Revisiting the genetic diversity of classical swine fever virus: A proposal for new genotyping and sub-genotyping schemes of classification.

Journal:	*Transboundary and Emerging Diseases*
Manuscript ID	TBED-RC-106-18
Manuscript Type:	Rapid Communication
Date Submitted by the Author:	19-Feb-2018
Complete List of Authors:	Rios, Liliam; Dalhousie University Faculty of Medicine, Reiman Cancer Research Laboratory Núñez, Jose ; Institut de Recerca i Tecnologia Agroalimentaries, Centre de Recerca en Sanitat Animal (CReSA, IRTA- UAB) Diaz de Arce, Heidy; Hospital Italiano de Buenos Aires, Juan D. Perón 4190 Ganges, Llilianne; OIE Reference Laboratory for Classical Swine Fever Perez, Lester J; Dalhousie University Faculty of Medicine, DMNB
Subject Area:	Virus, Disease control, Veterinary epidemiology

Revisiting the genetic diversity of classical swine fever virus: A proposal for new genotyping and sub-genotyping schemes of classification.

Liliam Rios[1], José I. Núñez[2], Heidy Díaz de Arce[3], Llilianne Ganges[4], Lester J. Pérez[5]*

Short title: New genetic classification for classical swine fever virus

[1]*University of New Brunswick, Saint John, New Brunswick, E2L4L5, Canada.*

[2]*IRTA-CReSA. Centre de Recerca en Sanitat Animal, Barcelona, 08193, Spain.*

[3]*Hospital Italiano de Buenos Aires, Juan D. Perón 4190, C1181ACH Buenos Aires, Argentina*

[4]*OIE Reference Laboratory for Classical Swine Fever, IRTA-CReSA, Barcelona, Spain.*

[6]*Dalhousie University, Dalhousie Medicine New Brunswick, Saint John, New Brunswick, E2L4L5, Canada.*

*Corresponding author: Tel.:+1 (506)-636-6977 fax:+1 (506)-636-6258

E-mail address: lester.perez@dal.ca (Pérez, L.J.)

Abstract

Classical swine fever (CSF) is a highly contagious febrile viral disease caused by CSF virus (CSFV) and it is considered one of the most important infectious diseases that affect domestic pigs and wild boar. Previous molecular epidemiology studies have revealed that the diversity of CSFV comprises three main genotypes and different subgenotypes defined by using a reliable cut-off to accurately classify CSFV at genotype and subgenotype levels. However, a growing number of CSFV both complete genome and full E2 gene sequences have been submitted to GenBank (more than 500 sequences are currently available, revised on December 1st, 2017). Therefore, the aim of the current study was to revisit the taxonomy of CSFV at genotype and subgenotype levels, to unify nomenclature and to provide an update to the classification of CSFV. We propose here a new genotyping scheme with five well defined CSFV-genotypes (CSFV-Genotype1-5) and 14 subgenotypes (seven for each of the CSFV-Genotype1 and CSFV-Genotype2). The findings showed in the current study are relevant for molecular epidemiology approaches and will help to better understand the genetic diversity and spreading of CSFV at a global scale. The update in the classification of CSFV will allow the scientific community to establish more accurately the links among different outbreaks of the disease.

Introduction

Classical swine fever (CSF) is a highly contagious febrile viral disease, considered one of the most important infectious diseases that affect domestic pigs and wild boar (*Sus scrofa*) (Blome et al., 2017). Because of its huge economic impact, the disease is notifiable to the World Organisation of Animal Health (OIE) (Moennig et al., 2013). Even though CSF has been successfully eradicated from some countries including Canada, United States, Australia and New Zealand; it remains to have a severe impact on Asia, Eastern Europe and most of South and Central America as well as the Caribbean (Postel et al., 2017).

CSF is caused by CSF virus (CSFV), a small-enveloped RNA virus of the genus *Pestivirus* included into the *Flaviviridae* family. The CSFV genome consists of a single plus-stranded RNA, which contains one large open reading frame (ORF) flanked by two untranslated regions (UTRs). The ORF encodes a polyprotein of approximately 3900 amino acids which is subsequently processed by cellular and viral proteases into mature proteins: four structural proteins (C, Erns, E1 and E2) and 8 non-structural proteins (Npro, P7, NS2, NS3, NS4A, NS4B, NS5A, NS5B)(Meyers and Thiel, 1996).

Previous molecular epidemiology studies have revealed that the diversity of CSFV comprises three main genotypes and different subgenotypes (Postel et al., 2013). To establish an international consensus for CSFV's classification system, a recent study assessed the reliability of the phylogenetic markers most commonly used in molecular epidemiology studies of CSFV (Rios et al., 2017). Thus, the phylogenetic marker based on full E2 gene was found to be the best phylogenetic marker, capable of reproducing the same phylogenetic and evolutionary information as the complete viral genome (Rios et al., 2017). In addition, Rios et al. (2017), using the combination of Pairwise Sequence Comparison (PASC), Sequence Demarcation Tool (SDT) analyses and pairwise distance calculation, determined a reliable cut-off to accurately classify CSFV at genotype and subgenotype levels (Rios et al., 2017). Rios et al. (2017) also investigated the evolutionary forces driving the genetic diversity and evolution of CSFV, including the conception of a structural model for E2 protein and other intensive computational analyses. Thus, collecting all these relevant data required a considerable amount of time, consequently,

the dataset of CSFV sequences used in this study included only those available until April 2016. However, a growing number of CSFV both complete genome and full E2 gene sequences have been submitted to GenBank (more than 500 sequences are currently available, revised on December 1st, 2017), most of them updated after June 2016 when the study published by Rios et al. (2017) was already accomplished. In addition, at the time that the study reported by Rios et al., (2017) was under revision two new subgenotypes (1.5 and 1.6) from CSFV isolates that circulated in Brazil were reported (Silva et al., 2017). This last finding is indicative that the genetic diversity of CSFV could be broader than it was previously reported (Postel et al., 2012, Rios et al., 2017). Hence, the current study was prompted by the increasing number of sequences of CSFV available on GenBank, and follows on the previously published proposals for the classification of CSFV based on 15% of genetic distance to differ among genotypes and 9% of genetic distance to consider new subgenotypes (Rios et al., 2017). Therefore, the aim of the current study was to revisit the taxonomy of CSFV at genotype and subgenotype levels, to unify nomenclature and to provide an update to the classification of CSFV. We propose here a new genotyping scheme with five well defined CSFV-genotypes (CSFV-Genotype1-5) and 14 subgenotypes (seven for each of the CSFV-Genotype1 and CSFV-Genotype2). The findings showed in the current study are relevant for molecular epidemiology approaches and will help to better understand the genetic diversity and spreading of CSFV at a global scale.

Material and methods

Dataset

All available CSFV sequences from both the E2 full gene and the complete genome were downloaded from GenBank on December 1st, 2017 (Supplementary material TableS1). The genome region analysed in this study included the full E2 gene as previously proposed by Rios et al. (2017). After removing poor quality and redundant sequences (Supplementary material TableS1, sequence highlighted in gray), a total of 517 sequences of the E2 gene were included in the study (Supplementary material TableS1).

Multiple alignment and Model selection

All sequences were aligned using the MUltiple Sequence Comparison by Log- Expectation (MUSCLE) software freely available at: https://www.ebi.ac.uk/Tools/msa/muscle. The software jModelTest 2.0 was used to estimate the best-fit model using the Akaike and Bayesian information criteria (AIC and BIC) (Darriba et al., 2012). The best-fit model selected was used for phylogenetic analysis and genetic distance calculation.

Phylogenetic analysis

Phylogenetic analyses were performed following the methodology suggested by Rios et al. (2017), briefly: searches for recombinant sequences and crossover regions were performed to remove the sequences with a possible recombinant event, using Geneconv, RDP, MaxChi, Chimera, BootScan, SiScan, 3Seq and LARD, all implemented in RDP3 Beta 4.1 (Martin and Rybicki, 2000). Phylogenetic relationships of the CSFV strains using the E2 complete gene marker were analyzed using a Maximum Likelihood (ML) approach. The sequences JX428945/NC_018713 belonging to the *Pestivirus Aydin* were used as outgroup.

Calculation of pairwise nucleotide p-distances and PAirwise Sequence Comparison (PASC) analysis.

Pairwise nucleotide p-distances were calculated using MEGA7 (Kumar et al., 2016). Different matrix of nucleotide divergence between groups were generated using an alpha-value=0.66 and 1000 bootstrap replicates to estimate variance. To confirm the reliability of the cut-off previously defined by Rios et al. (2017) for the different lineages of CSFV, a PASC analysis was performed. Thus, all 517 unique CSFV sequences for E2 full gene were submitted to the web tool DIVEIN (Deng et al., 2010) and a histogram based on computing the divergence/diversity among and within CSFV lineages was accomplished.

