

Genetic diversity and evolution of the emerging picornavirus *Senecavirus A*

Lok R. Joshi^{1,2†}, Kristin A. Mohr¹, Danielle Gava², Gerald Kutish³, Alaire S. Buysse¹, Fabio A. Vannucci⁴, Pablo E. Piñeyro⁵, Beate M. Crossley⁶, John J. Schiltz⁷, Melinda Jenkins-Moore⁷, Leo Koster⁷, Rachel Tell⁷, Rejane Schaefer², Douglas Marthaler^{4,8} and Diego G. Diel^{1,2,*†}

Abstract

Senecavirus A (SVA) is an emerging picornavirus that causes vesicular disease (VD) in swine. The virus has been circulating in swine in the United States (USA) since at least 1988, however, since 2014 a marked increase in the number of SVA outbreaks has been observed in swine worldwide. The factors that led to the emergence of SVA remain unknown. Evolutionary changes that accumulated in the SVA genome over the years may have contributed to the recent increase in disease incidence. Here we compared full-genome sequences of historical SVA strains (identified before 2010) from the USA and global contemporary SVA strains (identified after 2011). The results from the genetic analysis revealed 6.32% genetic divergence between historical and contemporary SVA isolates. Selection pressure analysis revealed that the SVA polyprotein is undergoing selection, with four amino acid (aa) residues located in the VP1 (aa 735), 2A (aa 941), 3C (aa 1547) and 3D (aa 1850) coding regions being under positive/diversifying selection. Several aa substitutions were observed in the structural proteins (VP1, VP2 and VP3) of contemporary SVA isolates when compared to historical SVA strains. Some of these aa substitutions led to changes in the surface electrostatic potential of the structural proteins. This work provides important insights into the molecular evolution and epidemiology of SVA.

INTRODUCTION

Senecavirus A (SVA) is an emerging pathogen, causing vesicular disease (VD) in swine, that is clinically indistinguishable from foot-and-mouth disease (FMD). The virus belongs to the genus *Senecavirus* of the family *Picornaviridae* [1]. SVA is a non-enveloped virus containing an icosahedral capsid that is ~27 nm in diameter. The SVA genome consists of a positive-sense single-stranded RNA molecule of ~7.2 kb, which encodes a single 2181 amino acid (aa) polyprotein. The

viral polyprotein is flanked by 5' and 3' untranslated regions (UTRs). The 5' UTR contains the internal ribosomal entry site (IRES), a secondary RNA structure that is important for cap-independent translation of viral RNA [2]. Upon translation, the viral polyprotein is processed by the 3C^{pro}, a virus-encoded protease, resulting in 12 viral proteins (5'-L-VP4-V P2-VP3-VP1-2A-2B-2C-3A-3B-3C-3D-3'). The VPs (VP1–4) compose the viral capsid and are involved in receptor binding

Received 29 May 2019; Accepted 11 November 2019; Published 20 December 2019

Author affiliations: ¹Department of Veterinary and Biomedical Sciences, Animal Disease Research and Diagnostic Laboratory, South Dakota State University, Brookings, SD 57007, USA; ²Embrapa Swine and Poultry, Concórdia, Santa Catarina, Brazil; ³Department of Pathobiology, University of Connecticut, Storrs, CT 06269, USA; ⁴Department of Population Medicine, University of Minnesota, St Paul, MN 55455, USA; ⁵Department of Veterinary Diagnostic and Production Animal Medicine, College of Veterinary Medicine, Iowa State University, Ames, IA, USA; ⁶California Animal Health and Food Safety Laboratory System, Department of Medicine and Epidemiology, School of Veterinary Medicine, University of California, Davis, CA, USA; ⁷National Veterinary Services Laboratories, Animal and Plant Health Inspection Service, USDA, Ames, IA, USA; ⁸Veterinary Diagnostic Laboratory, Kansas State University, Manhattan, KS, USA.

*Correspondence: Diego G. Diel, dgdiel@cornell.edu

Keywords: Senecavirus A; Seneca Valley virus; SVV; evolution; genetic diversity.

Abbreviations: ADRDL, Animal Disease Research and Diagnostic Laboratory; BIC, Bayesian information criteria; CAHFS, California Animal Health and Food Safety Laboratory; CCoV, canine enteric coronaviruses; FEL, fixed effects likelihood; FMD, foot-and-mouth disease; FSIS, Food Safety and Inspection Service; FUBAR, Fast Unconstrained bayesian Approximation for Inferring Selection; GTR, general time-reversible; IRES, internal ribosomal entry site; MEME, mixed effects model of evolution; NVSL, National Veterinary Services Laboratories; ORF, open reading frame; PEDV, porcine epidemic diarrhea virus; SHIC, Swine Health Information Center; SLAC, single-likelihood ancestor counting; SVA, *Senecavirus A*; UMN, University Of Minnesota; UTR, untranslated region; VD, vesicular disease.

†Present address: Department of Population Medicine and Diagnostic Sciences, Animal Health Diagnostic Center, College of Veterinary Medicine, Cornell University, Ithaca, NY, USA.

and entry, whereas the non-structural proteins are involved in virus genome transcription and replication [2].

Historically, SVA has been detected in the US swine population since at least 1988, when the United States Department of Agriculture National Veterinary Services Laboratories (NVSL) isolated a picorna-like virus from samples obtained in a farm in Minnesota with a history of piglet diarrhoea and stillbirths. Electron microscopy revealed a picorna-like virus particle that upon biochemical testing was shown to share characteristics with enteroviruses (NVSL, unpublished data). In 2002, an unknown picornavirus was isolated as a cell culture contaminant in Maryland and designated Seneca Valley Virus-001 (SVV-001); the complete genome sequence of this virus was published in 2008 [2] and the virus was recently renamed Senecavirus A (SVA) [1]. Work by Knowles and collaborators established that SVA strain SVV-001 shared considerable sequence identity with seven historical picorna-like viruses isolated from pigs in the USA exhibiting diverse clinical manifestations, including VD (Table 1) [3]. While the first published report of SVA being associated with VD occurred in 2007 [4], the virus has been intermittently isolated from cases of swine showing vesicular lesions that date at least as far back as 1997 (Table 1). In the 2007 case, vesicular lesions were observed and SVA was detected by reverse transcriptase PCR and virus isolation in pigs originating from Manitoba, Canada at a harvest facility in Minnesota [4] (NVSL, unpublished data). Another detection of SVA in a case of vesicular lesions in swine was reported in Indiana in 2010 [5]. Despite the association of SVA with VD [4, 5], attempts to reproduce the disease in pigs following experimental inoculations with historical SVA strains were not successful (NVSL, unpublished data) [6, 7].

In November 2014, several outbreaks of SVA associated with VD were reported in Brazil [8, 9]. During June–July 2015, SVA outbreaks were reported in China [10] and the USA [11–14]. Subsequently, SVA was detected in Colombia [15], Thailand [16] and Vietnam [17]. Recent studies with contemporary SVA strains have successfully reproduced VD in experimentally inoculated pigs, confirming the aetiological role of SVA in VD in swine [18, 19]. These studies revealed that SVA infection is characterized by the development of vesicles on the snout, oral mucosa, sole, interdigital space and/or the coronary bands of infected animals [18, 19]. Lesions are first observed between days 3–5 post-inoculation (p.i.) and usually resolve by day 14–16 p.i. [19]. Interestingly, a study comparing the pathogenicity of a contemporary SVA strain and the historic SVA strain SVV-001 revealed that this historic isolate is not pathogenic to finishing pigs (~50 kg) [7].

