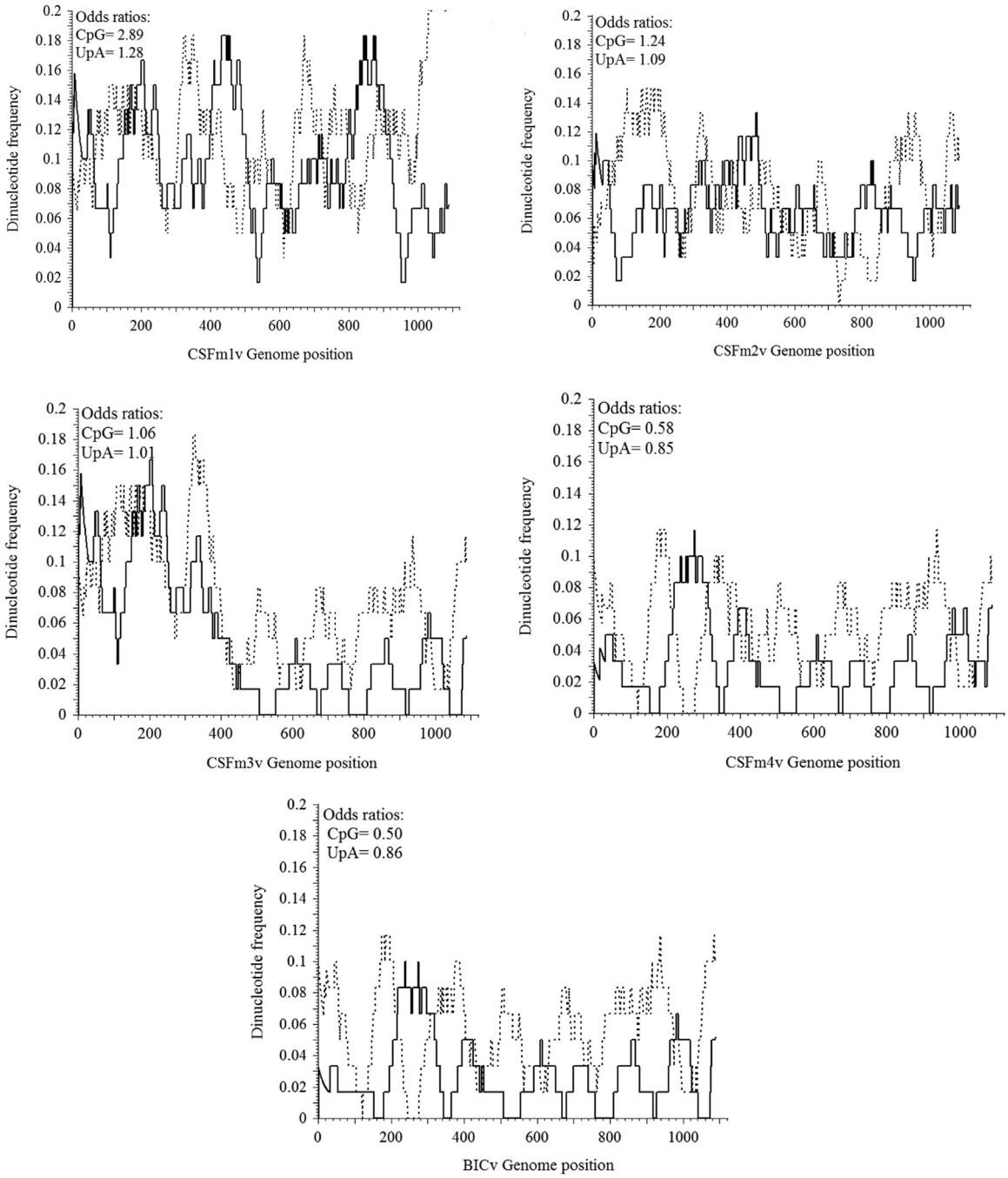

Fig. 3. Comparison of CpG and UpA dinucleotide frequencies along the E2 genomic region between BICv and CSFm1-4 mutants. Solid and dotted lines represents CpG and UpA dinucleotides respectively. Frequency values and odds ratio calculations were obtained using the program Composition Scan in the SSE package version 1.2 (Simmonds, 2012) as described in Section 2.

infected with BICv presented a rise in body temperature and clinical signs of CSF starting around 3 DPI. Animals quickly developed classical symptoms of severe disease, being euthanized around 5 to 6 DPI. CSFm3v and CSFm4v viruses presented a virulence phenotype almost indistinguishable from that of the parental BICv (Table 2, Figs. 5A and B and 6). All animals infected with these two viruses presented clinical signs of CSF starting at 3–4 DPI, with clinical presentation and severity similar to those observed in animals inoculated with BICv. Conversely, animals inoculated with CSFm2v survived the infection and remained clinically normal throughout the observation period (21 days). No rise in body temperature was observed, and only a transient and very mild diarrhea during the first week after infection was observed in these animals (Table 2, Figs. 5A and B and 6).