Results and Discussion

The PASC analyses, based on the E2 gene from the 517 E2 full gene sequences, displayed a multimodal curve (Fig. 1A), similar to the results obtained by Rios et al. (2017). Threshold values of 91% and 86% of identity allow to separate all the subgenotypes and genotypes of CSFV, respectively (Fig. 1A). Thus, the cut-off values were consistent with those previously obtained by Rios et al. (2017). It's

important to denote that the cut-off values were not changed by the effect of the number of sequences employed. For other viral agents such as porcine circovirus type 2 (PCV2), when the taxonomy was revisited due to conflicting results obtained by different research groups (Franzo et al., 2015), it was evidenced that the increase in the number of sequences analysed yielded incompatible cut-off values (Franzo et al., 2015) compared to those previously established (Grau-Roma et al., 2008). Thus, the fact that the cut-off values to define CSFV genotypes and subgenotypes were not altered despite the 5-fold increase of the total of sequences analysed, ensures the accurate classification for this viral agent.

ML tree, based on the complete E2 gene from 517 CSFV sequences, identified five main lineages (CSFV genotypes 1-5) and different sublineages (Fig. 1B). In addition, the genetic divergence among the different proposed CSFV-Genotypes ranged between 15.6%-19.1% (Fig. 1C). Thus, beside the historically recognized CSFV-Genotypes 1-3 (Postel et al., 2012), two new genotypes are proposed in this study. One of the new proposed genotypes was formed by the British CSFV strain "Congenital Tremor" isolated in 1964 (sequence ID:JQ411575) (Vilcek et al., 1996). This strain was found to be one of the most distinct strains in a phylogenetic study performed in 1996 (Vilcek et al., 1996), and it was later misplaced as outgroup, in molecular epidemiology studies of CSFV (Postel et al., 2012, Postel et al., 2017). In the current study, CSFV strain "Congenital Tremor" showed a genetic divergence, compared to the remaining CSFV-genotypes, between 15.7%-17.4% (Fig. 1C). It is also important to denote that in phylogenetic analysis at species level, where rooted trees are analyzed, those sequences from species closer to the specie in study must be used as outgroup (Perez et al., 2011, Martinez et al., 2012, Barrera et al., 2017, Holland et al., 2003). Since CSFV strain "Congenital Tremor" (sequence ID: JQ411575) showed less that 20% of genetic divergence when compared to the remaining CSFV genotypes, this strain is a CSFV member and therefore, its use as outgroup is not appropriate when phylogenetic analyses are conducted at species level. However, the branch formed by this strain is divergent enough from CSFV-genotype 2 to be considered a new genotype. The fact that the number of full sequences of CSFV E2 gene has increased considerably in the GenBank Database could have helped resolve this new topology

showing CSFV strain "Congenital Tremor" as a divergent lineage. Therefore, we propose this lineage to be designated as a new CSFV-genotype (CSFV-genotype 4) (Fig. 1B and C).

In addition to the CSFV strain "Congenital Tremor", the CSFV strains JJ9811 and YI9908 isolated from Korea during 1998 and 1999, respectively, formed a statistically supported independent lineage (Fig.1B). This lineage showed a genetic divergence compared to the remaining CSFV-genotypes that ranged between 15.6%-19.1% (Fig. 1C). The CSFV strains JJ9811 and YI9908 have been previously classified as genotype 3, subgenotype 3.2 (Lim et al., 2016). However, this previous classification was based on the phylogenetic analysis using the segment of E2 comprising 190 nt (E2-190 marker) (Lowings et al., 1996). In a previous report, it was shown that the phylogenetic marker E2-190 (Lowings et al., 1996) was associated with loss of phylogenetic information, besides, this marker was unable to reproduce the same topologies as the complete genome of CSFV or the E2-complete gene marker (Rios et al., 2017). A notable example showing the misclassification leaded by the use of the E2-190 marker is the case of the Cuban CSFV isolates, which were historically classified as subgenotype 1.2 (Diaz de Arce et al., 1999, de Arce et al., 2005, Perez et al., 2012). However, these isolates were re-classified as subgenotype 1.4 when the phylogenetic analysis was accomplished using the E2-complete gene marker showed a genetic segregation between 9.8–15.8% to sequences of subgenotype 1.2 (Postel et al., 2013). Thus, the use of the marker E2-190 (Lowings et al., 1996), could have lead to a misclassification of the CSFV strains JJ9811 and YI9908 into the subgenotype 3.2 (Lim et al., 2016). It is also relevant to consider that the lineage formed by the CSFV strains JJ9811 and YI9908 showed a genetic divergence of 16.6% with the CSFV strains belonging to the genotype 3 (Fig. 1B and C). Thus, both analyses (the topology and the genetic divergence of the lineage formed by the CSFV strains JJ9811 and YI9908) support the divergence of this new lineage and we propose to designate it as a new CSFV-genotype (CSFV-genotype 5) (Fig. 1 and Supplementary Material Fig.S1).

At intra-genotype level, a new scheme of the genetic diversity was also revealed (Fig. 2, Fig. 3 and Table 1). In the case of CSFV-genotype 1, seven subgenotypes were found, consisting of the four (1.1-1.4) previously recognized (Postel et al., 2013, Rios et al., 2017), the two new subgenotypes (1.5-1.6)

recently described circulating in Brazil (Silva et al., 2017), and a new subgenotype reported for the first time in the current study designated as CSFV- subgenotype 1.7 (Fig.2 and Table 1). The new subgenotype 1.7 was strongly supported by bootstrap values and the genetic divergence showed in comparison with the remaining subgenotypes (Fig.2 and Table 1). Thus, the CSFV subgenotype 1.7 showed a genetic divergence ranged from 9.2% to 13.1% (Table 1).

The new 1.7 subgenotype comprised 20 sequences (KX586754-KX586772, and KX586774), all from viral strains circulating in Ecuador during 2012 to 2015. Moreover, another sequence included in this study (KX586773) of a strain circulating in Ecuador in March of 2000 clustered together with the CSFV subgenotype 1.1. Hence, a switch from the subgenotype 1.1 to subgenotype 1.7 occurred in Ecuador during 2000 to 2012, caused by an event not yet reported.

Even though a phylodynamic study focused on Ecuadorian CSFV strains has been recently reported (Garrido Haro et al., 2018), no epidemiological information about this new cluster was discussed. Likewise, it's important to highlight that the phylodynamic analysis in this study was very limited, since only the B/C domain of the E2 gene (190 nt) was considered. Although Garrido-Haro et al. (2018) sequenced the complete gene E2 from Ecuadorian strains, these authors included in their analysis Peruvian strains (Genbank Acc. No. HM070972, HM070975, HM070976, HM070977, HM070982 and HM070988 (See Figure 1 in Garrido-Haro et al. (2018)), which are partial E2 sequences, framing the B/C domain region. Therefore, all the inferences performed in the different analyses in Garrido-Haro et al. (2018) (time for the most recent common ancestor (tMRCA), Bayesian Skyline Plot (BSP) and evolutionary rates) were restricted to the B/C domain region, which has been recently described to bias the results for phylogenetic and phylodynamic approaches in CSFV (Rios et al., 2017). Therefore, further studies will be required to get a better understanding about the events supporting the switch of the subgenotype 1.1 to 1.7 in Ecuador. We also remark that the new cluster containing the CSFV-strains circulating in Ecuador was previously classified as subgenotype 1.6 together with strains that circulated in Brazil and Peru during the years 2008 and 2009 (Garrido Haro et al., 2018). However, analysing in detail the results reported by Garrido-Haro et al. (2018), it can be noted that the Ecuadorian strains formed a

segregated cluster from the Brazilian and Peruvian strains. In addition, these authors didn't perform a genetic divergence analysis leading to a misclassification of this new group of strains circulating in Ecuador.

The results obtained in the current study from the genetic divergence evaluation among all the statistically supported lineages (Fig. 2), showed that no additional subgenotypes within the CSFV-genotype 1 were supported (Supplementary material TableS2).

Regarding CSFV-genotype 2, besides the three previously reported subgenotypes (Postel et al., 2012, Rios et al., 2017), another four subgenotypes were identified (Fig. 3, Table 1) from the 19 sublineages assessed (Supplementary material TableS3). The lineage formed by sequences of CSFV strains that circulated in India during 2012-2013 formed a statistically supported cluster (Fig. 3 and Supplementary material TableS1). These sequences showed a genetic divergence of 11.9% with the subgenotype 2.1, where the strains were previously located (Ahuja et al., 2015), and a genetic divergence compared to the remaining subgenotypes into CSFV-genotype 2 that ranged between 9.1% and 13.8% (Table 1). Hence, we propose this lineage to be defined as subgenotype 2.4 (Fig. 3). Likewise, the lineage consisting of CSFV strains that circulated in China during 2008 to 2013 and in Viet Nam in 2014 (Fig. 3 and Supplementary material TableS1) formed a statistically supported cluster (Fig. 3 and Supplementary material TableS1) and showed a genetic divergence of 9.1% with the subgenotype 2.1, where the strains were previously located, thus we propose to define this lineage as CSFV subgenotype 2.5 (Fig.3). It is important to highlight that both CSFV subgenotypes 2.4 and 2.5 were previously defined as subgenotypes 2.1d and 2.1c, respectively (Gong et al., 2016). Gong et al. (2016) also proposed another eight subgenotypes for a total of ten new subgenotypes all diversified from the subgenotype 2.1 (Gong et al., 2016). However, a detailed analysis accomplished in Rios et al. (2017) showed that neither the genetic divergence showed by the lineages nor the statistical values in the topology resolved were enough to support the classification of these lineages as new subgenotypes (Rios et al., 2017).