The factors that contributed to the emergence of SVA in swine populations around the world remain unknown. Importantly, genetic changes have often been associated with the emergence or re-emergence of viral pathogens [20]. For example, high rates of recombination in the spike gene led to the emergence of lethal canine enteric coronaviruses (CCoV), a virus that previously caused only mild gastrointestinal clinical signs in affected animals [21]. Similarly, genetic changes in

the spike gene of porcine epidemic diarrhoea virus (PEDV) may have contributed to the emergence of highly virulent strains of the virus [22–24]. Therefore, it is possible that the evolutionary changes that accumulated over the years in the genome of contemporary SVA strains [13] may have contributed to the differences in the pathogenicity between contemporary and historical SVA strains [7], and, perhaps, the increased incidence of SVA-induced VD [25]. Here, we performed genetic and evolutionary analysis of historical and contemporary SVA isolates using complete genome sequences from viruses isolated between 1988 and 2017 across different geographical locations.

METHODS

SVA-positive samples and isolates

Multiple samples and/or isolates from pigs presenting characteristic VD lesions that tested positive for SVA at the Animal Disease Research and Diagnostic Laboratory at SDSU, the Veterinary Diagnostic Laboratory at the UMN, or the California Animal Health and Food Safety Laboratory System at UC Davis were included in the present study. Additionally, NVSL virus isolates from pigs exhibiting diverse clinical presentations, including VD, were included. Samples collected in Brazil and submitted to Embrapa Swine and Poultry were also included. All samples and/or virus isolates were obtained between 1988 and 2017. A brief description of the viruses sequenced in our study and their GenBank accession numbers are presented in Table 1.

Sequencing

Complete genome sequences of contemporary SVA strains were obtained using the Illumina MiSeq platform or Sanger sequencing as previously described [9, 13, 19]. Briefly, viral RNA was extracted, and cDNA was prepared using reverse transcriptase. The cDNA was subjected to Illumina sequencing or to PCR amplification followed by Sanger sequencing. For next-generation sequencing (NGS), libraries were prepared according to the Nextera XT (Illumina) library preparation protocol and then sequenced using the Illumina MiSeq platform [9, 19]. For Sanger sequencing, nine pairs of overlapping primers (sequences available upon request) covering the full SVA genome were used for PCR amplification and the amplicons were subjected to DNA sequencing [19]. Historical SVA whole-genome sequences were generated using the Ion Torrent Personal Genome Machine, according to the manufacturer's protocols (Thermo Scientific). Briefly, RNA libraries were constructed using the Ion Total RNA-Seq kit v2 with 10 µl input purified nucleic acid. Libraries were amplified using the Ion PGM Hi-Q Chef kit on the Ion Chef System and sequencing was performed using the Ion PGM 200 v2 Sequencing kit on an Ion 316 chip. All sequences generated in our study, independent of the sequencing platform used, were subjected to QA/QC standards set by the instruments' manufacturers. Only sequences passing these standards were used to assemble complete SVA genomes.

Table 1. Description of the SVA isolates sequences in the present study.

Isolate name	Accession no.	Year of Isolation	Country of Isolation (state)	Laboratory of origin	Remarks
USA/NC88-23626/1988	MN233026	1988	USA (NC)	NVSL	
USA/MN88-36695/1988	MN233027	1988	USA (MN)	NVSL	Diarrhoea and stillborn piglets, no other lesions on necropsy
USA/89-47552/1989	MN233028	1989	USA (IA)	NVSL	
USA/NJ90-10324/1989	MN233029	1989	USA (NJ)	NVSL	Lameness, death, no other lesions on necropsy
USA/IA90-23664/1990	MN233030	1990	USA (IA)	NVSL	
USA/IL92-48963/1992	MN233031	1992	USA (IL)	NVSL	Herd with chronic respiratory and reproductive problems; samples from 8-week-old pig that was poor-doing with no significant post-mortem lesions
USA/IL94-9356/1993	MN233032	1993	USA (IL)	NVSL	
USA/LA97-98061/1997	MN233034	1997	USA (LA)	NVSL	Vesicular disease
USA/MN99-29256/1999	MN233033	1999	USA (MN)	NVSL	
USA/IL00-66289/2000	MN233017	2000	USA (IL)	NVSL	
USA/CA01-131395/2001	MN233018	2001	USA (CA)	NVSL	Vesicular disease
USA/IL 01-84124/2001	MN233019	2001	USA (IL)	NVSL	Isolated from pigs following experimental inoculation with isolate USA/IL00-66289/2000.
USA/SC05-363649/2004	MN233020	2004	USA (SC)	NVSL	Feral pig sent in for PRV testing
USA/IA05-401302/2005	MN233021	2005	USA (IA)	NVSL	Vesicular disease
USA/TN06-429971/2006	MN233022	2006	USA (TN)	NVSL	
CAN/07-503297/2007	MN233023	2007	Canada	NVSL	Vesicular disease
USA/MI11-055910/2011	KC667560	2011	USA (MI)	NVSL	Neurological signs in adult sow
USA/MO15-029085/2015	MN233024	2015	USA (MO)	NVSL	Vesicular disease
USA/KS15-031348/2015	MN233025	2015	USA (KS)	NVSL	Vesicular disease
SVA/US/MN/04/2015	MK333629	2015	USA (MN)	UMN	Vesicular fluid
SVA/US/MN/05/2015	MK333630	2015	USA (MN)	UMN	Vesicular fluid
SVA/US/MN/06/2015	MK333631	2015	USA (MN)	UMN	Vesicular fluid
SVA/US/MN/07/2015	MK333632	2015	USA (MN)	UMN	Vesicular fluid
SVA/US/IL/51-4	MH634527	2015	USA (IL)	ADRDL	Vesicular foot lesion
SVA/US/IL/51-7	MH634529	2015	USA (IL)	ADRDL	Vesicular fluid
SVA/US/IL/72-1	MH634530	2015	USA (IL)	ADRDL	Vesicular foot lesion
SVA/US/IL/72-2	MH634531	2015	USA (IL)	ADRDL	Vesicular snout lesion
SVA/US/IL/72-3	MH634532	2015	USA (IL)	ADRDL	Vesicular snout lesion
SVA/US/IL/81-4a	MH634533	2015	USA (IL)	ADRDL	Vesicular lesion swab
SVA/US/IL/81-5a	MH634534	2015	USA (IL)	ADRDL	Vesicular lesion swab
SVA/BRA/SC/011_15	MF615501	2015	Brazil (SC)	Embrapa	Vesicular snout lesion
SVA/BRA/SC/128_15	MF615506	2015	Brazil (SC)	Embrapa	Vesicular coronary band lesion