As expected, in general viremias accompanied the evolution of the clinical disease. Therefore, viremias in animals inoculated with BICv are clearly detected by 3 DPI reaching the highest titers by the time of death at 5–7 DPI (Fig. 6). Animals infected with mutant viruses CSFm3v or CSFm4v exhibited viremia kinetics almost indistinguishable from that induced by parental BICv, presenting high titers that remained until death of the animal. In contrast, mutant virus CSFm2v induced a mild and transient viremia lasting during the first week post-inoculation, with virus titers being significantly lower (1000 to 10,000 folds depending of the sample time considered) than those found in animals inoculated with parental BIC virus (Fig. 6).

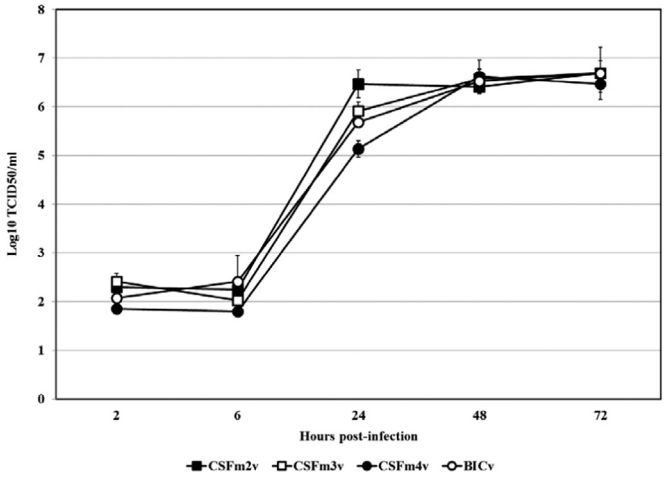

Fig. 4. *In vitro* growth characteristics of CSFm2-4v mutant viruses and parental BICv. SK6 cell cultures were infected (MOI=0.01) with CSFm2-4v with CSFm2v or BICv. Virus yield obtained at the indicated times post-infection were titrated in SK6 cell cultures. Data represent means and standard deviations from three independent experiments. Sensitivity of virus detection: $\geq$ 1.8 $TCID_{50}$/ml.

Table 2
Swine survival and fever response following infection with CSFm mutant viruses or virulent parental BICv.

Virus	No. of survivors/total No.	Mean time to death (days ± SD)	Fever	
			No. of days to onset (days ± SD)	Duration (days ± SD)
BICv	0/5	5. 6 (1.34)	3.4 (0.89)	2.2 (1.79)
CSFm2v	5/5	–	–	–
CSFm3v	0/5	6.4 (0.89)	3.6 (0.89)	2.82 (0.84)
CSFm4v	0/5	5.8 (1.1)	3.4 (0.55)	2.4 (1.14)

3.6. Comparative growth of CSFm2 mutant virus in swine macrophage cell cultures

Cells derived from the macrophages lineage are the main target for CSFV replication during the infection in swine. Since the CSFm2 mutant virus clearly replicates with lower efficiency than BICv during swine infection it was important to assess the comparative ability of CSFm2v to replicate in swine macrophages. *In vitro* growth characteristics of the CSFm2v mutant virus relative to parental BICv were evaluated in a multistep growth curve. Primary swine macrophage cell cultures were infected at an MOI of 0.01 $TCID_{50}$ per cell. Viruses were adsorbed for 1 h (time zero), and samples were collected during 72 h post-infection. Samples were titrated in SK6 cells and the presence of virus detected by immunoperoxidase staining. The growth of CSFm2 virus was significantly lower than BICv, showing titers approximately 100-fold lower than the parental virus, depending on the sampling time considered (Fig. 7). Similar results were obtained in a single-step growth curve experiment comparatively performed in primary swine macrophage cell cultures using CSFm2v and BICv at a MOI of 1 and harvested at 36 h post-inoculation. CSFm2v reached titers approximately 100 times lower (4.25_{log10} $TCID_{50}$ ± SD=0.16) than the parental BICv (6.42_{log10} $TCID_{50}$ SDL: 0.16) (data not shown). Therefore, attenuated CSFm2 virus has an evident disadvantage replicating in swine macrophages compared with the parental virulent BICv.

3.7. Transcriptional activation profile of immunologically relevant genes in swine macrophages infected with CSFm2 virus

To further understand possible mechanisms responsible for CSFm2v attenuation, the pattern of activation of immunologically relevant genes in swine macrophages infected with mutant and parental viruses was analyzed using quantitative real-time PCR followed by melting curve analysis. To assess changes in cellular