In the current study, a supported diversification from the branch that originated the subgenotype 2.1 yielded two new lineages (subgenotype 2.4 and subgenotype 2.5) (Fig. 3). The effect of the number of

taxa on the support of the nodes in the phylogenetic tree has been previously reported as key element to be considered in the phylogenetic tree reconstruction (Simion et al., 2017, Philippe et al., 2011). In addition, it is well known that including more taxa allows a better detection of multiple substitutions, decreasing the amount of non-phylogenetic signal while preserving phylogenetic signal, which can be translated in a better resolution of the topology (Philippe et al., 2011). Therefore, the fact that the number of sequences used in the current study has been increased approximately in 3-fold in comparison with the sequences used in Gong et al. (2016) (517 sequences vs 160 sequences used in Gong et al. (2016)) clearly improved the support of the node in the topologies obtained. Likewise, it has been previously demonstrated, taking into account the postulates of the neutral theory of evolution, that genetic diversity increases with a larger effective population size (Hague and Routman, 2016). Hence, the taxa number increase could have been determinant in obtaining the genetic distance values for the subgenotypes 2.4 and 2.5 (Table 1 and supplementary material TableS3). Despite these novel results, it is also relevant to denote that, non-additional subgenotypes were found from the diversification of the subgenotype 2.1 (Figure 3 and supplementary material TableS3). Thus, in this regard, the present work reinforces the results previously described by Rios et al. (2017) which evidenced that subgenotypes 2.1 a,b,g,h,i and j defined by Gong et al. (2016) are not distinct enough to be regarded as new subgenotypes.

In addition to the two new subgenotypes (2.4 and 2.5) which emerged from the same ancestor than subgenotype 2.1, another two subgenotypes (2.6 and 2.7) were defined into the CSFV genotype 2 (Fig. 3 and Table 1). These two proposed subgenotypes, emerged from the same ancestor than subgenotype 2.2 (Fig. 3). The new proposed subgenotype 2.6 consisted of strains circulating in Viet Nam in 2014 (Supplementary information TableS1) with an ancestral CSFV strain that circulated in Italy in 1998 (Supplementary information TableS1). A recent report by Hung et al. (2017) described the same topological reconstruction for these Vietnamese strains emerging from the ancestral CSFV-strain *CSF0573-Parma* circulating in Italy in 1998 (Hung, 2017). However, since Hung et al. (2017) only employed a total of 29 sequences of CSFV, it was not possible to obtain a genetic divergence of this cluster with an accurate resolution compared to subgenotype 2.2 (Hung, 2017). In the current study, the

independent segregation of this lineage (defined as 2.6) was statistically supported by a 100% of bootstrap value. In addition, the new proposed subgenotype 2.6 showed a genetic distance of 13.0% compared with subgenotype 2.2 (Table 1) and the genetic divergence compared with the remaining subgenotypes was ranged between 9.3% and 13.0%. Therefore, based on all the results obtained, we consider this lineage as a new subgenotype designated as 2.6 (Fig.3 and Table 1).

Surprisingly, the strain Bergen isolated in Netherlands in 1977 formed an independent lineage, statistically supported with 85% of bootstrap value (Fig. 3) and showed a genetic divergence of 10.8% compared to subgenotype 2.2 (Table 1), where it was previously included (Postel et al., 2012). Relevant aspects need to be clarified regarding this result. First, the lineage proposed as the new subgenotype 2.7 was composed by two non-identical sequences from the CSFV Bergen strain. This strain has four sequences on Genbank database: one sequence for the complete E2 gene (JQ411587), another sequence for the complete genome (KJ619377) and two sequences for the NS5B region (U30720 and AF182909). However, both E2 sequences for this viral strain (JQ411587 and the E2 sequence extracted from the complete genome (KJ619377)) are not identical, therefore, they were analyzed independently. Second, in a previous report, it was highlighted that CSFV-strain Netherlands/JQ411587 "Bergen" (CSF0906) partially displayed a higher genetic similarity to some genotype 2.1 isolates than to different 2.2 isolates, disturbing the segregation of 2.1 and 2.2 isolates (Postel et al., 2012). However, since Postel et al. (2012) only used 33 CSFV sequences, they also faced the trouble of acquiring both an accurate resolution and a proper genetic divergence of this new cluster compared the subgenotype 2.2.

On the other hand, for the remining genotypes (CSFV-genotype 3, CSFV-genotype 4 and CSFV-genotype 5) a diversification in additional subgenotypes was not detected (Supplementary material Fig. S2, Table S4 and S5).

In the current study, a new classification scheme for CSFV is proposed. The increased number of CSFV sequences available on GenBank database, especially of the full E2 gene, facilitated obtaining a better resolution for the topology of CSFV-tree. In addition, the establishment of a reliable cut-off value by Rios et al. (2017) made possible to accurately define genotypes and subgenotypes for CSFV. Similar

approaches have been accomplished for other viral agents. Thus, the growing number of sequences for infectious bursal disease virus (IBDV) and the use of phylogenetic methodologies have enabled a new classification of this viral agent into seven genogroups, updating the previous classification which only recognized three groups (Michel and Jackwood, 2017). Likewise, for porcine circovirus type 2 (PCV2), a new genotype has been added to the previous taxonomic classification, after the analysis of approximately 3300 new sequences of the complete genome of this virus (Franzo et al., 2016), which were submitted to GenBank database after the first taxonomical classification for PCV2 had been accomplished (Grau-Roma et al., 2008). The current study also highlights the importance of submitting non-redundant sequences for CSFV. Although a new classification scheme is provided here, it is relevant to denote that some phylogenetic clades have better representation of viral isolates than others. Thus, we encourage the different research groups to increase their molecular epidemiology studies regarding CSFV, which can stimulate the acquisition of new representative CSFV sequences. Finally, the results presented here will facilitate future analyses focused on elucidating evolutionary relationships among different CSFV isolates. The update in the classification of CSFV will allow the scientific community to establish more accurately the links among different outbreaks of the disease.

Conflict of interest

The authors declare no conflict of interest.

Author contribution

L.J.P. designed the research; L.J.P. and L.R performed the phylogenetic analysis; L.J.P., L.R., H.D-A., L.G., and J.I.N. analyzed and interpreted the data; L.J.P. and L.R. wrote the paper; L.G., J.I.N. and H.D-A edit the paper and provided intellectual inputs. All the authors read and approved the final version of the manuscript.

Guarantor Statement

Dr. Lester J. Pérez, is the guarantor of this work, had full access to all the data, and takes full responsibility for the integrity of data and the accuracy of data analysis.

References

Ahuja, A., U. Bhattacharjee, A. K. Chakraborty, A. Karam, S. Ghatak, K. Puro, S. Das, I. Shakuntala, N. Srivastava, S. V. Ngachan and A. Sen, 2015: Complete genome sequence of classical Swine Fever virus subgenogroup 2.1 from assam, India. ***Genome announcements,*** **3.**

Barrera, M., A. Garrido-Haro, M. S. Vaca, D. Granda, A. Acosta-Batallas and L. J. Pérez, 2017: Tracking the Origin and Deciphering the Phylogenetic Relationship of Porcine Epidemic Diarrhea Virus in Ecuador. ***BioMed research international,*** **2017, 7.**

Blome, S., C. Staubach, J. Henke, J. Carlson and M. Beer, 2017: Classical Swine Fever-An Updated Review. ***Viruses,*** **9.**

Darriba, D., G. L. Taboada, R. Doallo and D. Posada, 2012: jModelTest 2: more models, new heuristics and parallel computing. ***Nature methods,*** **9, 772.**

de Arce, H. D., L. Ganges, M. Barrera, D. Naranjo, F. Sobrino, M. T. Frias and J. I. Nunez, 2005: Origin and evolution of viruses causing classical swine fever in Cuba. ***Virus research,*** **112, 123-131.**

Deng, W., B. S. Maust, D. C. Nickle, G. H. Learn, Y. Liu, L. Heath, S. L. Kosakovsky Pond and J. I. Mullins, 2010: DIVEIN: a web server to analyze phylogenies, sequence divergence, diversity, and informative sites. ***Biotechniques,*** **48, 405-408.**

Diaz de Arce, H., J. I. Nunez, L. Ganges, M. Barreras, M. Teresa Frias and F. Sobrino, 1999: Molecular epidemiology of classical swine fever in Cuba. ***Virus research,*** **64, 61-67.**

Franzo, G., M. Cortey, A. Olvera, D. Novosel, A. M. Castro, P. Biagini, J. Segales and M. Drigo, 2015: Revisiting the taxonomical classification of Porcine Circovirus type 2 (PCV2): still a real challenge. ***Virology journal,*** **12, 131.**

Franzo, G., M. Cortey, J. Segales, J. Hughes and M. Drigo, 2016: Phylodynamic analysis of porcine circovirus type 2 reveals global waves of emerging genotypes and the circulation of recombinant forms. ***Molecular phylogenetics and evolution,*** **100, 269-280.**

Garrido Haro, A. D., M. Barrera Valle, A. Acosta and J. F. F, 2018: Phylodynamics of classical swine fever virus with emphasis on Ecuadorian strains. ***Transboundary and emerging diseases*****.**