Continued

Table 1. Continued

Isolate name	Accession no.	Year of Isolation	Country of Isolation (state)	Laboratory of origin	Remarks
SVA/BRA/SC/235_15	MF615507	2015	Brazil (SC)	Embrapa	Vesicular fluid
SVA/BRA/SC/244_15	MF615508	2015	Brazil (SC)	Embrapa	Small intestine
SVA/BRA/SC/245_15	MF615509	2015	Brazil (SC)	Embrapa	Vesicular coronary band lesion
SVA/BRA/SC/072_16	MF615502	2016	Brazil (SC)	Embrapa	Vesicular lesion on snout and coronary band
SVA/BRA/SC/075_16	MF615503	2016	Brazil (SC)	Embrapa	Lung
SVA/BRA/SC/077_16	MF615504	2016	Brazil (SC)	Embrapa	Vesicular coronary band lesion
SVA/BRA/SC/078_16	MF615505	2016	Brazil (SC)	Embrapa	Vesicular lesion on snout and coronary band
SVA/US/MN/09/2016	MK333633	2016	USA (MN)	UMN	Vesicular fluid
SVA/US/MN/10/2016	MK333634	2016	USA (MN)	UMN	Vesicular fluid
SVA/US/MN/11/2016	MK333635	2016	USA (MN)	UMN	Vesicular fluid
SVA/US/MN/12/2016	MK333636	2016	USA (MN)	UMN	Vesicular fluid
SVA/US/MN/13/2016	MK333637	2016	USA (MN)	UMN	Vesicular fluid
SVA/US/CA/17-42-3D	MH634506	2017	USA (CA)	CAHFS/ADRDL	Arrived at plant with vesicular lesions on coronary bands and snouts; pooled swab from affected hogs
SVA/US/CA/17-65-4D	MH634507	2017	USA (CA)	CAHFS/ADRDL	Ruptured snout lesion
SVA/US/CA/17-59-3D	MH634508	2017	USA (CA)	CAHFS/ADRDL	Large healing erosion on snout, small erosion on RR coronary band
SVA/US/CA/17-60-2D	MH634509	2017	USA (CA)	CAHFS/ADRDL	Erosion on snout and coronary band, pooled sample
SVA/US/CA/17-06-3D	MH634510	2017	USA (CA)	CAHFS/ADRDL	Coronary lesions and snout lesion
SVA/US/CA/17-57-11D	MH634511	2017	USA (CA)	CAHFS/ADRDL	2 snout lesions approximately 0.5 cm in diameter, both chronic
SVA/US/CA/17-38-5D	MH634512	2017	USA (CA)	CAHFS/ADRDL	Arrived at plant with vesicular lesions on coronary bands and snouts; pooled swab from affected hogs
SVA/US/CA/17-95-9D	MH634513	2017	USA (CA)	CAHFS/ADRDL	Approximately 90% of pigs affected, vesicles on coronary bands
SVA/US/CA/17-28-4D	MH634514	2017	USA (CA)	CAHFS/ADRDL	Intact vesicles on snout, pooled sample (2 pigs)
SVA/US/CA/17-58-2D	MH634515	2017	USA (CA)	CAHFS/ADRDL	Vesicular lesion on snout, erosions on coronary bands and lameness
SVA/US/CA/17-97-3D	MH634516	2017	USA (CA)	CAHFS/ADRDL	Approximately 60% affected hogs, vesicles on coronary bands
SVA/US/CA/17-96-4D	MH634517	2017	USA (CA)	CAHFS/ADRDL	Coronary band sloughing
SVA/US/CA/17-43-3D	MH634518	2017	USA (CA)	CAHFS/ADRDL	Arrived at plant with coronary lesions, pooled sample

Continued

Table 1. Continued

Isolate name	Accession no.	Year of Isolation	Country of Isolation (state)	Laboratory of origin	Remarks
SVA/US/CA/17-06-5D	MH634519	2017	USA (CA)	CAHFS/ADRDL	Arrived at plant with coronary lesions, pooled sample
SVA/US/CA/17-43-5D	MH634520	2017	USA (CA)	CAHFS/ADRDL	Snout and foot lesions reported by FSIS at slaughter establishment
SVA/US/CA/17-94-9D	MH634521	2017	USA (CA)	CAHFS/ADRDL	Arrived at plant with vesicular lesions on coronary bands and snouts, pooled swab
SVA/US/CA/17-96-3D	MH634522	2017	USA (CA)	CAHFS/ADRDL	Approximately 90% affected: vesicles on coronary bands

NVSL, National Veterinary Services Laboratories; UMN, University of Minnesota; ADRDL, Animal Disease Research and Diagnostic Laboratory; CAHFS, California Animal Health and Food Safety Laboratory. FSIS, Food Safety and Inspection Service.

Genome assembly

The quality of the sequence reads was assessed using FastQC. Lower quality reads were trimmed using FastX version 0.0.14 (http://hannonlab.cshl.edu/fastx_toolkit/). FastX was ran using the fastq_quality_filter with -q set to 27 and -p set to 80. SVA genomes were *de novo* assembled using Ray v2.3 [26], Velvet 1.2.10 [27], Newbler 2.9 [28] or SeqMan NGen 13.0. (DNASTAR). Contigs were identified using BLAST searches.

Alignment and phylogenetic analysis

A total of 117 whole SVA genome sequences (61 newly sequenced and 56 sequences obtained from GenBank) were aligned using MUSCLE v 3.5 [29]. Selection of the best-fit model of nucleotide substitution was performed using MEGA x and the Bayesian information criteria (BIC). The phylogenetic tree was constructed in MEGA x [30] using the maximum-likelihood method, the general time-reversible (GTR) model and bootstrap tests of 500 replicates.

Pairwise distance analysis was performed in MEGA x. A heatmap showing pairwise distance comparison was generated using ClustVis [31]. The SVA sequences were clustered using Euclidean distance and complete linkage implemented with ClustVis.

Selection pressure analysis

Site-specific selection pressures for the SVA polyprotein (1–2181) were measured as the ratio of non-synonymous (d_N) to synonymous (d_S) nucleotide substitutions per site (d_N/d_S) [32]. Tests for positive/diversifying selection were performed using four methods: single-likelihood ancestor counting (SLAC), fixed effects likelihood (FEL), fast unconstrained Bayesian approximation for inferring selection (FUBAR) and mixed effects model of evolution (MEME) [32] implemented in HyPhy 2.2 using the Datammonkey server [33]. These methods estimate selection in a phylogenetic context [[32]]. Sites presenting a *P*-value <0.05 when analysed by SLAC, FEL or MEME were reported. FUBAR results obtained through

a Bayesian approach were considered significant when the posterior probability was greater than 0.9. Multiple testing or false discovery rate correction on individual site results was not performed here as: (1) the methods used in our analyses are calibrated so no excessive false positives are generated and (2) key assumptions of methods for correcting false discovery rates cannot be applied to the site-specific testing performed [32].

Protein structure analysis

The aa sequences corresponding to the SVA capsid proteins (VP4, VP2, VP3 and VP1, aa 80–937; VPs) of four historical and four contemporary SVA isolates was selected for protein structure modelling. The VPs aa sequences were aligned using MUSCLE and visualized using Jalview 2 [34]. The secondary structure of capsid proteins was predicted using the homology based modelling server SWISS-MODEL [35] using the ‘User-Template’ module. The aa sequence 3CJI.pdb was used as a reference template as it represents the structure of Seneca Valley Virus-001 capsid available in the Protein Data Bank (<https://www.rcsb.org/structure/3CJI>). The molecular structures were visualized using the University of California San Francisco (UCSF) Chimera package [36].

RESULTS

SVA strains included in the analysis

Several historical and contemporary SVA strains obtained between 1988 and 2017 were sequenced in this study (Table 1). The earliest SVA complete genome sequences include USA/NC88-23626/1988 and USA/MN88-36695/1988, obtained in 1988 from pigs from North Carolina and Minnesota, respectively. The USA/MN88-36695/1988 was isolated from tissues collected from a farm with history of piglet diarrhoea and stillbirths. SVA strain USA/LA97-98061/1997 was isolated in 1997 from a pig presenting vesicular lesions on the feet, legs and mouth. When this isolate was used in experimental