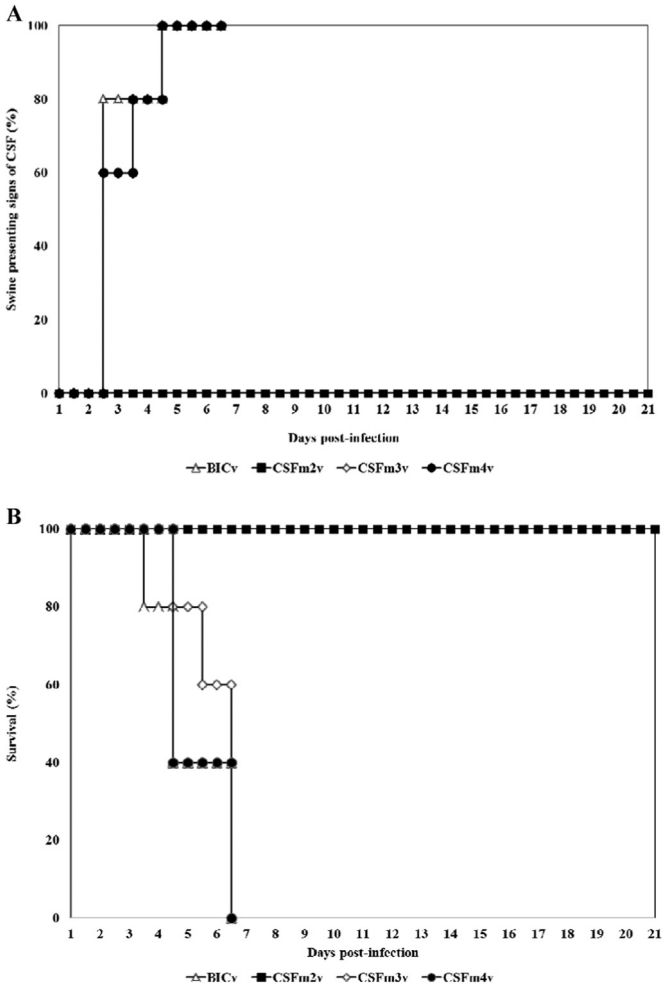

Fig. 5. Morbidity (A) and mortality (B) in swine infected with mutant viruses CSFm2-4v or parental BICv. Groups of animals (n=5) were infected intranasally (IN) with 10^5 $TCID_{50}$. Presence of clinical signs was recorded for an observational period of 21 DPI.

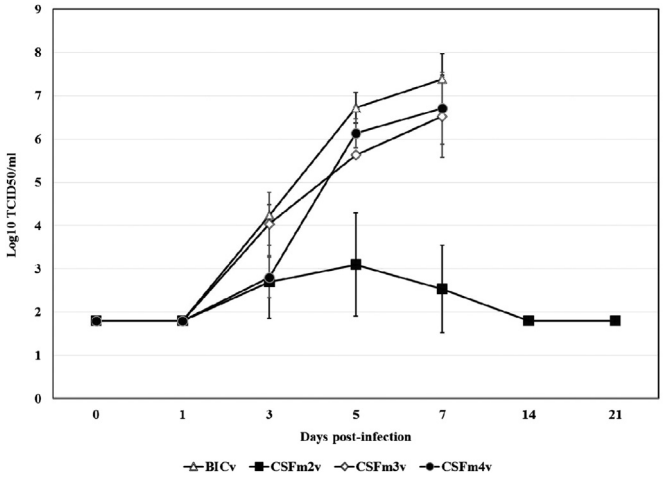

Fig. 6. Viremia titers in swine infected with mutant viruses CSFm2-4v or parental BICv. Groups of animals (n=5) were infected IN with 10^5 TCID $TCID_{50}$. Viremia data represent the mean log_{10} TCID $TCID_{50}$/ml and standard deviations of titers from all animals in the group. Sensitivity of virus detection: $\geq log_{10}$ 1.8 TCID $TCID_{50}$/ml.

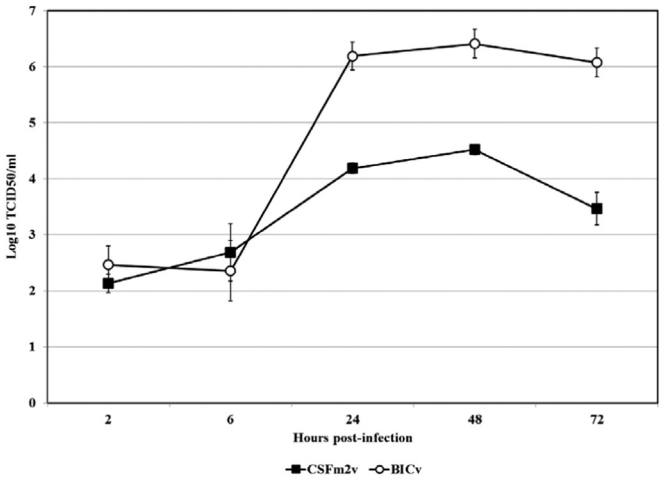

Fig. 7. *In vitro* growth characteristics of CSFm2 mutant virus and parental BICv in swine macrophages. Primary swine macrophages cell cultures were infected (MOI=0.01) with CSFm2v or BICv. Virus yield obtained at the indicated times post-infection were titrated in SK6 cell cultures. Data represent means and standard deviations from three independent experiments. Sensitivity of virus detection: $\geq$ 1.8 $TCID_{50}$/ml.

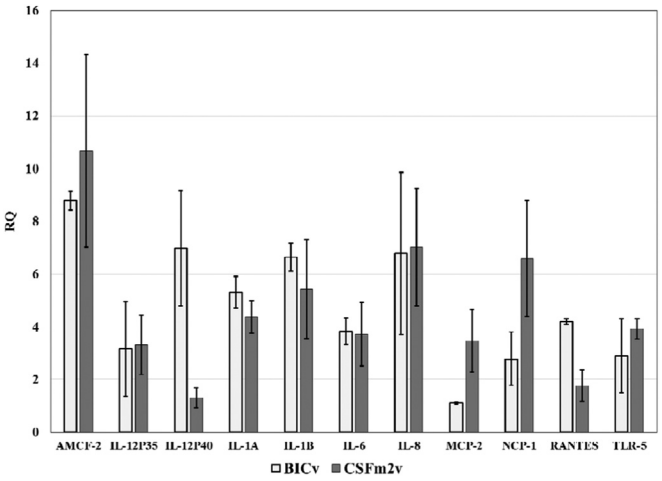

Fig. 8. Gene expression changes in peripheral blood-derived macrophages after infection with either CSFm2 or BIC viruses. Gene expression quantification was assessed by quantitative reverse transcription real-time polymerase chain reaction (qrt-PCR). Values are represented as relative quantities (RQ) of mRNA accumulation (estimated by $2^{-\Delta\Delta Ct}$) with their corresponding SD. RQ values over 3 are considered as positive.