Gong, W., J. Wu, Z. Lu, L. Zhang, S. Qin, F. Chen, Z. Peng, Q. Wang, L. Ma, A. Bai, H. Guo, J. Shi and C. Tu, 2016: Genetic diversity of subgenotype 2.1 isolates of classical swine fever virus. ***Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases,*** **41, 218-226.**

Grau-Roma, L., E. Crisci, M. Sibila, S. Lopez-Soria, M. Nofrarias, M. Cortey, L. Fraile, A. Olvera and J. Segales, 2008: A proposal on porcine circovirus type 2 (PCV2) genotype definition and their relation with postweaning multisystemic wasting syndrome (PMWS) occurrence. ***Veterinary microbiology,*** **128, 23-35.**

Hague, M. T. and E. J. Routman, 2016: Does population size affect genetic diversity? A test with sympatric lizard species. ***Heredity,*** **116, 92-98.**

Holland, B. R., D. Penny and M. D. Hendy, 2003: Outgroup misplacement and phylogenetic inaccuracy under a molecular clock--a simulation study. ***Systematic biology,*** **52, 229-238.**

Hung, N. P., Lan, N.T., Nga, B.T.T., Truong, T., Phan, L.V., 2017: Genetic Characterization of E2 gene of classical swine fever virus circulating in Nam Dinh and Hai Duong Provinces. ***Vietnam Journal of Agricultural Science,*** **15, 8.**

Kumar, S., G. Stecher and K. Tamura, 2016: MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. ***Mol Biol Evol,*** **33, 1870-1874.**

Lim, S. I., S. H. Han, H. Hyun, J. A. Lim, J. Y. Song, I. S. Cho and D. J. An, 2016: Complete Genome Sequence Analysis of Acute and Mild Strains of Classical Swine Fever Virus Subgenotype 3.2. ***Genome announcements,*** **4.**

Lowings, P., G. Ibata, J. Needham and D. Paton, 1996: Classical swine fever virus diversity and evolution. ***The Journal of general virology,*** **77 (Pt 6), 1311-1321.**

Martin, D. and E. Rybicki, 2000: RDP: detection of recombination amongst aligned sequences. ***Bioinformatics,*** **16, 562-563.**

Martinez, N., P. E. Brandao, S. P. de Souza, M. Barrera, N. Santana, H. D. de Arce and L. J. Perez, 2012: Molecular and phylogenetic analysis of bovine coronavirus based on the spike glycoprotein gene. ***Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases,*** **12, 1870-1878.**

Meyers, G. and H. J. Thiel, 1996: Molecular characterization of pestiviruses. ***Advances in virus research,*** **47, 53-118.**

Michel, L. O. and D. J. Jackwood, 2017: Classification of infectious bursal disease virus into genogroups. ***Archives of virology,*** **162, 3661-3670.**

Moennig, V., P. Becher and M. Beer, 2013: Classical swine fever. ***Developments in biologicals,*** **135, 167-174.**

Perez, L. J., H. D. de Arce, M. Cortey, P. Dominguez, M. I. Percedo, C. L. Perera, J. Tarradas, M. T. Frias, J. Segales, L. Ganges and J. I. Nunez, 2011: Phylogenetic networks to study the origin and evolution of porcine circovirus type 2 (PCV2) in Cuba. ***Veterinary microbiology,*** **151, 245-254.**

Perez, L. J., H. Diaz de Arce, C. L. Perera, R. Rosell, M. T. Frias, M. I. Percedo, J. Tarradas, P. Dominguez, J. I. Nunez and L. Ganges, 2012: Positive selection pressure on the B/C domains of the E2-gene of classical swine fever virus in endemic areas under C-strain vaccination. ***Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases,*** **12, 1405-1412.**

Philippe, H., H. Brinkmann, D. V. Lavrov, D. T. Littlewood, M. Manuel, G. Worheide and D. Baurain, 2011: Resolving difficult phylogenetic questions: why more sequences are not enough. ***PLoS biology,*** **9, e1000602.**

Postel, A., S. Austermann-Busch, A. Petrov, V. Moennig and P. Becher, 2017: Epidemiology, diagnosis and control of classical swine fever: Recent developments and future challenges. ***Transboundary and emerging diseases*****.**

Postel, A., S. Schmeiser, J. Bernau, A. Meindl-Boehmer, G. Pridotkas, Z. Dirbakova, M. Mojzis and P. Becher, 2012: Improved strategy for phylogenetic analysis of classical swine fever virus based on full-length E2 encoding sequences. ***Veterinary research,*** **43, 50.**

Postel, A., S. Schmeiser, C. L. Perera, L. J. Rodriguez, M. T. Frias-Lepoureau and P. Becher, 2013: Classical swine fever virus isolates from Cuba form a new subgenotype 1.4. ***Veterinary microbiology,*** **161, 334-338.**

Rios, L., L. Coronado, D. Naranjo-Feliciano, O. Martinez-Perez, C. L. Perera, L. Hernandez-Alvarez, H. Diaz de Arce, J. I. Nunez, L. Ganges and L. J. Perez, 2017: Deciphering the emergence, genetic diversity and evolution of classical swine fever virus. ***Scientific reports,*** **7, 17887.**

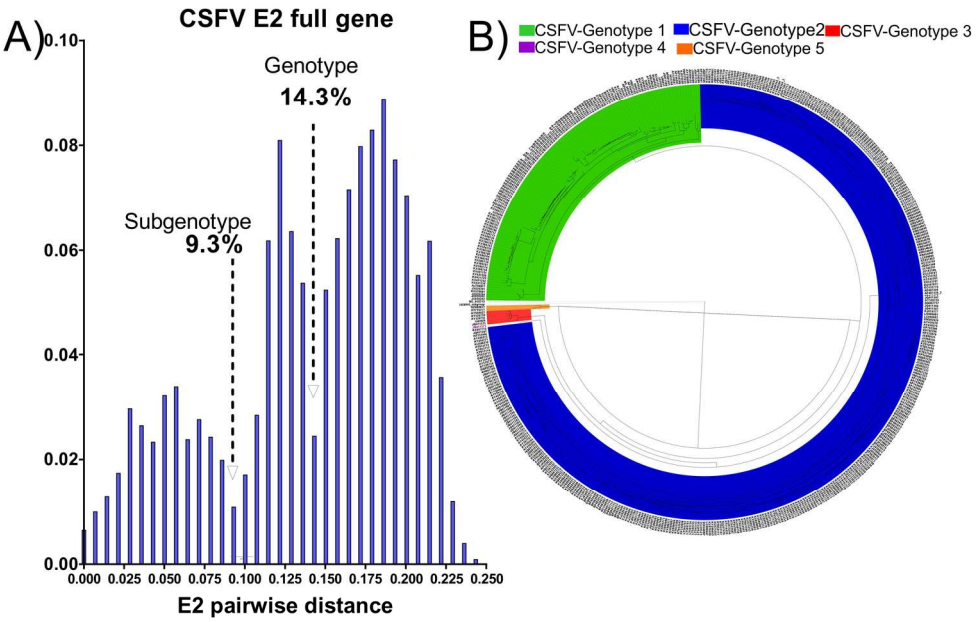
Silva, M. N., D. M. Silva, A. S. Leite, A. L. Gomes, A. C. Freitas, J. W. Pinheiro-Junior, R. S. Castro and A. L. Jesus, 2017: Identification and genetic characterization of classical swine fever virus isolates in Brazil: a new subgenotype. ***Archives of virology,*** **162, 817-822.**

Simion, P., H. Philippe, D. Baurain, M. Jager, D. J. Richter, A. Di Franco, B. Roure, N. Satoh, E. Queinnec, A. Ereskovsky, P. Lapebie, E. Corre, F. Delsuc, N. King, G. Worheide and M. Manuel, 2017: A Large and Consistent Phylogenomic Dataset Supports Sponges as the Sister Group to All Other Animals. ***Current biology : CB,*** **27, 958-967.**

Vilcek, S., T. Stadejek, A. Ballagi-Pordany, J. P. Lowings, D. J. Paton and S. Belak, 1996: Genetic variability of classical swine fever virus. ***Virus research,*** **43, 137-147.**

Table 1. Genetic distances based on full-length E2 gene sequences of CSFV for different subgenotypes**.** Values above the diagonal represent the standard error, values below the diagonal represent the p-distance values obtained using MEGA7 and 1000 bootstrap replicates

Subgenotype	1.1	1.2	1.3	1.4	1.5	1.6	1.7	2.1	2.2	2.3	2.4	2.5	2.6	2.7
1.1		*0.007*	*0.008*	*0.009*	*0.002*	*0.001*	*0.002*							
1.2	0.096		*0.006*	*0.004*	*0.008*	*0.009*	*0.005*							
1.3	0.104	0.102		*0.009*	*0.001*	*0.001*	*0.009*							
1.4	0.116	0.107	0.120		*0.011*	*0.001*	*0.001*							
1.5	0.096	0.131	0.138	0.134		*0.007*	*0.009*							
1.6	0.095	0.118	0.137	0.134	0.098		*0.009*							
1.7	0.096	0.104	0.128	0.131	0.108	0.106								
2.1									*0.008*	*0.012*	*0.011*	*0.002*	*0.001*	*0.008*
2.2								0.134		*0.008*	*0.012*	*0.012*	*0.012*	*0.012*
2.3								0.138	0.130		*0.012*	*0.013*	*0.012*	*0.012*
2.4								0.119	0.135	0.136		*0.012*	*0.012*	*0.001*
2.5								0.097	0.130	0.138	0.113		*0.004*	*0.001*
2.6								0.099	0.130	0.130	0.107	0.097		*0.001*
2.7								0.147	0.108	0.131	0.139	0.133	0.146	