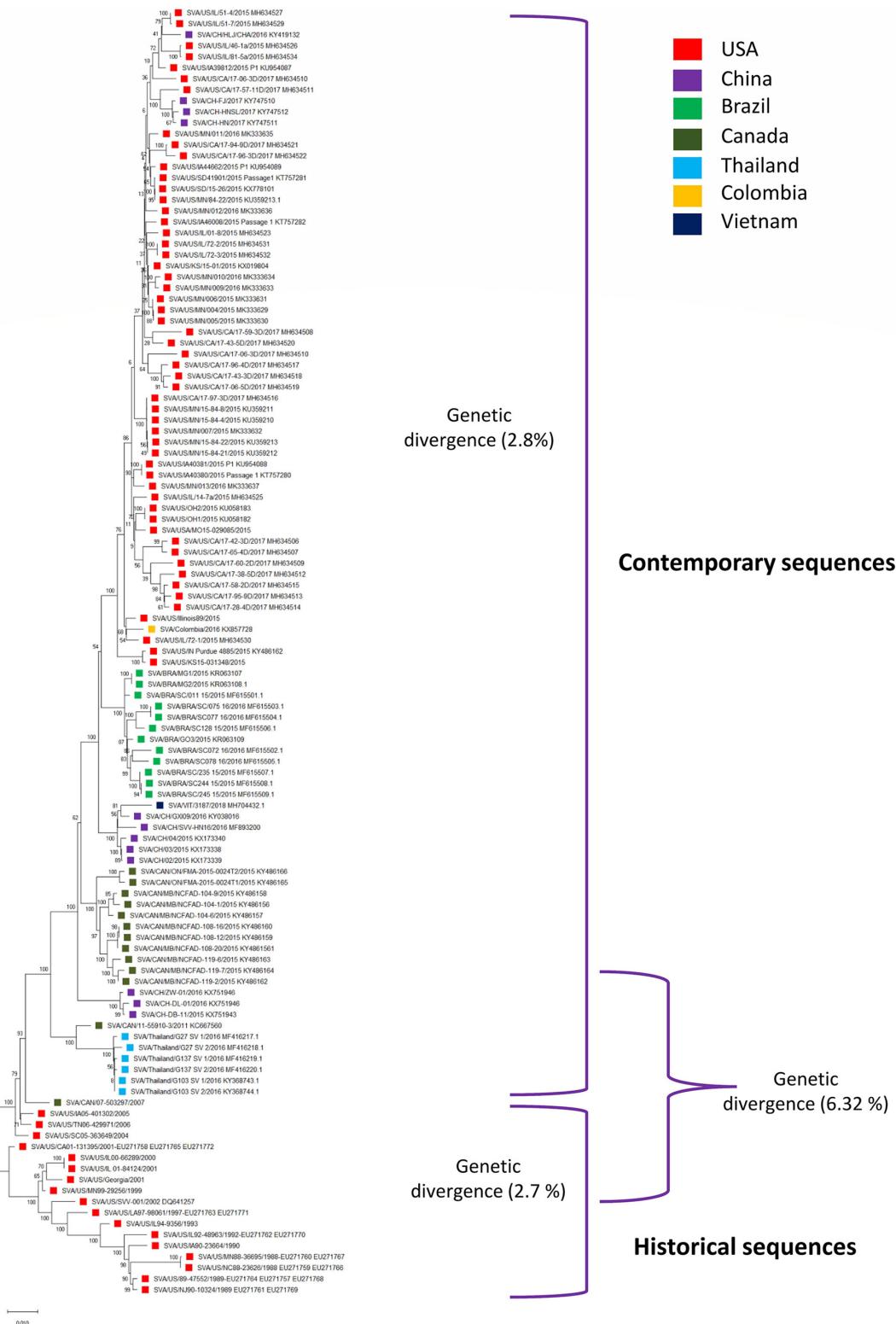


Fig. 1. Phylogenetic tree based on whole-genome sequences of historical (detected between 1988–2010) and contemporary SVA strains (detected between 2011–2017). The phylogenetic tree was constructed using the maximum-likelihood method in MEGA X using bootstrap tests of 500 replicates. Pairwise distance was calculated using MEGA X.

Table 2. Genetic distances between SVA sequences from different countries

Country*	% difference between the countries				
	USA	Canada	China	Brazil	Thailand
USA	–				
Canada	3.08†	–			
China	2.48	2.94	–		
Brazil	2.71	2.97	2.8	–	
Thailand	4.85	4.14	4.62	4.53	–

*The SVA isolates were grouped according to their geographical origin.

†The genetic distances between SVA isolates from different countries was inferred based on the number of base substitutions per site representing the average over all sequence pairs between different countries using MEGA X.

–No difference

inoculation studies by NVSL, no overt disease was observed in inoculated pigs. The SVA isolate USA/IL00-66289/2000 was also used in a pathogenesis study. While none of the pigs inoculated with this isolate presented clinical signs, the virus was reisolated from tissues, as well as from nasal and rectal swabs, from inoculated animals (NVSL, unpublished data). SVA strains USA/CA01-131395/2001 and USA/IA05-401302/2005 were also isolated from pigs presenting VD (NVSL, unpublished data).

Additionally, several contemporary SVA strains obtained between 2015 and 2017 in the USA and Brazil were included in this study. All but one of the contemporary SVA strains presented with a history of VD. The SVA strains isolated in California in 2017 were obtained from finishing pigs at slaughter houses. The Brazilian SVA strains were isolated from animals presenting VD in 2015 or 2016 in the state of Santa Catarina (Table 1), which holds the status of state free of FMDV without vaccination (http://www.oie.int/fileadmin/Home/eng/Animal_Health_in_the_World/map/A_Brazil.jpg).

Phylogenetic analysis

A total of 117 complete SVA genome sequences were included in the analysis. SVA sequences obtained up to 2010 were classified as historical, whereas all isolates obtained from 2011 onwards were classified as contemporary strains. Phylogenetic analysis using whole SVA genome sequences revealed that historical SVA strains obtained between 1989 and 2001 form a distinct phylogenetic clade (strains NJ90-10324/1989 to CA01-13195/2001). Additionally, historical SVA strains IA05-401302/2005 and TN06-429971/2006 form a separate clade, while SVA strain SC05-363649/2004 appears to be the ancestor to the contemporary SVA strains (Fig. 1).

Phylogenetic analysis revealed that most SVA isolates cluster according to their geographical origin, with viruses from each country forming separate clades in the tree. Interestingly, a few

Table 3. Genetic distances between SVA sequences obtained within each country

Country	Genetic distance (%)
USA	1.3†
Canada	1.4
China	2.4
Brazil	0.9
Thailand	0.2

*The SVA isolates were grouped according to their country of origin.

†The genetic distances between the SVA isolates within each countries were inferred using MEGA X. The values are presented as percentages.

SVA strains from China (CH/HLJ/CHA/2016, CH-FJ/2017, CH-HNSL/2017 and CH-HN/2017) and one from Colombia (Colombia/2016) were found in distant clades, clustering with contemporary US strains (Fig. 1). These results show the evolutionary divergence between historical and contemporary SVA strains. Clustering of contemporary SVA strains in separate phylogenetic groups based on the country of origin suggests that the majority of these viruses are evolving independently within the swine population of each affected country. The close phylogenetic relationship between a few Chinese, a Colombian and US contemporary strains, however, suggests that strains from one region may occasionally be transferred to swine populations of other geographic areas.

Pairwise distance comparison

The pairwise genetic distance between SVA strains ranged between 0 to 9.5% at the nucleotide level. The nucleotide distances between SVA isolates obtained in different countries or within each country are presented in Tables 2 and 3, respectively. The highest genetic distance (9.5%) was observed between a historical SVA strain from 1988 (SVA/USA/NC88-23626/1988) and a contemporary strain from 2017 (SVA/US/CA/17-60-2D/2017) (data not shown). A heatmap illustrating pairwise distances between available SVA sequences is presented in Fig. 2. High genetic distance, as indicated by red colour on the heat map, was observed between the contemporary and historical sequences. Clustering based on Euclidean distance shows that historical SVA isolates form three distinct clusters: (1) isolates obtained between 1988 and 1993, (2) isolates obtained between 1997 and 2002, and (3) isolates obtained between 2004 and 2007. The genetic distance between 1988–1993 and 1997–2002 historical SVA strains is ~3.8% (Table 4). These differences are illustrated in the heatmap (purple box, Fig. 2). The genetic distance between 1988–1993 and 2004–2007 historical strains is ~5.5% [Table 4, Fig. 2 (black box in the heatmap)]. Interestingly, contemporary SVA strains from Thailand are closely related to historical SVA strains obtained in the USA and Canada between 2004 and 2007 (Fig. 2). The genetic distance between contemporary