gene expression upon infection, primary porcine macrophage cell cultures were infected at MOI=1 $TCID_{50}$ per cell with either CSFm2 virus or BICv. Total cellular RNA was extracted from infected and mock-infected cells at 36 hpi. Steady state levels of mRNA accumulation were determined for 49 swine immunomodulatory genes as described elsewhere (Borca et al., 2008). This approach identified only eleven genes (AMCF-2, IL-12p35, IL-12p40, IL-1α, IL-1β, IL-6, IL-8, MCP-2, NCP-1, RANTES and TLR-5) actually showing a significant expression increase in macrophages infected with either CSFm2v or BICv when compared with the pattern of gene expression of uninfected macrophages (Fig. 8). Among these genes only four of them presented a differential expression between CSFm2v and BICv viruses. MCP-2, and NCP-1 has increased expression in macrophages infected with CSFm2 virus while IL-12p40 and RANTES were more actively transcribed in BICv-infected macrophages (Fig. 8).

3.8. CSFm2 mutant virus protects pigs against lethal CSFV challenge

The ability of attenuated CSFm2v virus to induce protection against virulent BICv was assessed in early and late vaccination-exposure experiments. Groups of pigs (n=5) were IN inoculated with 10^5 $TCID_{50}$ of CSFm2v and IN challenged at 3 or 21 DPI with 10^5 $TCID_{50}$ virulent BICv. Mock-vaccinated control pigs receiving BICv (n=5) developed anorexia, depression and fever by 5 days post-challenge (DPC) (Table 3), and died or were euthanized *in extremis* by 7 DPC. CSFm2 virus induced complete clinical protection in animals challenged at 21 days post-immunization. All pigs survived infection with the virulent parental BICv and remained without fever or demonstrating any other CSF-associated clinical signs during the 21 day observational period (Table 3). Animals infected with CSFm2v and challenged with BICv 3 days later presented a heterogeneous behavior. Three of the animals remained clinically normal with the exception of a very weak and transient rise in body temperature in one of them during the observational period. The other two animals became sick at the same time that the animals in the mock-vaccinated group did although they presented a clinical disease more extended, being euthanized around 10 DPC.

Viremia in vaccinated and challenged animals was examined at different times post-challenge. As expected, in mock-vaccinated control animals, viremia was observed within 4 dpc, with virus titers remaining typically high (6.84_{log10} $TCID_{50}$ ± SD=0.42) at the last time point tested before animals died or were euthanized (Fig. 9). Conversely, animals inoculated with CSFm2v and challenged with BICv at 21 DPI did not produce detectable viremia. BICv was not detected in blood samples during the 21-day observation period (Figs. 7 and 8). Alternatively, in the group of CSFm2v-infected animals challenged with BICv at 3 DPI, viremia's kinetics follows the presence of clinical signs. Animals showing severe disease and death presented viremias similar to those in animals mock-vaccinated (7.12_{log10} $TCID_{50}$ ± SD=0.71 at day 7 post challenge) while in those remaining clinically normal, transient viremias were detected at low levels (3.29 $TCID_{50}$ ± SD=0.29) by 4 dpc (Fig. 8). These results indicate that protection induced by CSFm2v was complete, preventing both the presentation of CSF-related clinical signs and the replication of the challenge virus when challenge was conducted at 21 days after initial infection.

Table 3
Swine survival and fever response in CSFm-vaccinated animals after challenge with parental virulent BICv.

Challenge group (at dpi)	No. of survivors/ total	Mean time to death (± SD)	Fever		
			No. of days to onset (± SD)	Duration no. of days (± SD)	Maximum daily temp, °F (± SD)
Mock	0/5	6.8 (0.45)	5 (0.0)	1.8 (0.45)	105.8 (0.22)
CSFm2v 3 dpi	3/5	10 (1.41)	5 (0.0)	5 (1.41)	105.6 (0.81)
CSFm2v 21 dpi	5/5	–	–	0	102.2 (0.27)

4. Discussion

The introduction of synonymous mutations in virus genomes has been used to experimentally develop attenuated virus strains by de-optimizing the codon usage bias or the codon pair bias. This

