WILEY | Hindawi

Research Article

Tracing the Origin of Genotype II African Swine Fever Virus in China by Genomic Epidemiology Analysis

Yong Zhang,[1,2,3] **Qinghua Wang,**[1] **Zhongyi Zhu,**[4] **Shujuan Wang,**[1] **Shuyang Tu,**[5]
Yongqiang Zhang,[1] **Yanli Zou,**[1] **Yutian Liu,**[1] **Chunju Liu,**[1] **Weijie Ren,**[1] **Dongxia Zheng,**[1]
Yunling Zhao,[1] **Yongxin Hu,**[1] **Lin Li,**[1] **Chuan Shi,**[1,3,6] **Shengqiang Ge,**[1] **Peng Lin,**[3,6]
Fengping Xu,[2,3] **Jinmin Ma,**[4] **Xiaodong Wu,**[1] **Hongchao Ma** **,**[1] **Zhiliang Wang** **,**[1]
and Jingyue Bao [1]

[1]*China Animal Health and Epidemiology Center, Qingdao 266032, China*
[2]*Lars Bolund Institute of Regenerative Medicine, Qingdao-Europe Advanced Institute for Life Sciences, BGI-Qingdao, BGI-Shenzhen, Qingdao 266555, China*
[3]*College of Life Sciences, University of Chinese Academy of Sciences, Beijing 518083, China*
[4]*Pathogenesis Pharmaceutical Technology, BGI-Shenzhen, Shenzhen 518083, China*
[5]*National Facility for Protein Science in Shanghai, Zhangjiang Laboratory, Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai 201210, China*
[6]*BGI-Qingdao, BGI-Shenzhen, Qingdao 266555, China*

Correspondence should be addressed to Hongchao Ma; mhcqd999@163.com, Zhiliang Wang; zlwang111@163.com, and Jingyue Bao; baojingyue88@163.com

The pandemic spread of African swine fever (ASF) has caused serious effects on the global pig industry. Virus genome sequencing and genomic epidemiology analysis play an important role in tracking the outbreaks of the disease and tracing the transmission of the virus. Here we obtained the full-length genome sequence of African swine fever virus (ASFV) in the first outbreak of ASF in China on August 3rd, 2018 and compared it with other published genotype II ASFV genomes including 9 genomes collected in China from September 2018 to October 2020. Phylogenetic analysis on genomic sequences revealed that genotype II ASFV has evolved into different genetic clusters with temporal and spatial correlation since being introduced into Europe and then Asia. There was a strong support for the monophyletic grouping of all the ASFV genome sequences from China and other Asian countries, which shared a common ancestor with those from the Central or Eastern Europe. An evolutionary rate of 1.312×10^{-5} nucleotide substitutions per site per year was estimated for genotype II ASFV genomes. Eight single nucleotide variations which located in MGF110-1L, MGF110-7L, MGF360-10L, MGF505-5R, MGF505-9R, K145R, NP419L, and I267L were identified as anchor mutations that defined genetic clusters of genotype II ASFV in Europe and Asia. This study expanded our knowledge of the molecular epidemiology of ASFV and provided valuable information for effective control of the disease.

1. Introduction

African swine fever (ASF) is a lethal contagious viral disease of domestic pigs and wild boars, which also affects African wild suits (warthogs and bushpigs) in an asymptomatic carrier state. Soft ticks serve as a natural reservoir and transmit the disease to the vertebrate host. African swine fever virus (ASFV) was first identified in Kenya in 1921. Since then, it has been reported in Africa, Europe, and South America [1]. In 2007, ASFV was introduced into the Georgian Republic and then swept across Europe and Asia [2–6]. In 2021, ASF was reported in Dominica and Haiti [7]. With no vaccine or any treatment available, affected pigs have been culled to try to contain the outbreaks. More than

50 countries are now affected by ASF, causing the death or culling of more than 9 million pigs. The pandemic spread of ASF has caused serious effects on the global pig industry, even leading to a critical global heparin shortage [8, 9].

ASFV is a large, encapsulated double-stranded DNA virus belonging to the *Asfivirus* genus of the *Asfarviridae* family. *Asfarviridae* family, together with the families *Mimiviridae, Pithoviridae, Marseilleviridae, Pandoraviruses, Phycodnaviridae, Iridoviridae,* and *Poxviridae,* composing the group of nucleocytoplasmic large DNA viruses (NCLDV), are also known as giant viruses [10]. The main primary target cells of ASFV belong to the monocyte/ macrophage lineage. Upon ASFV infection, the highest titres are observed in tissues containing a large component of the mononuclear phagocyte system (reticuloendothelial cellular elements) like the spleen and lymph nodes [11]. Therefore, it is recommended in *African swine fever detection and diagnosis, a manual for veterinarians published by* the *FAO* that spleen and lymph nodes are the most important target organs being used to check for the presence of ASFV [12].

The genome of ASFV is a linear double-stranded DNA (dsDNA) molecule in the length of 171 kB to 193 kB with terminal inverted repeats and hairpin loops. It has a conserved central region (CCR) of about 125 kB, while the ends are variable in size. The left variable region and the right variable region contain different number of five multigene family (MGF) genes: MGF 100, 110, 300, 360, and 505/530, named after the average number of encoded amino acids. The CCR contains genes involving virus replication, assembly, and host cell function modulation. Phylogenetic analysis based on the partial sequence of the B646L gene (440 bp, nucleotide positions 103,630-104,069 of the ASFV genome) has revealed 24 genotypes of ASFV in Africa, from genotype I to genotype XXIIII. Genotype I is also endemic in Sardinia, Italy. ASFV strains recently circulating in Europe and Asia are classified as genotype II [5]. Several molecular markers have been used to typing genotype II ASFV in Europe and Asia [13–15]. However, the limited information provided by a partial genomic sequence was not sufficient for tracing the transmission of ASFV.