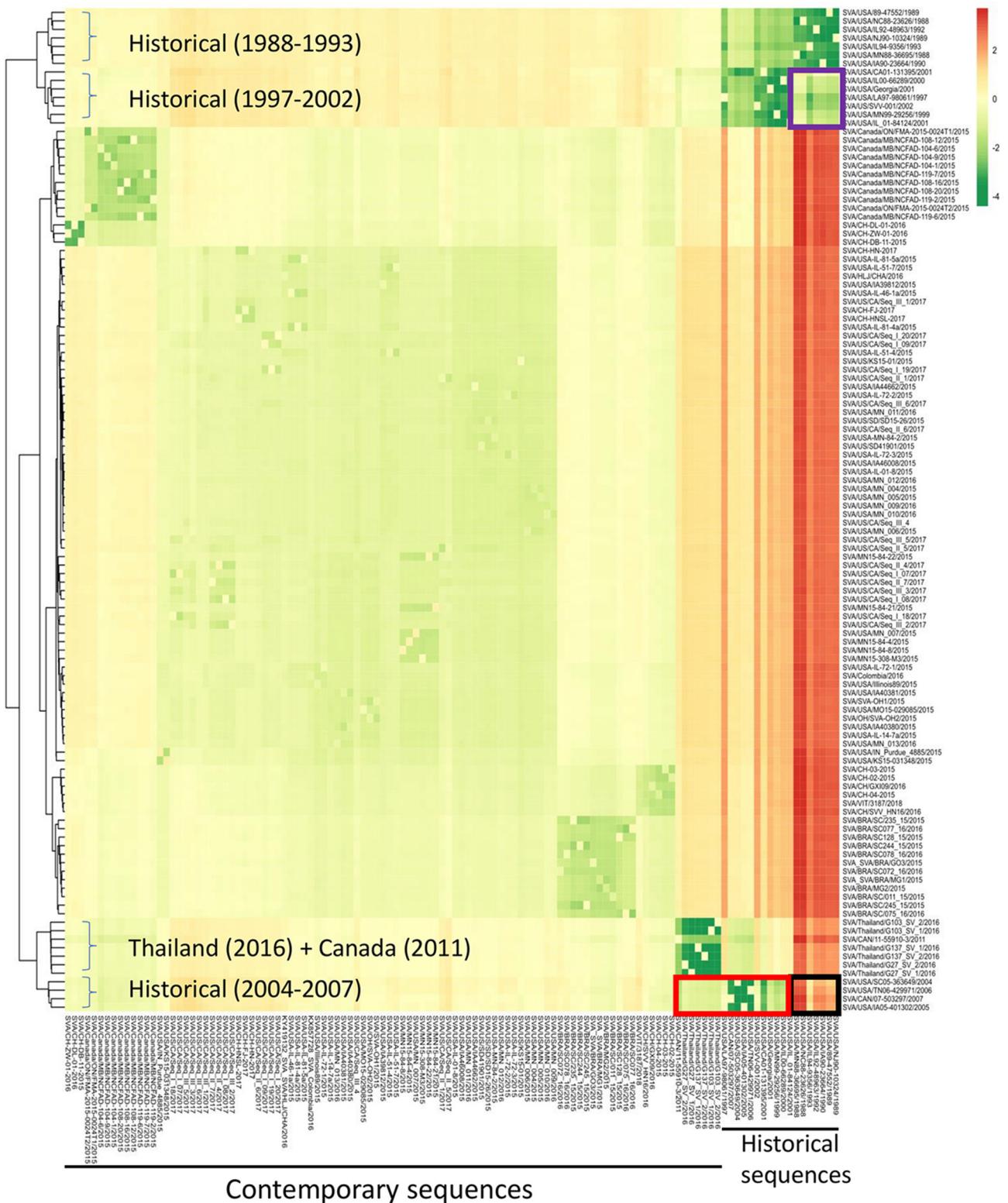


Fig. 2. Heatmap showing pairwise genetic distance. The pairwise distance between the sequences/strains was calculated using MEGA x. Heatmap representation of pairwise distance was created using ClustVis. Clustering of the SVA strains was performed on the basis of Euclidean distance.

Table 4. SVA genetic variation between historical isolates

Strains	Historical (1988–1993)	Historical (1997–2002)	Historical (2004–2007)
Historical (1988–1993)	–		
Historical (1997–2002)	3.8	–	
Historical (2004–2007)	5.5	2.90	–

–No difference

Thai strains and 2004–2007 historical strains is only 3.3%, whereas, contemporary SVA strains from Thailand differ by 5.0 and 7.3% from other historical SVA strains (1997–2002 and 1988–1993) [Fig. 2 (red box in heatmap)]. Similar to the results from the phylogenetic analysis (Fig. 1), contemporary SVA strains from Thailand cluster with a strain isolated from Canada in 2011 (SVA/CAN-11-55910-3/2011) (Fig. 2).

Selection pressure analysis

The overall non-synonymous to synonymous substitution rate (d_N/d_S) over the entire SVA open reading frame (ORF) was estimated at 0.1, which suggests predominantly purifying selection [32]. Most of the codons in the SVA ORF do not seem to be under strong negative/purifying nor positive/diversifying selection (Table 5). Of the 2181 codons that comprise the viral polyprotein, the maximum number of

Table 6. Sites under positive selection identified by SLAC, FUBAR, FEL and MEME methods

Methods ^a	Positive selected sites
SLAC	735, 941
FUBAR	735, 941, 1547, 1850
FEL	475, 664, 735, 770, 941, 1003, 1148, 1427, 1480, 1547, 1780, 1547, 1780, 1850, 1941, 2112
MEME	34, 43, 56, 144, 242, 260, 289, 306, 428, 429, 430, 440, 475, 482, 484, 486, 491, 492, 497, 511, 515, 582, 595, 603, 648, 654, 664, 683, 735, 736, 770, 793, 828, 894, 902, 903, 941, 948, 988, 993, 999, 1003, 1036, 1059, 1079, 1148, 1176, 1344, 1348, 1349, 1355, 1357, 1362, 1366, 1378, 1379, 1382, 1383, 1411, 1412, 1427, 1438, 1469, 1472, 1476, 1478, 1480, 1488, 1547, 1574, 1582, 1589, 1679, 1716, 1729, 1743, 1755, 1765, 1767, 1780, 1810, 1839, 1850, 1851, 1866, 1909, 1941, 2051, 2060, 2075, 2112, 2154

^a Sites in bold indicate the codons that were identified to be under positive selection by at least three of the four methods used.

codons presenting strong evidence of negative selection was 742, using the FEL method, and the maximum number of positively selected codons was 92, as detected with the MEME method (Table 5).

Only four residues (735, 941, 1547 and 1850) located in the coding region of SVA VP1, 2A, 3C and 3D proteins, respectively, were detected as being under strong positive selection using at least three (FUBAR, FEL and MEME) of the four

Table 5. Number of codons under positive and negative selection identified by SLAC, FEL, FUBAR and MEME methods

Proteins	Codon position	Length of protein	SLAC		FEL		FUBAR*		MEME*	
			Negative	Positive	Negative	Positive	Positive	Positive	Positive	Positive
L	1–79	79	7	0	15	0	0	0	3	
VP4	80–150	71	9	0	19	0	0	0	1	
VP2	151–434	284	65	0	100	0	0	0	7	
VP3	435–673	239	31	0	93	2	0	0	16	
VP1	674–937	264	37	1	119	2	1	1	9	
2A	938–946	9	1	0	4	1	1	1	1	
2B	947–1074	128	17	0	57	1	0	0	7	
2C	1075–1396	322	50	0	158	1	0	0	14	
3A	1397–1486	90	9	0	27	2	0	0	9	
3B	1487–1509	23	3	0	9	0	0	0	1	
3C	1510–1719	210	30	1	74	1	1	1	6	
3D	1720–2181	462	38	0	67	4	1	1	18	
Total		2181	297	2	742	14	4	4	92	

*FUBAR and MEME report only sites under positive selection.