C)

Genotype	1	2	3	4	5
1		0.012	0.013	0.006	0.002
2	0.176		0.014	0.014	0.013
3	0.166	0.189		0.012	0.014
4	0.157	0.174	0.165		0.015
5	0.156	0.191	0.166	0.170	

Fig. 1. Representation of frequency distribution of pairwise distance, the phylogenetic tree and genetic distance for all the five-main lineage of CSFV. A) PASC results: the cut-off values genotype (14.3%) and subgenotype (9.3%) of genetic divergence were denoted. B) All no-redundant genomes were analysed using ML-method, the GenBank IDs for all the sequence is shown, the main lineages proposed as CSFV-genotypes are denoted (CSFV-genotype 1: green, CSFV-genotype 2: blue, CSFV-genotype 3: red, CSFV-genotype 4: purple, and CSFV-genotype 5: orange (this tree has been used with representation purposes only, an additional tree showing the significance value for the nodes is shown in Supplementary material Fig.S1.C) P-distance between CSFV-genotypes, 1, 2, 3, 4 and 5: Indicates the CSFV genotypes 1, 2, 3, 4 and 5, respectively. Values above the diagonal represent the standard error, values below the diagonal represent the p-distance values obtained using MEGA7 and 1000 bootstrap replicates.

177x155mm (300 x 300 DPI)

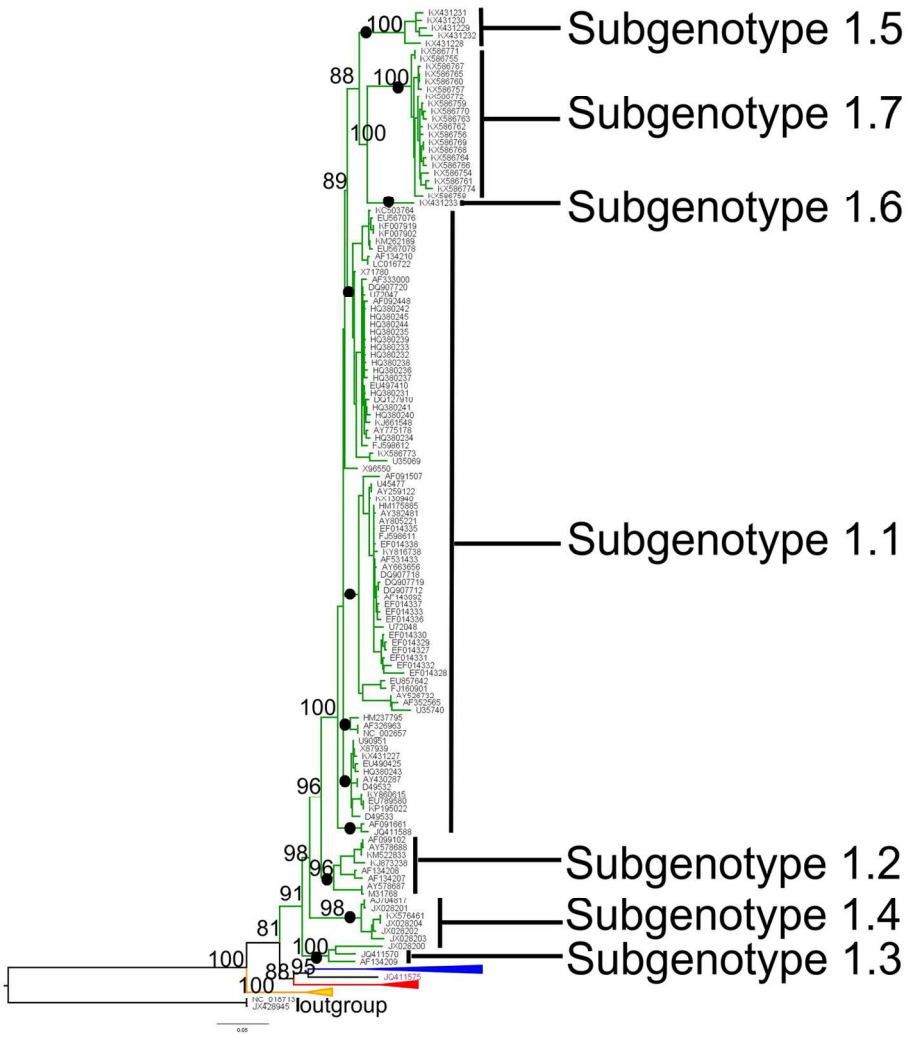


Fig. 2. Phylogenetic tree for CSFV genotype 1. All non-redundant genomes were analysed using ML-method, the GenBank IDs for all the sequences are shown. All the lineages assessed within the CSFV genotype 1 (green) are denoted with black circles, the main sublineages proposed as CSFV subgenotypes within CSFV genotype 1 are denoted. All the remaining CSFV genotypes were collapsed: CSFV-genotype 2: blue, CSFV-genotype 3: red, CSFV-genotype 4: purple, and CSFV-genotype 5: orange. Numbers along the branches refer to the percentages of confidence and minor branch values were hidden.

139x153mm (300 x 300 DPI)

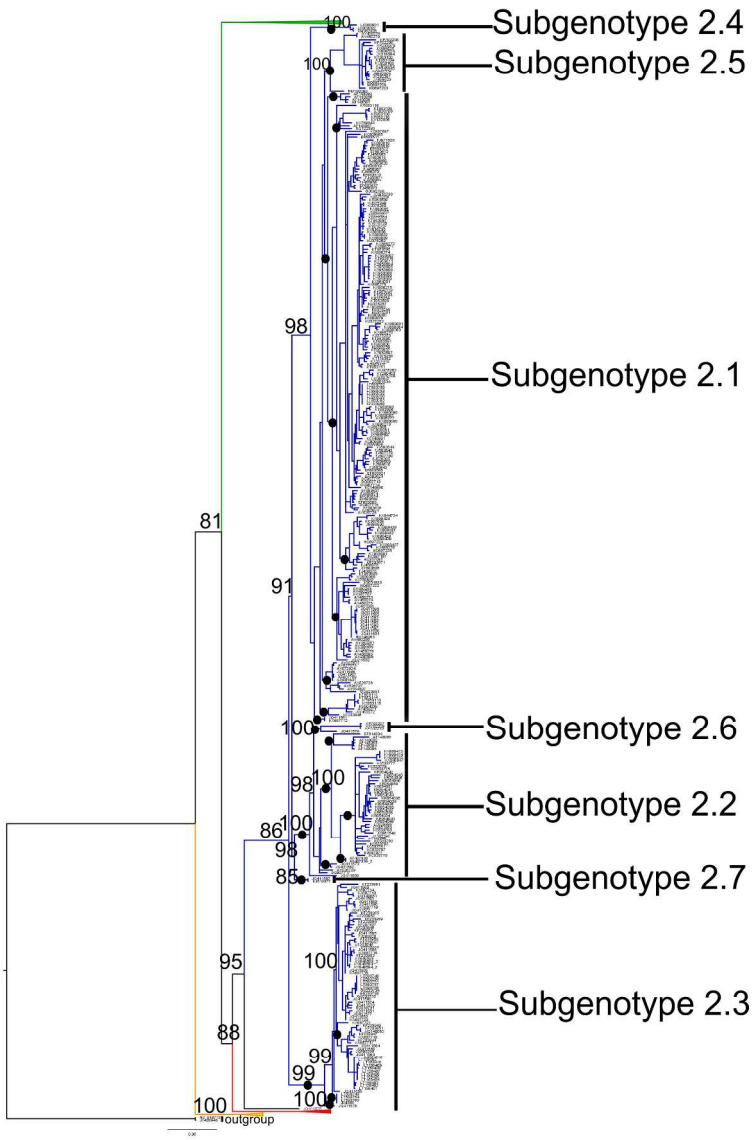


Fig. 3. Phylogenetic tree for CSFV genotype 2. All non-redundant genomes were analysed using ML-method, the GenBank IDs for all the sequences are shown. All the lineages assessed within the CSFV genotype 2 (blue) are denoted with black circles, the main sublineages proposed as CSFV subgenotypes within CSFV genotype 2 are denoted. All the remaining CSFV genotypes were collapsed: CSFV-genotype 1: green, CSFV-genotype 3: red, CSFV-genotype 4: purple, and CSFV-genotype 5: orange. Numbers along the branches refer to the percentages of confidence and minor branch values were hidden

203x270mm (300 x 300 DPI)