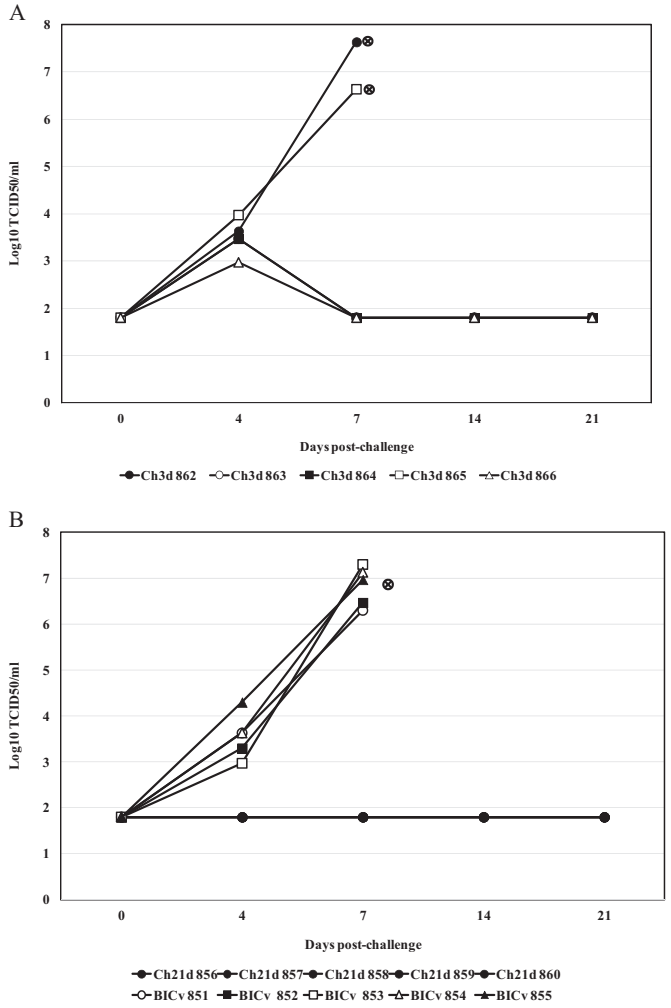

Fig. 9. Viremia titers in swine infected with CSFm2v virus and challenged with BICv. Three groups of animals ($n=5$) were either infected IN with 10^5 $TCID_{50}$ CSFm2v virus (animal tags # 856–860 and 862–866) or mock infected (animal tags # 851–855) and challenge IN with 10^5 $TCID_{50}$ of BICv at either (A) 3 (Ch3d 862–866) or (B) 21 (Ch21d 856–860) days post-CSFm2v infection. Viremia data represent individual animal data for those mock vaccinated (851–855) or challenged at 3 DPI (Ch3d 862–866) and the mean log10 TCID50/ml and standard deviations of animals challenged at 21 DPI (Ch21d 856–860). Sensitivity of virus detection: $\geq \log_{10}$ 1.8 TCID50/ml.

methodology has been used to develop several attenuated RNA virus strains utilizing poliovirus (Nogales et al., 2014; Mueller et al., 2006; Burns et al., 2006), PRRSV (Gao et al., 2015), dengue virus (Shen et al., 2015), influenza Mueller et al. (2010), respiratory syncytial virus (Le Nouen et al., 2014; Meng et al., 2014) and vesicular stomatitis virus (Wang et al., 2015).

The mechanisms causing virus attenuation remain poorly understood and may depend on many factors. Initial reports indicated that virus attenuation can be due to translation impairment although more recently it has been suggested that changes in the RNA sequence may increase innate immune responses in the host. Different parameters raised from the analysis of the nucleotide sequence of the mutated genes have been used to predict virus attenuation.

Although RSCU and CPB represent two different concepts, our methodology based on de-optimizing the RSCU indirectly altered the CPB. CPB de-optimization has been correlated with alteration in the ability of the mutant viruses to replicate and produce disease in diverse viruses such as PRRS (Gao et al., 2015), influenza (Mueller et al., 2010), dengue (Shen et al., 2015), human respiratory syncytial virus (Le Nouen et al., 2014) and VSV (Wang et al., 2015). Importantly, the increase in the usage of the underrepresented codon pairs of the attenuated CSFm2 virus, which showed the lowest CBP score among all CSFm mutants, is consistent with those findings. Interestingly, it has been reported that clustering of CSFV isolates based on their CPBs clearly discriminates between highly virulent strains, vaccine strains and moderately virulent strains (Leifer et al., 2013).

It has been reported that genomes of RNA viruses have low CpG and UpA dinucleotide frequencies (Rima and McFerran, 1997; Rothberg and Wimmer, 1981) compared with mammalian genomes (Alff-Steinberger, 1987). Previous reports showed that poliovirus (Mueller et al., 2006; Burns et al., 2006) and echovirus mutants (Atkinson et al., 2014) that have increased CpG and UpA frequencies display decreased replication kinetics compared to wild-type virus. Conversely, virus mutants where CpGs and UpAs were removed displayed increased replication compared to parental wild-type virus (Atkinson et al., 2014). A comparative analysis of CpG and UpA dinucleotide frequencies and attenuation in different virus models indicates correlation between increased frequencies and attenuation (Tulloch et al., 2014). As has been shown using a PRRSV system (Gao et al., 2015), the CSFm mutant construct that displayed the highest CpG and UpA dinucleotide frequency, CSFm1, was not able to produce viable virus progeny. Importantly, of the CSFm mutant viruses that were able to replicate, the mutant presenting the highest frequency of CPG and UPA dinucleotides, CSFm2v, is the only one with an attenuated phenotype in swine.