SLAC, single-likelihood ancestorcounting; FEL, fixed effects likelihood; FUBAR, fast, unconstrained Bayesian approximation for inferring selection; MEME, mixed effects model of evolution.

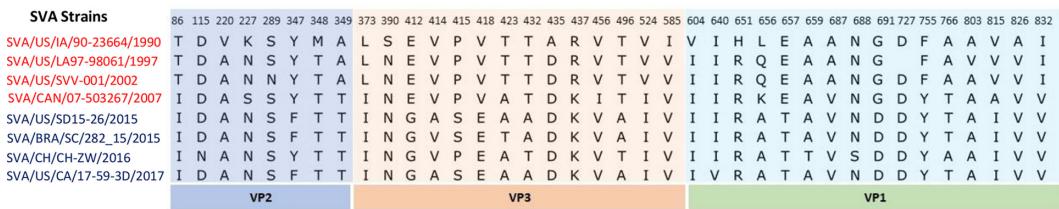


Fig. 3. Differences in amino acid (aa) sequence between historical and contemporary SVA strains. Four historical and four contemporary SVA strains were selected for protein structure analysis. Amino acid alignment was performed using MUSCLE. The figure shows the aa positions within the SVA capsid proteins where aa substitutions were noted.

methods (SLAC, FUBAR, FEL and MEME) used in our study (Table 6). Most importantly, two of these residues (735 and 941) were also detected under strong positive selection using the more conservative SLAC method (Table 5).

SVA capsid protein structure analysis

Four historical and four contemporary isolates were selected for structural analysis of the region encoding the capsid proteins. Sequence alignment highlighting the aa differences within VP2–VP1 proteins of these isolates is shown in Fig. 3. The SVA capsid is formed by four VPs, including VP4 (80–150), VP2 (151–434), VP3 (435–673) and VP1 (674–937). Among these proteins no aa changes were observed in VP4 protein between the selected strains. However, differences in 8, 14 and 16 aa positions were observed in VP2, VP3 and VP1 proteins, respectively.

These aa sequences were modelled on the SWISS-MODEL homology modelling server using the SVA strain SVV-001 capsid structure (Protein Data Bank: 3cji) as template. Although the basic alpha-helix and beta sheet structures

in VPs of all six viral strains is conserved, there is a clear difference in the size of a central pocket of this polypeptides (Fig. 4). A wider gap was observed in this region of the VPs of historical isolates when compared to contemporary strains, which presents no gap or a much narrower gap in the central protein pocket (Fig. 4). Slight differences in the size of the VPs' central pocket were also observed among historical strains (Fig. 4).

A total of 38 aa changes were observed within SVA VPs' coding region of the selected viral isolates (Fig. 3). Of these, four aa positions (412, 418, 657 and 691) are clearly distinct between historical and contemporary strains (Table 7, Fig. 3). To assess whether these aa substitutions affect the electrostatic potential of the proteins, the VPs of these isolates were modelled based on their electrostatic potential. Substitutions observed in three of the four residues (E412G, E657T and G691D) resulted in changes in the electrostatic potential between historical and contemporary strains (Fig. 5). Notably, G691D and E697T substitutions lie within the VP1 protein, while the E412G change is located within the VP3 protein

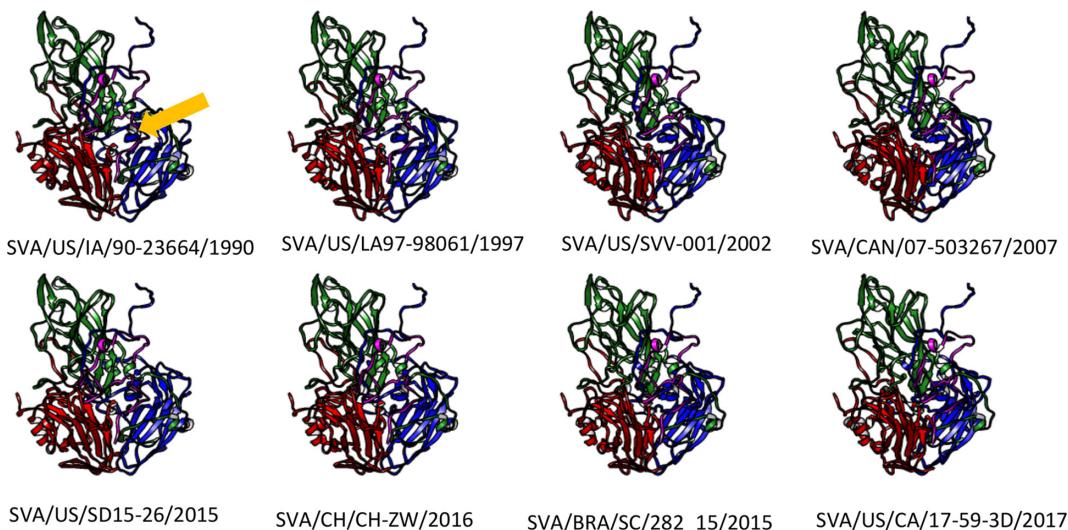


Fig. 4. SVA capsid ribbon structure. The secondary structure of VPs the SVA capsid coding polypeptides was predicted using SWISS-MODEL. The protein structures were visualized using Chimera. The colour of the ribbon represents different capsid proteins: (a) green, VP1; (b) red, -VP2; (c) blue, VP3; (d) magenta, VP4. The central pocket is indicated by the yellow arrow.

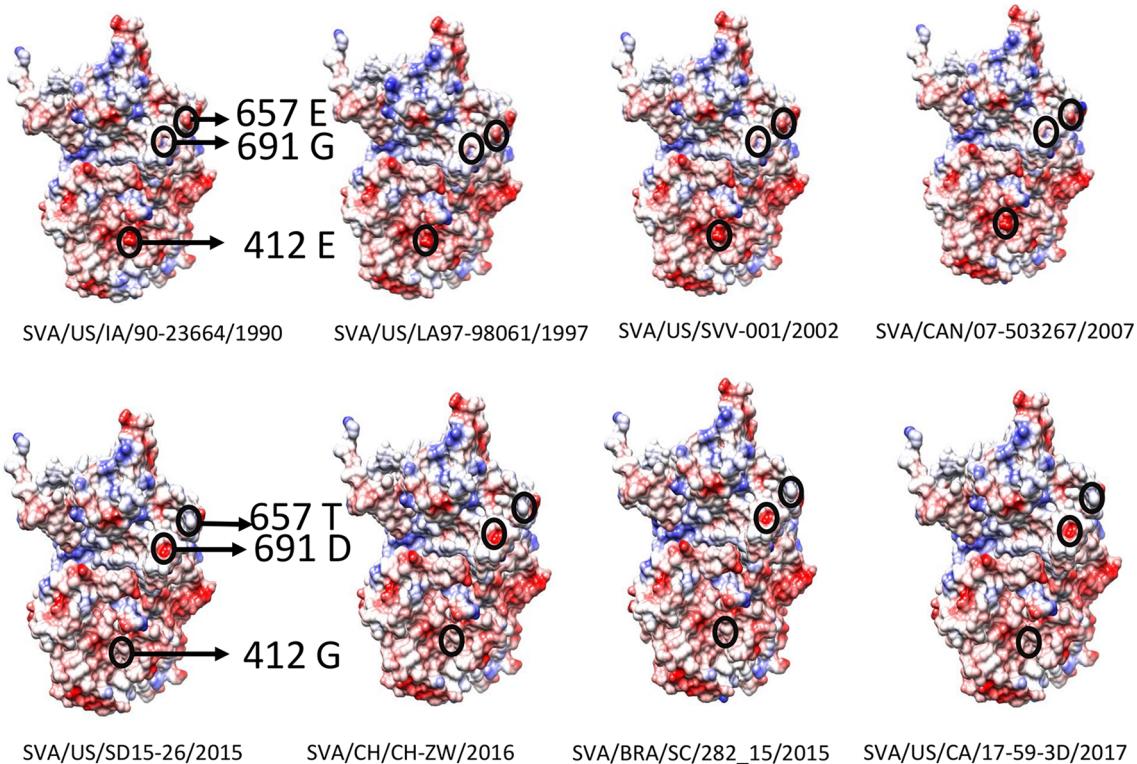


Fig. 5. Electrostatic potential on the surface of SVA. SVA VPs structures were modelled on the basis of electrostatic potential. Electrostatic potential was calculated using Coulomb's method in Chimera. Red colour on the surface represents negative potential, white colour represents neutral and blue colour represents positive potential. Positions with aa substitutions are demarcated by black circles.