GenBank ID	Country	Collection Date
AF352565	Taiwan	Not available
AB897785	Japan	2012
D49533	Japan	Not available
D49532	Japan	Not available
AF407339	China	Not available
LT593749	Not available	Not available
LT593750	Not available	Not available
LT593751	Not available	Not available
LT593757	Not available	Not available
LT593758	Not available	Not available
LT593748	Not available	Not available
LT593754	Not available	Not available
LT593760	Not available	Not available
KY849593	Serbia	2005
KY849594	Serbia	2006
LT593752	Not available	Not available
LT593753	Not available	Not available
LT593755	Not available	Not available
LT593756	Not available	Not available
LT593759	Not available	Not available
LT593761	Not available	Not available
LT593762	Not available	Not available
JQ595295	Belgium	1993-1994
AY072924	Denmark	Not available
LC016722	Thailand	1993
JQ411563	Germany	1992
JQ411564	Germany	1994
JQ411565	Austria	1994
JQ411568	Poland	1995
JQ411572	Czech Republic	1996
JQ411574	Germany	1997
JQ411576	Germany	1995
JQ411577	Germany	1997
JQ411580	Germany	1998
JQ411581	Germany	1998
JQ411583	Slovakia	2000
JQ411584	Germany	2000
JQ411585	Spain	2001
JQ411586	Croatia	2006
JQ411559	Germany	1984
JQ411560	Germany	1989
JQ411561	Germany	1989
JQ411562	Austria	1990

JQ411566	Germany	1997
JQ411567	Netherlands	1997
JQ411569	Poland	1995
JQ411570	Malaysia	1986
JQ411571	Japan	1974
JQ411573	Czech Republic	1994
JQ411575	United Kingdom	1964
JQ411578	Germany	1982
JQ411579	Italy	1998
JQ411582	United Kingdom	2000
JQ411587	Netherlands	Not available
JQ411588	Italy	1951
JQ411589	Hungary	2007
JQ411590	Slovakia	2007
JQ411591	Lithuania	2009
JX028200	Guatemala	Not available
JX028202	Cuba	2009
JX028204	Cuba	2010
JX028201	Cuba	1958
JX028203	Cuba	2011
AY526726	Taiwan	2003
AY526727	Taiwan	1994
AY526728	Taiwan	2001
AY526729	Taiwan	1995
AY526730	Taiwan	1996
AY526731	Taiwan	1990
AY526732	Taiwan	
U35069	Taiwan	Not available
U35740	Taiwan	Not available
X71780	United Kingdom	Not available
AJ704817	Cuba	
EU567076	India	
EU567077	India	
EU567078	India	
FJ160901	India	
KF007902	India	2012
KF007919	India	2011
HQ697222	China	2008
HQ697223	China	2008
HQ697224	China	2009
HQ697225	China	2009
HQ697226	China	2010
HQ697227	China	2009
HQ697228	China	2009

HCU03290	Not available	Not available
AH007708	Not available	Not available
DQ907712	Not available	Not available
DQ907713	Not available	Not available
DQ907714	Not available	Not available
DQ907715	Not available	Not available
DQ907716	Not available	Not available
DQ907719	Not available	Not available
NM_001204369	Not available	Not available
KX130940	Indonesia	2013
AF143082	China	Not available
AF143092	China	Not available
AF143083	China	Not available
AF143084	China	Not available
AF143085	China	Not available
AF143086	China	Not available
AF143087	China	Not available
AF143088	China	Not available
AF143089	China	Not available
AF143090	China	Not available
AF143091	China	Not available
KY990413	Not available	Not available
KY990414	Bangladesh	2015
KY990415	Bangladesh	2015
HCU72047	Not available	Not available
HCU72048	Not available	Not available
KC533781	India	2009
KC533787	India	2010
AY430287	Not available	Not available
DQ907717	Not available	Not available
DQ907718	Not available	Not available
DQ907720	Not available	Not available
KC533782	India	Not available
AY027672	Not available	Not available
AY027673	Not available	Not available
EF683605	China	Not available
EF683606	China	Not available
EF683607	China	Not available
EF683608	China	Not available
EF683609	China	Not available
EF683610	China	Not available
EF683611	China	Not available
EF683612	China	Not available
EF683613	China	Not available

EF683614	China	Not available
EF683615	China	Not available
EF683616	China	Not available
EF683617	China	Not available
EF683618	China	Not available
EF683619	China	Not available
EF683620	China	Not available
EF683621	China	Not available
EF683622	China	Not available
EF683623	China	Not available
FJ456865	China	2007
FJ456866	China	2007
FJ456867	China	2006
FJ456868	China	2007
FJ456869	China	2005
FJ456870	China	2004
FJ456871	China	2005
FJ456872	China	2006
FJ456873	China	2006
FJ456874	China	2006
FJ456875	China	2004
FJ456876	China	2007
FJ582642	China	2008
FJ582643	China	2008
FJ582644	China	2008
FJ598609	China	2008
FJ598610	China	2008
FJ598611	China	2008
FJ598612	China	2008
FJ607779	China	2008
FJ607780	China	2008
FJ977628	China	2009
HM190299	South Africa	2005
HQ317681	China	2010
HQ380232	China	2009
HQ380233	China	2010
HQ380234	China	2009
HQ380235	China	2008
HQ380236	China	2008
HQ380237	China	2008
HQ380238	China	2009
HQ380239	China	2009
HQ380240	China	2009
HQ380241	China	2009

HQ380242	China	2010
HQ380243	China	2010
HQ380244	China	2009
HQ380245	China	2009
JN882005	China	2011
JQ001833	China	2011
JQ001834	China	2011
JN886990	China	2011
JQ411592	Lithuania	2011
JQ411593	Lithuania	2011
JQ411594	Lithuania	2011
JQ411595	Lithuania	2011
JQ411596	Lithuania	2011
JQ411597	Lithuania	2011
JQ411598	Lithuania	2011
JQ411599	Lithuania	2011
JQ411600	Lithuania	2011
JQ411601	Lithuania	2011
JX162240	Nepal	2011
JX162241	Nepal	2011
JX898523	China	2011
JX898524	China	2012
JX898525	China	2012
KC597187	China	2012
KC809979	China	2011
KC809980	China	2012
KC809981	China	2011
KC809982	China	2011
KC809983	China	2011
KC809984	China	2011
KC809985	China	2011
KC809986	China	2012
KC867687	China	2011
KC867688	China	2012
KC867689	China	2012
KF297337	Romania	2007
KF233944	Bulgaria	1997
KF233945	Croatia	1997
KF233946	Not available	2000
KF233947	Bulgaria	2002
KF233948	Romania	2004
KF233949	Bulgaria	2006
KF233950	Serbia	2006
KF233951	Bulgaria	2007

KF233952	Bulgaria	2007
KF233953	Croatia	2007
KF233954	Romania	2006
KF233955	Romania	2006
KF233956	Romania	2007
KF233957	Bulgaria	2008
KF233958	Israel	2009
KF233959	Bulgaria	2009
KF233960	Serbia	2010
KF233961	Latvia	2013
KJ661548	China	2008
LC000001	India	2013
LC000002	India	2012
LC029886	India	2012
KP702206	Viet Nam	2014
KP702207	Viet Nam	2014
KP702208	Viet Nam	2014
KP702209	Viet Nam	2014
KP702210	Viet Nam	2014
KR054034	India	Not available
KR054035	India	Not available
KR054036	India	Not available
KR054037	India	Not available
KR054038	India	Not available
KR054039	India	Not available
KR054040	India	Not available
KR054041	India	Not available
KR054042	India	Not available
KR054043	India	Not available
KR054044	India	Not available
KR054045	India	Not available
KR054046	India	Not available
KR054047	India	Not available
KR054048	India	Not available
KR054049	India	Not available
KR054050	India	Not available
KR054051	India	Not available
KR054052	India	Not available
KT953587	China	2015
KT953588	China	2015
KT953589	China	2014
KT953590	China	2015
KT953591	China	2015
KT953592	China	2013

KT953593	China	2015
KT953594	China	2015
KT953595	China	2015
KT953596	China	2015
KT953597	China	2015
KT953598	China	2015
KT953599	China	2015
KT953600	China	2015
KT953601	China	2015
KT953602	China	2015
KT953603	China	2015
KT953604	China	2014
KT953605	China	2015
KT953606	China	2015
KT953607	China	2015
KT953608	China	2015
KT953609	China	2015
KT953610	China	2015
KT953611	China	2015
KT853102	China	2011
KT853103	China	2011
KT853104	China	2011
KT853105	China	2011
KT853106	China	2011
KT853107	China	2011
KT853108	China	2011
KT853109	China	2011
KT853110	China	2011
KT853111	China	2011
KT853112	China	2011
KT853113	China	2011
KT853114	China	2011
KT853115	China	2011
KT853116	China	2011
KU375249	China	2014
KU375250	China	2014
KU375251	China	2014
KU375252	China	2012
KU375253	China	2013
KU375254	China	2015
KU375255	China	2015
KU375256	China	2014
KU375257	China	2014
KU375258	China	2014