It is not clear what the molecular mechanisms are underlying decreased virus replication and, more importantly, attenuation of viruses suffering from these synonymous de-optimizations. It has been proposed that de-optimization of codon pair bias produces an alteration in the translation efficiency, affecting the rate of protein production which may result in slower virus replication rates (Coleman et al., 2008; Mueller et al., 2010; Yang et al., 2013). Alternatively, it has been suggested that virus attenuation induced by increased CpG and UpA frequencies can be mediated by an enhanced host innate immune response (Burns et al., 2006; Zsak et al., 1996; Atkinson et al., 2014). Both hypotheses may be combined and virus attenuation may result because viruses replicating at much slower rate, not causing significant damage to a host cell, will subsequently allow the induction of an efficient innate immune response.

We do not know the host/virus mechanisms mediating attenuation during the infection of CSFm2v virus in swine. Attenuated CSFm2v yield in swine macrophage cultures was clearly reduced when compared to parental virulent BICv. Accordingly, CSFm2v virus replication during infection in swine is severely decreased when compared with BICv. Similar results were reported in de-optimized attenuated viruses tested in experimental (Nogales et al., 2014; Meng et al., 2014; Mueller et al., 2006; Wang et al., 2015) as well as natural hosts (Le Nouen et al., 2014; Gao et al., 2015). In regards to a possible stimulation of the innate immune response by the attenuated de-optimized virus as described by others (Burns et al., 2006; Tulloch et al., 2014; Atkinson et al., 2014; Blaszczyk et al., 2000), we found a few changes in the pattern of immune-relevant genes activated in swine macrophages infected by CSFm2v. Increased expression of chemokines MCP-2 and NCP-1 may account for increased recruitment of cells mediating the host innate immune response which in turn may result in decreased virus replication. For example the monocyte chemotactic protein 2 (MCP-2), has been described as a potent inhibitor of HIV-1 strains by interacting with the cell surface receptor CCR5 that is used by this virus as a co-receptor (Blaszczyk et al., 2000).

In summary, it is shown here that introduction of an specific

type of silent mutations in massive amounts through determined area of CSFV genome encoding for structural glycoprotein E2 produced a severe attenuation of the parental virulent phenotype during swine infection. Importantly, infection with the attenuated virus induced protection of the infected animals against challenge with the virulent parental virus. This is the first report showing the development of an attenuated CSFV strain by codon usage de-optimization, and one of the few examples of virus attenuated using this methodology that is assessed in its natural host.

Acknowledgments

We thank the Plum Island Animal Disease Center Animal Care Unit staff for excellent technical assistance. We particularly like to thank Melanie Prarat for editing the manuscript. We also thank Dr Peter Simmonds for allowing the use of SSE program used for CPB analysis and dinucleotide calculation (Simmonds, 2012). This work was partially supported by the State of Kansas National Bio and Agro-Defense Facility Fund (NBAF) and Institute for Infectious Animal Diseases (IIAD) Career Development Program.

References

Alff-Steinberger, C., 1987. Codon usage in *Homo sapiens*: evidence for a coding pattern on the non-coding strand and evolutionary implications of dinucleotide discrimination. J. Theor. Biol. 124, 89–95.

Atkinson, N.J., Witteveldt, J., Evans, D.J., Simmonds, P., 2014. The influence of CpG and UpA dinucleotide frequencies on RNA virus replication and characterization of the innate cellular pathways underlying virus attenuation and enhanced replication. Nucleic Acids Res. 42, 4527–4545.

Baker, S.F., Nogales, A., Martinez-Sobrido, L., 2015. Downregulating viral gene expression: codon usage bias manipulation for the generation of novel influenza A virus vaccines. Future Virol. 10, 715–730.

Becher, P., Avalos Ramirez, R., Orlich, M., Cedillo Rosales, S., Konig, M., Schweizer, M., Stalder, H., Schirrmeier, H., Thiel, H.J., 2003. Genetic and antigenic characterization of novel pestivirus genotypes: implications for classification. Virology 311, 96–104.

Behura, S.K., Severson, D.W., 2013. Codon usage bias: causative factors, quantification methods and genome-wide patterns: with emphasis on insect genomes. Biol. Rev. Camb. Philos. Soc. 88, 49–61.

Blaszczyk, J., Coillie, E.V., Proost, P., Damme, J.V., Opdenakker, G., Bujacz, G.D., Wang, J.M., 2000. Complete crystal structure of monocyte chemotactic protein-2, a CC chemokine that interacts with multiple receptors. Biochemistry 46, 14075–14081.

Borca, M.V., Gudmundsdottir, I., Fernandez-Sainz, I.J., Holinka, L.G., Risatti, G.R., 2008. Patterns of cellular gene expression in swine macrophages infected with highly virulent classical swine fever virus strain Brescia. Virus Res. 138, 89–96.

Burns, C.C., Shaw, J., Campagnoli, R., Jorba, J., Vincent, A., Quay, J., Kew, O., 2006. Modulation of poliovirus replicative fitness in HeLa cells by deoptimization of synonymous codon usage in the capsid region. J. Virol. 80, 3259–3272.

Cheng, B.Y., Ortiz-Riano, E., Nogales, A., de la Torre, J.C., Martinez-Sobrido, L., 2015. Development of live-attenuated arenavirus vaccines based on codon deoptimization. J. Virol. 89, 3523–3533.

Coleman, J.R., Papamichail, D., Skiena, S., Futcher, B., Wimmer, E., Mueller, S., 2008. Virus attenuation by genome-scale changes in codon pair bias. Science 320, 1784–1787.