sequence. Together, these results suggest that the differences in aa sequences in the SVA VPs may have led to the changes in the secondary structure and the surface electrostatic potential of the capsid proteins of the virus.

DISCUSSION

In the present study, we obtained the complete genome sequences of several historical and contemporary SVA strains

Table 7. SVA VP amino acid (aa) sequence comparison between historical and contemporary isolates

aa positions	Historical	Contemporary
412	E	G
418	V	E
657	E	T
691	G	D
Sequences used in analysis	SVA/US/IA/90-23664/1990 SVA/USA/LA97-98061/1997 SVA/US/SVV-001/2002 SVA/Canada/07-503267/2007	SVA/US/SD15-26/2015 SVA/BRA/SC/282_15/2015 SVA/China/CH-ZW/2016 USA/CA/Seq-III-5/2017

*Positions with consistent aa changes between the four historical and the four contemporary isolates.

detected and/or isolated between 1988 and 2017 in different geographical areas of the world and conducted a large-scale genetic and evolutionary analysis of SVA and . Our results show that most SVA strains form phylogenetic clades based on their geographical origin, with the exception of a few Chinese strains and one Colombian SVA strain, which clustered with contemporary USA strains. These observations indicate a close phylogenetic relationship between the Chinese and the Colombian strains with contemporary American SVA strains, suggesting a potential transfer of these strains between the distinct geographical regions. Currently, the origin or the initial source of the SVA outbreak that is affecting major swine producing countries around the world is unknown. The phylogenetic relationship between the sequences included in the present study suggest that the historical strain SVA/US/SC05-363649/2004 (originally isolated from a feral pig; Table 1) may be the ancestor of the contemporary strains currently circulating in the swine population. The scarcity of sequences obtained between 2004 and 2014, however, complicates any definitive conclusions about the origin of contemporary SVA strains.

The analysis performed in this study revealed that SVA is undergoing selection. Interestingly, our analysis shows that negative or purifying selection appears to be the main force at play on the SVA genome. Negative selection on viral genomes may lead to the removal of less adapted virus variants, thereby

maintaining variant viruses with a higher fitness within the susceptible host population [37]. Positive selection on structural genes, on the other hand, is known to drive virus evolution and diversity, as these proteins play critical roles in receptor binding and are the main targets of host immune responses [38, 39]. The results presented here show that site 735 located within the VP1 protein coding region is under strong positive selection (Table 6). Interestingly, at least three of the inference methods used in our study (FUBAR, FEL and MEME) also detected strong positive/diversifying selection in three non-structural proteins, including 2A (site 941), 3C (site 1547) and 3D (site 1850). These observations suggest that while negative/purifying selection seems to be the main force driving SVA diversity, positive selection may also contribute to the molecular evolution of SVA. These selection pressures could potentially have contributed to the recent emergence of the virus and perhaps the increased pathogenicity of contemporary SVA strains [7].

Marked differences in the SVA VP aa sequences between historical and contemporary SVA strains were observed in our analysis, indicating continuous viral evolution. The mutations observed in these proteins occurred mainly in VP1, VP2 and VP3 proteins. These proteins form the outer capsid of the virus and contain major antigenic sites [40]. VP1 also plays an important role in receptor binding [41], which keep this protein under strong immunological pressure. The results here corroborate the findings of previous studies [42, 43] in which the highest number of non-synonymous mutations in contemporary SVA strains from China were observed in VP3 and VP1. Additionally, similar to our findings, the VP2 of Chinese SVA isolates also presented two aa substitutions when compared to historical SVA isolates, whereas no differences were found in VP4. Since VP4 is present on the inner side of the capsid [40], it is possible that this protein is not subjected to the same immunological pressure as the external capsid proteins (VP1, VP2 and VP3). Some of the aa substitutions observed in P1 result in changes of the electrostatic potential of the capsid proteins (Fig. 5). Since virus–receptor attachment is affected by the surface charge of these molecules, these mutations could potentially affect virus entry mechanisms. This observation underscores the importance of structure-based protein analysis to infer evolutionary relationships between viral strains that are not easily visualized or identified by phylogenetic analysis [44, 45].

The recent re-emergence of SVA in swine worldwide raised important questions about the factors that may have led to the rapid increase in the incidence of SVA-induced vesicular disease around the world [8, 13, 15–17, 46]. In the USA alone over 400 cases of SVA have been confirmed since its re-emergence in July 2015 [25]. Although the prevalence of SVA in the USA is unknown, early serological studies conducted in the USA revealed a high prevalence (38%) of neutralizing antibodies for SVA in pigs [3]. A more recent serological investigation reported a seroprevalence of 12.2 and 34.0% in grower-finisher pigs and sows, respectively [46]. These observations suggest a broad circulation of the virus in the swine population of the country.

This study provides important information on the phylogenetic relationship and evolution of historical and contemporary SVA strains. The evolutionary changes observed at the nucleotide and aa level may have contributed to the emergence of SVA. Further studies employing molecular clones are required, however, to define the actual contribution of natural viral evolution to the virulence and pathogenicity of contemporary SVA strains in pigs.

Funding information

This study was funded by Swine Health Information Center (SHIC) project #15-192 and in part by the intramural research program of the US Department of Agriculture, Animal and Plant Health Inspection Service.

Acknowledgements

We would like to thank the California Department of Food and Agriculture veterinarians for collecting and providing samples used in this study. We are thankful for Dr Sabrina Swenson's collaboration and coordination on historical SVA strain sequencing, and for her input to and review of the manuscript. The findings and conclusions in this publication have not been formally disseminated by the US Department of Agriculture and should not be construed to represent any agency determination or policy.