KU375259	China	2014
KU375260	China	2014
KU375261	China	2014
KU375262	China	2014
KU375263	China	2015
KX431227	Brazil	2008
KX431228	Brazil	2001
KX431229	Brazil	2003
KX431230	Brazil	2003
KX431231	Brazil	2004
KX431232	Brazil	2009
KX431233	Brazil	2009
KX687712	Germany	1989
KX687713	Croatia	2002
KX687714	Croatia	2002
KX687715	Romania	1994
KX687716	Croatia	2007
KX687717	Bulgaria	1998
KX687718	Bulgaria	2002
KX687719	Slovakia	2007
KX687720	Germany	1993
KX687721	Hungary	1992
KX759642	China	2009
KX759643	China	2006
KX257416	China	Not available
KX586754	Ecuador	2015
KX586755	Ecuador	2014
KX586756	Ecuador	2013
KX586757	Ecuador	2014
KX586758	Ecuador	2014
KX586759	Ecuador	2014
KX586760	Ecuador	2015
KX586761	Ecuador	2015
KX586762	Ecuador	2012
KX586763	Ecuador	2015
KX586764	Ecuador	2015
KX586765	Ecuador	2015
KX586766	Ecuador	2015
KX586767	Ecuador	2015
KX586768	Ecuador	2015
KX586769	Ecuador	2015
KX586770	Ecuador	2015
KX586771	Ecuador	2013
KX586772	Ecuador	2012

KX586773	Ecuador	2000
KX586774	Ecuador	2015
KY816735	Not available	2014
KY816736	Not available	2014
KY816737	Not available	2014
KY816738	Not available	2014
KY816739	Not available	2014
KX844724	Not available	2014
KX880080	China	2015
KX880081	China	2014
KX880082	China	2016
KX880083	China	2015
KX880084	China	2014
KX880085	China	2016
KX880086	China	2016
KX880087	China	2016
KX880088	China	2015
KX880089	China	2016
KX898427	China	2014
KX898428	China	2014
KX898429	China	2014
KX898430	China	2014
KX898431	China	2014
KX898432	China	2014
KX898433	China	2014
KX886270	China	2015
KX886271	China	2016
KX886272	China	2016
KX886273	China	2015
KX886274	China	2015
KX886275	China	2015
KX886276	China	2015
KX886277	China	2016
KX886278	China	2015
KX886279	China	2014
HCU43924	Not available	Not available
AF134207	Thailand	Not available
AF134208	Thailand	Not available
AF134209	Thailand	Not available
AF134210	Thailand	Not available
KC533777	India	2011
KC533779	India	2011
AY450271	Not available	Not available
AY450272	Not available	Not available

AY450273	Not available	Not available
AY450274	Not available	Not available
AY450275	Not available	Not available
AY450276	Not available	Not available
AY450277	Not available	Not available
AY450278	Not available	Not available
AY450279	Not available	Not available
AY450280	Not available	Not available
AY450281	Not available	Not available
AY450282	Not available	Not available
AY450283	Not available	Not available
AY450284	Not available	Not available
EF014327	Not available	Not available
EF014328	Not available	Not available
EF014329	Not available	Not available
EF014330	Not available	Not available
EF014331	Not available	Not available
EF014332	Not available	Not available
EF014333	Not available	Not available
EF014334	Not available	Not available
EF014335	Not available	Not available
EF014336	Not available	Not available
EF014337	Not available	Not available
EF014338	Not available	Not available
KC533783	India	2006
AY430095	Not available	Not available
AY430096	Not available	Not available
KX345847	Bangladesh	2015
KP195022	India	2014
AF091507	Not available	Not available
AF091661	Italy	1951
AF092448	China	Not available
AF099102	Russia	Not available
AF326963	Germany	1965
AF333000	China	Not available
AF531433	China	Not available
AY259122	Switzerland	Not available
AY367767	China	Not available
AY382481	China	Not available
AY554397	Taiwan	Not available
AY578687	Italy	2001
AY578688	Not available	2001
AY646427	Taiwan	1994
AY663656	China	2003

AY775178	China	1945
AY805221	China	Not available
DQ127910	China	2004
EU490425	France	Not available
EU497410	China	2006
EU789580	Japan	1980
EU857642	India	Not available
FJ265020	Spain	2001
FJ529205	China	2010
GQ122383	China	2006
GQ902941	Denmark	Not available
GQ923951	China	2009
GU233731	Germany	2006
GU233732	Germany	2005
GU233733	Germany	2009
GU233734	Germany	2009
GU324242	Germany	2004
GU592790	China	2009
HM175885	China	2008
HM237795	Czech Republic	Not available
HQ148061	Croatia	2002
HQ148062	Bulgaria	2007
HQ148063	Lithuania	2009
HQ380231	China	2009
J04358	Germany	1999
JQ268754	China	2010
JQ861548	India	2011
JX218094	China	2012
JX262391	China	2011
KC149990	South Korea	2011
KC149991	South Korea	2011
KC503764	India	2011
KC533775	India	2006
KC533776	India	Not available
KC533793	India	2011
KC851953	India	2012
KF669877	South Korea	1998
KJ619377	Netherlands	1977
KJ873238	USA	1994
KM262189	India	2009
KM362426	India	Not available
KM522833	USSR	Not available
KP233070	China	2013
KP233071	China	2013

KP343640	China	2011
KT119352	China	2014
KT716271	South Korea	1999
KU504339	China	2011
KU556758	China	2015
KX064281	China	2015
KX576461	Cuba	2010
KX870109	South Korea	2016
KY132096	China	2011
KY290453	South Korea	2016
KY860615	India	2013
LC086647	Mongolia	2014
LT158401	Not available	Not available
LT158402	Not available	Not available
LT158403	Not available	Not available
LT158404	Not available	Not available
LT158405	Not available	Not available
LT158406	Not available	Not available
LT158407	Not available	Not available
LT158408	Not available	Not available
LT158409	Not available	Not available
LT158410	Not available	Not available
LT158502	Not available	Not available
M31768	Netherlands	Not available
NC_00265	Switzerland	2000
U35069	Taiwan	Not available
U35740	Taiwan	Not available
U43924	Taiwan	Not available
U45477	Germany	Not available
U72047	China	Not available
U72048	Not available	Not available
U90951	France	Not available
X87939	Switzerland	Not available
X96550	Switzerland	Not available
LT593749	Duplicated	
LT593750	Duplicated	
LT593751	Duplicated	
LT593757	Duplicated	
LT593758	Duplicated	
LT593748	Duplicated	
LT593754	Duplicated	
LT593760	Duplicated	
LT593753	Duplicated	
LT593755	Duplicated	

LT593756	Duplicated	
LT593761	Duplicated	
LT593762	Duplicated	
KF007902	Duplicated	
KF007919	Duplicated	
HQ697227	Duplicated	
HQ697228	Duplicated	
KY990413	Duplicated	
KY990414	Duplicated	
KY990415	Duplicated	
KC533781	Duplicated	
KC533787	Duplicated	
HQ380235	Duplicated	
HQ380244	Duplicated	
HQ380245	Duplicated	
JQ411593	Duplicated	
JQ411594	Duplicated	
JQ411595	Duplicated	
JQ411596	Duplicated	
JQ411597	Duplicated	
JQ411598	Duplicated	
JQ411599	Duplicated	
JQ411600	Duplicated	
JQ411601	Duplicated	
KT853104	Duplicated	
KT853108	Duplicated	
KT853109	Duplicated	
KT853110	Duplicated	
KT853112	Duplicated	
KT853114	Duplicated	
KT853113	Duplicated	
KT853115	Duplicated	
KU375249	Duplicated	
KU375250	Duplicated	
KX586760	Duplicated	
KX586765	Duplicated	
KX586768	Duplicated	
KX586769	Duplicated	

Table S2. Genetic distances based on full-length E2 gene sequences of all the lineages assessed within the CSFV genotype 1**.** Values above the diagonal represent the standard error, values below the diagonal represent the p-distance values obtained using MEGA7 and 1000 bootstrap replicates.

Subgenotype	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10	1.11
1.1		0.007	0.008	0.009	0.002	0.001	0.002	0.005	0.005	0.006	0.004
1.2	0.096		0.006	0.004	0.008	0.009	0.005	0.007	0.007	0.008	0.007
1.3	0.104	0.102		0.009	0.001	0.010	0.009	0.008	0.008	0.009	0.008
1.4	0.116	0.107	0.120		0.011	0.010	0.010	0.010	0.009	0.010	0.010
1.5	0.096	0.131	0.138	0.134		0.007	0.009	0.008	0.008	0.008	0.008
1.6	0.095	0.118	0.137	0.134	0.098		0.009	0.008	0.008	0.009	0.008
1.7	0.096	0.104	0.128	0.131	0.108	0.106		0.008	0.008	0.009	0.009
1.8	0.041	0.083	0.099	0.113	0.089	0.081	0.085		0.005	0.006	0.005
1.9	0.049	0.083	0.102	0.111	0.100	0.090	0.091	0.044		0.006	0.005
1.10	0.055	0.100	0.113	0.123	0.106	0.101	0.100	0.061	0.069		0.006
1.11	0.032	0.078	0.098	0.104	0.087	0.081	0.088	0.039	0.046	0.054	

Table S3. Genetic distances based on full-length E2 gene sequences of all the lineages assessed within the CSFV genotype 2**.** Values above the diagonal represent the standard error, values below the diagonal represent the p-distance values obtained using MEGA7 and 1000 bootstrap replicates.