Edwards, S., Moennig, V., Wensvoort, G., 1991. The development of an international reference panel of monoclonal antibodies for the differentiation of hog cholera virus from other pestiviruses. Vet. Microbiol. 29, 101–108.

Gao, L., Wang, L., Huang, C., Yang, L., Guo, X.K., Yu, Z., Liu, Y., Yang, P., Feng, W.H., 2015. HP-PRRSV is attenuated by de-optimization of codon pair bias in its RNA-dependent RNA polymerase nsp9 gene. Virology 485, 135–144.

Huang, X., Madan, A., 1999. CAP3: a DNA sequence assembly program. Genome Res. 9, 868–877.

Jayaraj, S., Reid, R., Santi, D.V., 2005. GeMS: an advanced software package for designing synthetic genes. Nucleic Acids Res. 33, 3011–3016.

Jenkins, G.M., Holmes, E.C., 2003. The extent of codon usage bias in human RNA viruses and its evolutionary origin. Virus Res. 92, 1–7.

Konig, M., Lengsfeld, T., Pauly, T., Stark, R., Thiel, H.J., 1995. Classical swine fever virus: independent induction of protective immunity by two structural glycoproteins. J. Virol. 69, 6479–6486.

Krug, P.W., Holinka, L.G., O'Donnell, V., Reese, B., Sanford, B., Fernandez-Sainz, I., Gladue, D.P., Arzt, J., Rodriguez, L., Risatti, G.R., Borca, M.V., 2015. The progressive adaptation of a georgian isolate of African swine fever virus to vero cells leads to a gradual attenuation of virulence in swine corresponding to major modifications of the viral genome. J. Virol. 89, 2324–2332.

Kudla, G., Lipinski, L., Caffin, F., Helwak, A., Zylicz, M., 2006. High guanine and cytosine content increases mRNA levels in mammalian cells. PLoS Biol. 4, e180.

Le Nouen, C., Brock, L.G., Luongo, C., McCarty, T., Yang, L., Mehedi, M., Wimmer, E., Mueller, S., Collins, P.L., Buchholz, U.J., DiNapoli, J.M., 2014. Attenuation of human respiratory syncytial virus by genome-scale codon-pair deoptimization. Proc. Natl. Acad. Sci. USA 111, 13169–13174.

Leifer, I., Ruggli, N., Blome, S., 2013. Approaches to define the viral genetic basis of classical swine fever virus virulence. Virology 438, 51–55.

Luan, S., Pan, W., Li, T., Yang, H., Zhang, B., Li, F., Chen, L., 2009. Rescued influenza A virus with codon deoptimized NS1 gene is attenuated both in vitro and in vivo. Sheng Wu Gong Cheng Xue Bao 25, 720–726.

Meng, J., Lee, S., Hotard, A.L., Moore, M.L., 2014. Refining the balance of attenuation and immunogenicity of respiratory syncytial virus by targeted codon deoptimization of virulence genes. MBio 5, e01704–e01714.

Meng, J., Lee, S., Hotard, A.L., Moore, M.L., 2014. Refining the balance of attenuation and immunogenicity of respiratory syncytial virus by targeted codon deoptimization of virulence genes. MBio 5, e01704–e01714.

Mueller, S., Papamichail, D., Coleman, J.R., Skiena, S., Wimmer, E., 2006. Reduction of the rate of poliovirus protein synthesis through large-scale codon deoptimization causes attenuation of viral virulence by lowering specific infectivity. J. Virol. 80, 9687–9696.

Mueller, S., Coleman, J.R., Papamichail, D., Ward, C.B., Nimnual, A., Futcher, B., Skiena, S., Wimmer, E., 2010. Live attenuated influenza virus vaccines by computer-aided rational design. Nat. Biotechnol. 28, 723–726.

Nakamura, Y., Gojobori, T., Ikemura, T., 2000. Codon usage tabulated from international DNA sequence databases: status for the year 2000. Nucleic Acids Res. 28, 292.

Ni, Y.Y., Zhao, Z., Opriessnig, T., Subramaniam, S., Zhou, L., Cao, D., Cao, Q., Yang, H., Meng, X.J., 2014. Computer-aided codon-pairs deoptimization of the major envelope GP5 gene attenuates porcine reproductive and respiratory syndrome virus. Virology 450-451, 132–139.

Nogales, A., Baker, S.F., Ortiz-Riano, E., Dewhurst, S., Topham, D.J., Martinez-Sobrido, L., 2014. Influenza A virus attenuation by codon deoptimization of the NS gene for vaccine development. J. Virol. 88, 10525–10540.

Pesole, G., Attimonelli, M., Liuni, S., 1988. A backtranslation method based on codon usage strategy. Nucleic Acids Res. 16, 1715–1728.

Puigbo, P., Guzman, E., Romeu, A., Garcia-Vallve, S., 2007. OPTIMIZER: a web server for optimizing the codon usage of DNA sequences. Nucleic Acids Res. 35, W126–W131.

Reed, L.J., Muench, H.A., 1938. A simple method of estimating fifty per cent endpoints. Am. J. Hyg. 27, 493–497.

Rima, B.K., McFerran, N.V., 1997. Dinucleotide and stop codon frequencies in single-stranded RNA viruses. J. Gen. Virol. 78 (Pt 11), 2859–2870.