References

- Zell R, Delwart E, Gorbatenya AE, Hovi T, King AMQ et al. ICTV virus taxonomy profile: Picornaviridae. *J Gen Virol* 2017;98:2421–2422.
- Hales LM, Knowles NJ, Reddy PS, Xu L, Hay C et al. Complete genome sequence analysis of Seneca Valley virus-001, a novel oncolytic picornavirus. *J Gen Virol* 2008;89:1265–1275.
- Knowles NJ, Hales LM, Jones BH, Landgraf JG, House JA et al. Epidemiology of seneca valley virus: identification and characterization of isolates from pigs in the United States. *XIVth Meeting of the European Study Group on Molecular Biology of Picornaviruses*. Saariselka; 2006. p. G2.
- Pasma T, Davidson S, Shaw SL. Idiopathic vesicular disease in swine in Manitoba. *Can Vet J* 2008;49:84–85.
- Singh K, Corner S, Clark S, Sherba G, Fredrickson R. Seneca Valley virus and vesicular lesions in a pig with idiopathic vesicular disease. *J Vet Sci Technol* 2012;3:1–3.
- Yang M, van Bruggen R, Xu W. Generation and diagnostic application of monoclonal antibodies against Seneca Valley virus. *J Vet Diagn Invest* 2012;24:42–50.
- Fernandes MHV, Maggioli MF, Joshi LR, Clement T, Faccin TC et al. Pathogenicity and cross-reactive immune responses of a historical and a contemporary Senecavirus a strains in pigs. *Virology* 2018;522:147–157.
- Leme RA, Zotti E, Alcântara BK, Oliveira MV, Freitas LA et al. *Senecavirus A*: an emerging vesicular infection in Brazilian pig herds. *Transbound Emerg Dis* 2015;62:603–611.
- Vannucci FA, Linhares DCL, Barcellos DESN, Lam HC, Collins J et al. Identification and complete genome of Seneca Valley virus in vesicular fluid and sera of pigs affected with idiopathic vesicular disease, Brazil. *Transbound Emerg Dis* 2015;62:589–593.
- Wu Q, Zhao X, Bai Y, Sun B, Xie Q et al. The first identification and complete genome of Senecavirus a affecting pig with idiopathic vesicular disease in China. *Transbound Emerg Dis*.
- Bracht AJ, O'Hearn ES, Fabian AW, Barrette RW, Sayed A. Real-Time reverse transcription PCR assay for detection of Senecavirus A in swine vesicular diagnostic specimens. *PLoS One* 2016;11:e0146211.
- Canning P, Canon A, Bates JL, Gerardy K, Linhares DCL et al. Neonatal mortality, vesicular lesions and lameness associated with Senecavirus A in a U.S. sow farm. *Transbound Emerg Dis* 2016;63:373–.

13. Joshi LR, Mohr KA, Clement T, Hain KS, Myers B et al. Detection of the emerging picornavirus Senecavirus A in pigs, mice, and houseflies. *J Clin Microbiol* 2016;54:1536–1545.
14. Hause BM, Myers O, Duff J, Hesse RA. Senecavirus A in pigs, United States, 2015. *Emerg Infect Dis* 2016;22:1323–1325.
15. Sun D, Vannucci F, Knutson TP, Corzo C, Marthaler DG. Emergence and whole-genome sequence of Senecavirus A in Colombia. *Transbound Emerg Dis* 2017;64:1346–1349.
16. Saeng-chuto K, Rodtian P, Temeeyasen G, Wegner M, Nilubol D. The first detection of Senecavirus A in pigs in Thailand, 2016. *Transbound Emerg Dis* 2018;65:285–288.
17. Arzt J, Bertram MR, Vu LT, Pauszek SJ, Hartwig EJ et al. First detection and genome sequence of Senecavirus A in Vietnam. *Microbiol Resour Announc* 2019;8:e01247–18.
18. Montiel N, Buckley A, Guo B, Kulshreshtha V, VanGeelen A et al. Vesicular disease in 9-Week-Old pigs experimentally infected with Senecavirus a. *Emerg Infect Dis* 2016;22:1246–1248.
19. Joshi LR, Fernandes MHV, Clement T, Lawson S, Pillatzki A et al. Pathogenesis of senecavirus a infection in finishing pigs. *J Gen Virol* 2016;97:3267–3279.
20. Morse SS. Factors in the emergence of infectious diseases. *Emerg Infect Dis* 1995;1:7–15.
21. Licita B, Duhamel G, Whittaker G. Canine enteric coronaviruses: emerging viral pathogens with distinct recombinant spike proteins. *Viruses* 2014;6:3363–3376.
22. Jarvis MC, Lam HC, Zhang Y, Wang L, Hesse RA et al. Genomic and evolutionary inferences between American and global strains of porcine epidemic diarrhea virus. *Prev Vet Med* 2016;123:175–184.
23. Oka T, Saif LJ, Marthaler D, Esseili MA, Meulia T et al. Cell culture isolation and sequence analysis of genetically diverse us porcine epidemic diarrhea virus strains including a novel strain with a large deletion in the spike gene. *Vet Microbiol* 2014;173:258–269.
24. Vlasova AN, Marthaler D, Wang Q, Culhane MR, Rossow KD et al. Distinct characteristics and complex evolution of PEDV strains, North America, may 2013–February 2014. *Emerg Infect Dis* 2014;20:1620–1628.
25. SHMP. Swine health monitoring project report 2018.
26. Boisvert S, Raymond F, Godzardis Élénie, Laviolette F, Corbeil J et al. Ray meta: scalable de novo metagenome assembly and profiling. *Genome Biol* 2012;13:R122.
27. Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 2008;18:821–829.
28. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 2005;437:376–380.
29. Edgar RC. Muscle: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004;32:1792–1797.
30. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* 2018;35:1547–1549.
31. Metsalu T, Vilo J. ClustVis: a web tool for visualizing clustering of multivariate data using principal component analysis and heatmap. *Nucleic Acids Res* 2015;43:W566–W570.
32. Spielman SJ, Weaver S, Shank SD, Magalis BR, Li M et al. Evolution of viral genomes: interplay between selection, recombination, and other forces. *Methods Mol Biol* 1910;2019:427–468.
33. Weaver S, Shank SD, Spielman SJ, Li M, Muse SV et al. Data-monkey 2.0: a modern web application for characterizing selective and other evolutionary processes. *Mol Biol Evol* 2018;35:773–777.
34. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 2009;25:1189–1191.
35. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* 2018;46:W296–W303.
36. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM et al. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* 2004;25:1605–1612.
37. Loewe L. Negative selection. *Nat Educ* 2008;1:59.
38. Lin J-Y, Chen T-C, Weng K-F, Chang S-C, Chen L-L et al. Viral and host proteins involved in picornavirus life cycle. *J Biomed Sci* 2009;16:103.
39. Maggioli MF, Lawson S, de Lima M, Joshi LR, Faccin TC et al. Adaptive immune responses following senecavirus a infection in pigs. *J Virol*;92.
40. Belnap DM, McDermott BM, Filman DJ, Cheng N, Trus BL et al. Three-dimensional structure of poliovirus receptor bound to poliovirus. *Proc Natl Acad Sci USA* 2000;97:73–78.
41. Ashkani J, Rees DJG. The critical role of VP1 in forming the necessary cavities for receptor-mediated entry of FMDV to the host cell. *Sci Rep* 2016;6:27140.
42. Xu W, Hole K, Goolia M, Pickering B, Salo T et al. Genome wide analysis of the evolution of Senecavirus A from swine clinical material and assembly yard environmental samples. *PLoS One* 2017;12:e0176964–17.
43. Zhao X, Wu Q, Bai Y, Chen G, Zhou L et al. Phylogenetic and genome analysis of seven senecavirus a isolates in China. *Transbound Emerg Dis* 2017;64:2075–2082.
44. Ströh LJ, Neu U, Blaum BS, Buch MHC, Garcea RL et al. Structure analysis of the major capsid proteins of human polyomaviruses 6 and 7 reveals an obstructed sialic acid binding site. *J Virol* 2014;88:10831–10839.
45. Ravantti J, Bamford D, Stuart DI. Automatic comparison and classification of protein structures. *J Struct Biol* 2013;183:47–56.
46. Houston E, Giménez-Lirola LG, Magtoto R, Mora-Díaz JC, Baum D et al. Seroprevalence of senecavirus A in sows and grower-finisher pigs in major swine producing-states in the United States. *Prev Vet Med*.

Five reasons to publish your next article with a Microbiology Society journal

1. The Microbiology Society is a not-for-profit organization.
2. We offer fast and rigorous peer review – average time to first decision is 4–6 weeks.
3. Our journals have a global readership with subscriptions held in research institutions around the world.
4. 80% of our authors rate our submission process as ‘excellent’ or ‘very good’.
5. Your article will be published on an interactive journal platform with advanced metrics.

Find out more and submit your article at microbiologyresearch.org.