Subgenotype	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	2.10	2.11	2.12	2.13	2.14	2.15	2.16	2.17	2.18	2.19
2.1		*0.008*	*0.012*	*0.011*	*0.002*	*0.001*	*0.008*	*0.006*	*0.005*	*0.008*	*0.009*	*0.008*	*0.011*	*0.013*	*0.011*	*0.013*	*0.012*	*0.013*	*0.013*
2.2	0.134		*0.008*	*0.012*	*0.012*	*0.012*	*0.012*	*0.012*	*0.012*	*0.012*	*0.012*	*0.012*	*0.010*	*0.010*	*0.008*	*0.006*	*0.009*	*0.012*	*0.012*
2.3	0.138	0.130		*0.012*	*0.013*	*0.012*	*0.012*	*0.012*	*0.012*	*0.013*	*0.012*	*0.012*	*0.011*	*0.011*	*0.012*	*0.012*	*0.011*	*0.008*	*0.005*
2.4	0.119	0.135	0.136		*0.012*	*0.012*	*0.001*	*0.012*	*0.011*	*0.012*	*0.011*	*0.011*	*0.012*	*0.012*	*0.013*	*0.013*	*0.013*	*0.013*	*0.013*
2.5	0.097	0.130	0.138	0.113		*0.004*	*0.001*	*0.009*	*0.009*	*0.010*	*0.010*	*0.008*	*0.011*	*0.012*	*0.011*	*0.012*	*0.012*	*0.013*	*0.013*
2.6	0.099	0.130	0.130	0.107	0.097		*0.001*	*0.008*	*0.009*	*0.010*	*0.010*	*0.008*	*0.011*	*0.013*	*0.012*	*0.012*	*0.012*	*0.014*	*0.013*
2.7	0.147	0.108	0.131	0.139	0.133	0.146		*0.007*	*0.007*	*0.009*	*0.008*	*0.007*	*0.010*	*0.012*	*0.011*	*0.012*	*0.012*	*0.013*	*0.012*
2.8	0.051	0.139	0.138	0.114	0.081	0.064	0.059		*0.007*	*0.006*	*0.009*	*0.007*	*0.010*	*0.012*	*0.011*	*0.012*	*0.012*	*0.013*	*0.012*
2.9	0.051	0.138	0.140	0.111	0.086	0.069	0.064	0.053		*0.010*	*0.008*	*0.009*	*0.011*	*0.012*	*0.012*	*0.013*	*0.012*	*0.014*	*0.013*
2.10	0.063	0.142	0.152	0.122	0.083	0.077	0.074	0.057	0.064		*0.010*	*0.009*	*0.011*	*0.013*	*0.012*	*0.013*	*0.013*	*0.014*	*0.013*
2.11	0.077	0.138	0.137	0.105	0.083	0.074	0.061	0.070	0.074	0.074		*0.007*	*0.010*	*0.012*	*0.011*	*0.012*	*0.012*	*0.013*	*0.012*
2.12	0.064	0.127	0.122	0.083	0.062	0.054	0.040	0.054	0.060	0.063	0.057		*0.010*	*0.012*	*0.011*	*0.012*	*0.012*	*0.013*	*0.012*
2.13	0.081	0.096	0.093	0.107	0.098	0.098	0.081	0.090	0.089	0.112	0.101	0.073		*0.010*	*0.010*	*0.011*	*0.009*	*0.012*	*0.011*
2.14	0.099	0.540	0.129	0.093	0.094	0.091	0.132	0.142	0.135	0.150	0.135	0.122	0.081		*0.010*	*0.011*	*0.009*	*0.012*	*0.012*
2.15	0.136	0.078	0.122	0.142	0.119	0.125	0.119	0.128	0.127	0.136	0.128	0.110	0.081	0.096		*0.008*	*0.008*	*0.013*	*0.012*
2.16	0.134	0.046	0.130	0.143	0.126	0.115	0.122	0.128	0.127	0.136	0.131	0.115	0.080	0.092	0.056		*0.009*	*0.013*	*0.012*
2.17	0.130	0.079	0.106	0.123	0.121	0.112	0.110	0.120	0.117	0.141	0.119	0.107	0.060	0.077	0.069	0.055		*0.012*	*0.011*
2.18	0.147	0.137	0.054	0.140	0.137	0.139	0.125	0.135	0.138	0.154	0.138	0.118	0.087	0.121	0.119	0.127	0.101		*0.007*
2.19	0.136	0.123	0.033	0.127	0.131	0.123	0.117	0.124	0.127	0.141	0.124	0.106	0.073	0.114	0.105	0.110	0.087	0.028	

Table S4. Genetic distances based on full-length E2 gene sequences of all the lineages assessed within the CSFV genotype 3**.** Values above the diagonal represent the standard error, values below the diagonal represent the p-distance values obtained using MEGA7 and 1000 bootstrap replicates.

Subgenotype	3.1	3.2	3.3
3.1		0.007	0.006
3.2	0.047		0.004
3.3	0.044	0.022	

Table S5. Genetic distances based on full-length E2 gene sequences of all the lineages assessed within the CSFV genotype 5**.** Values above the diagonal represent the standard error, values below the diagonal represent the p-distance values obtained using MEGA7 and 1000 bootstrap replicates.

Subgenotype	5.1	5.2
5.1		0.006
5.2	0.045	

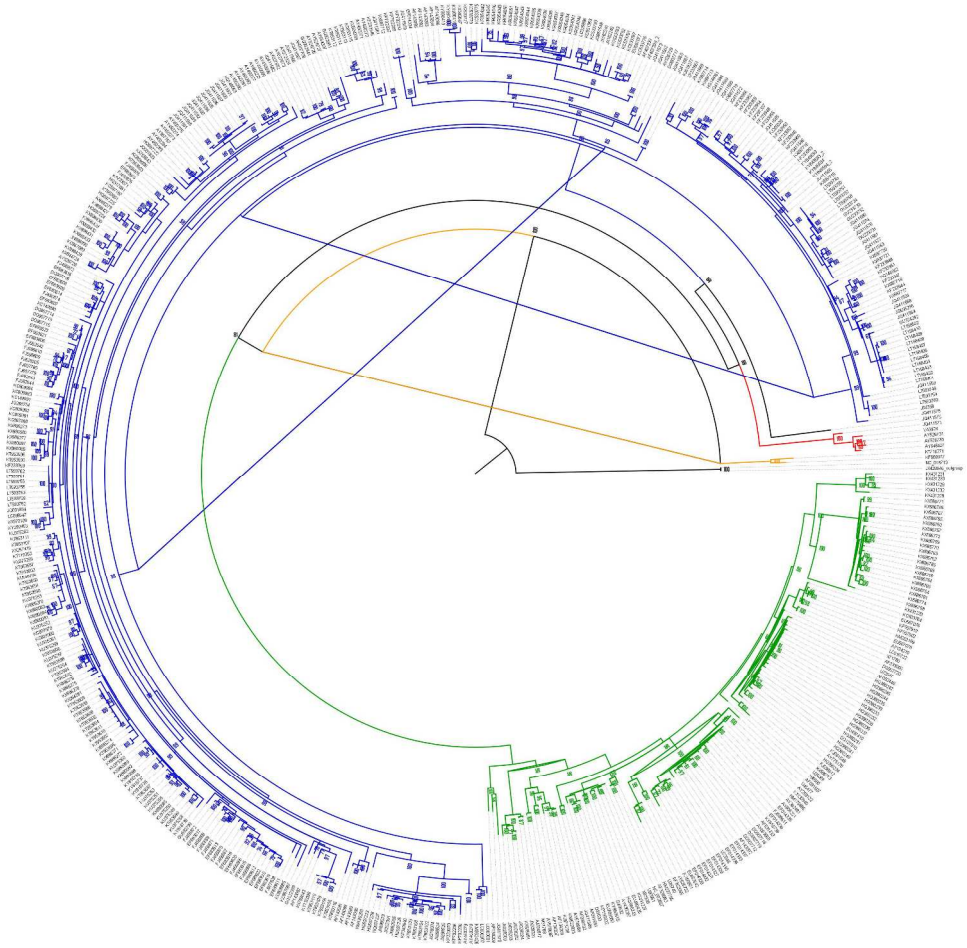


Figure S1

257x257mm (300 x 300 DPI)

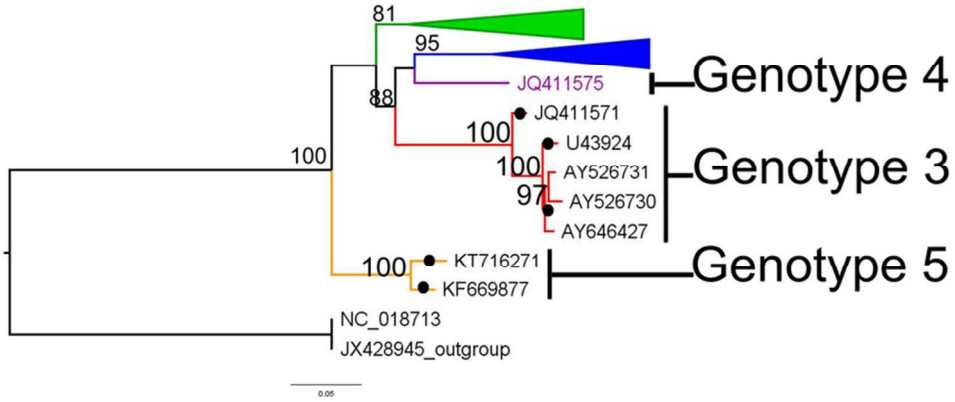


Figure S2

76x45mm (300 x 300 DPI)