Risatti, G.R., Holinka, L.G., Fernandez Sainz, I., Carrillo, C., Lu, Z., Borca, M.V., 2007b. N-linked glycosylation status of classical swine fever virus strain Brescia E2 glycoprotein influences virulence in swine. J. Virol. 81, 924–933.

Risatti, G.R., Borca, M.V., Kutish, G.F., Lu, Z., Holinka, L.G., French, R.A., Tulman, E.R., Rock, D.L., 2005a. The E2 glycoprotein of classical swine fever virus is a virulence determinant in swine. J Virol 79, 3787–3796.

Risatti, G.R., Holinka, L.G., Carrillo, C., Kutish, G.F., Lu, Z., Tulman, E.R., Sainz, I.F., Borca, M.V., 2006. Identification of a novel virulence determinant within the E2 structural glycoprotein of classical swine fever virus. Virology 355, 94–101.

Risatti, G.R., Holinka, L.G., Lu, Z., Kutish, G.F., Tulman, E.R., French, R.A., Sur, J.H., Rock, D.L., Borca, M.V., 2005b. Mutation of E1 glycoprotein of classical swine fever virus affects viral virulence in swine. Virology 343, 116–127.

Risatti, G.R., Holinka, L.G., Fernandez Sainz, I., Carrillo, C., Kutish, G.F., Lu, Z., Zhu, J., Rock, D.L., Borca, M.V., 2007a. Mutations in the carboxyl terminal region of E2 glycoprotein of classical swine fever virus are responsible for viral attenuation in swine. Virology 364, 371–382.

Rothberg, P.G., Wimmer, E., 1981. Mononucleotide and dinucleotide frequencies, and codon usage in poliovirion RNA. Nucleic Acids Res. 9, 6221–6229.

Sanger, F., Nicklen, S., Coulson, A.R., 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74, 5463–5467.

Sharp, P.M., Li, W.H., 1987. The codon adaptation index—a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res. 15, 1281–1295.

Sharp, P.M., Tuohy, T.M., Mosurski, K.R., 1986. Codon usage in yeast: cluster analysis clearly differentiates highly and lowly expressed genes. Nucleic Acids Res. 14, 5125–5143.

Shen, S.H., Stauft, C.B., Gorbatsevych, O., Song, Y., Ward, C.B., Yurovsky, A., Mueller, S., Futcher, B., Wimmer, E., 2015. Large-scale recoding of an arbovirus genome to rebalance its insect versus mammalian preference. Proc. Natl. Acad. Sci. USA 112, 4749–4754.

Simmonds, P., 2012. SSE: a nucleotide and amino acid sequence analysis platform. BMC Res. Notes 5, 50.

Terpstra, C., Woortmeyer, R., Barteling, S.J., 1990. Development and properties of a cell culture produced vaccine for hog cholera based on the Chinese strain. Dtsch. Tierarztl. Wochenschr. 97, 77–79.

Thiel, H.J., Stark, R., Weiland, E., Rumenapf, T., Meyers, G., 1991. Hog cholera virus: molecular composition of virions from a pestivirus. J. Virol. 65, 4705–4712.

Tulloch, F., Atkinson, N.J., Evans, D.J., Ryan, M.D., Simmonds, P., 2014. RNA Virus Attenuation by Codon Pair Deoptimisation is an Artefact of Increases in CpG/UpA Dinucleotide Frequencies. Elife 3:e04531.

van Gennip, H.G., van Rijn, P.A., Widjojoatmodjo, M.N., de Smit, A.J., Moormann, R.J., 2000. Chimeric classical swine fever viruses containing envelope protein E (RNS) or E2 of bovine viral diarrhoea virus protect pigs against challenge with

CSFV and induce a distinguishable antibody response. Vaccine 19, 447–459.
Van Gennip, H.G., Vlot, A.C., Hulst, M.M., De Smit, A.J., Moormann, R.J., 2004. Determinants of virulence of classical swine fever virus strain Brescia. J. Virol. 78, 8812–8823.
Wang, B., Yang, C., Tekes, G., Mueller, S., Paul, A., Whelan, S.P., Wimmer, E., 2015. Recoding of the vesicular stomatitis virus L gene by computer-aided design provides a live, attenuated vaccine candidate. MBio 6.
Weiland, E., Stark, R., Haas, B., Rumenapf, T., Meyers, G., Thiel, H.J., 1990. Pestivirus glycoprotein which induces neutralizing antibodies forms part of a disulfide-linked heterodimer. J. Virol. 64, 3563–3569.
Weiland, F., Weiland, E., Unger, G., Saalmuller, A., Thiel, H.J., 1999. Localization of pestiviral envelope proteins E(rns) and E2 at the cell surface and on isolated particles. J. Gen. Virol. 80 (Pt 5), 1157–1165.
Yang, C., Skiena, S., Futcher, B., Mueller, S., Wimmer, E., 2013. Deliberate reduction of hemagglutinin and neuraminidase expression of influenza virus leads to an ultraprotective live vaccine in mice. Proc. Natl. Acad. Sci. USA 110, 9481–9486.
Zsak, L., Lu, Z., Kutish, G.F., Neilan, J.G., Rock, D.L., 1996. An African swine fever virus virulence-associated gene NL-S with similarity to the herpes simplex virus ICP34.5 gene. J. Virol. 70, 8865–8871.