Genomic epidemiology analysis has been widely used to trace the transmission of RNA viruses such as SARS-CoV-2, Ebola, and Zika virus, which provided valuable information for the interpretation of field epidemiology data and the implementation of efficient control measures [16–19]. The development of next-generation sequencing (NGS) and third-generation sequencing techniques and hence the reduction of genome sequencing cost has made feasible large-scale sequencing for large dsDNA virus genomes [20, 21]. Genome sequence diversity analysis of 45 epidemiologically varied virus isolates has been conducted to reveal clues to the evolution of the variola (smallpox) virus [22]. Dynamic genome evolution and complex virocell metabolism of globally-distributed giant viruses was investigated by generating and analyzing 501 metagenome-assembled genomes of NCLDV from environments around the globe [10]. A curated dataset has recently been constructed, containing a total of 123 ASFV genome sequences including 64 genotype I genomes, 42 genotype II genomes, and 17 genomes

belonging to other genotypes [23]. The genome-wide genome similarity between ASFV genotype I and genotype II ranges between 82.4–94.7%. A few molecular epidemiology studies based on genome sequences have been conducted to reconstruct the evolution of ASFV globally or to trace the origin and evolution of genotype I ASFV in Sardinia [24–26]. Up to date, the evolutionary dynamics of genotype II ASFV, currently widespread in Eurasia remain unclear.

In China, the first outbreak of ASF was reported on August 3rd, 2018 in Liaoning province. The causative strain of this outbreak was identified as ASFV genotype II by molecular characterization based on the partial sequence of the B646L gene [2]. However, the full-length genome sequence of the ASFV that caused the first outbreak of ASF in China was not available. The first published genome sequence of ASFV in China was generated from field samples collected in an ASF outbreak on September 2nd, 2018 in Anhui province [27]. From then on, genome characterization of ASFV strains circulating in China has been described. To trace the origin of ASFV in China, it is essential to determine the complete genome sequence of ASFV in the first outbreak.

In this study, we generate a novel full-length genome sequence of ASFV from clinical samples collected in the first outbreak of ASF in China and assess the genome sequence diversity of genotype II ASFV in China. In addition, we present phylogenetic analysis to demonstrate that genotype II ASFV has evolved into different genetic clusters since being introduced into Europe and then Asia. Moreover, we identify anchor mutations that define the genetic clusters of genotype II ASFV. We verify a strong monophyletic relationship among all the ASFV genome sequences from China and other Asian countries, which share a common ancestor with those from Central or Eastern Europe. Our study serves as an example of how genomic epidemiology can be used to trace the origin of ASF and track the transmission of the virus. Our findings also provide insight into the genomic evolution of genotype II ASFV during the ongoing pandemic in Europe and Asia, highlighting its important role in controlling this lethal swine disease.

2. Materials and Methods

2.1. Genome Sequencing and Genome Sequence Comparison. On August 3rd, 2018, an outbreak of ASF was reported in a swine farm in Shenbei district, Shenyang city in Liaoning province. Eight infected pigs were sampled for diagnosis at the infected farm. Spleen, lymph node, and kidney samples were collected from each pig and sent to our laboratory for confirmation. Clinical samples were screened with real-time PCR with amplification targeting the B646L gene. The ASFV China/LN/2018/1 strain was grouped into genotype II by partial sequencing of the B646L gene. For each pig, the sample with the lowest Ct value in the ASFV real-time PCR assay was further used for genome sequencing. Genome sequences were determined directly from the clinical samples to avoid cell-culture-driven sequence variation, as previously reported [27]. An identical ASFV genome sequence was found in samples from 8 pigs. Therefore, the

ASFV genome from the spleen of a 6-month-old infected pig was used in this study to represent the first ASF outbreak in China. Nine genotype II ASFV genomes described in a recently published curated dataset were used for genome wide diversity analysis on genotype II ASFV in China. These genomes represented genotype II ASFV collected in China from September 2nd, 2018 to October 1st, 2020 (Table 1).

2.2. Genome Sequences Alignment and Recombination Analysis. A total of forty-two genotype II ASFV genomes described in a curated dataset were used for phylogenetic analysis [23]. The genome of LR812933.1/Arm/07/CBM/c2 was not included in this study for fear of possible contamination. The final dataset included a total of 41 genotype II ASFV genomes collected from 15 countries from 2007 to 2020. The GenBank accession number, the year, and the country of isolation were listed in Table 1. Multiple nucleotide alignments of genome sequences was conducted using the MAFFT software (version 7.475) [45]. The alignment of ASFV genomes was analyzed for recombination using seven different recombination detection methods (RDP, GENECONV, BootScan, MaxChi, Chimaera, SiScan, and 3Seq).

2.3. Phylogenetic Analysis. Maximum likelihood (ML) phylogenetic trees were estimated by RAxML v8.2.12 [46] using GTR + G nucleotide substitution model. ML bootstrapping was performed with 1000 resamples to assess the robustness of tree topologies. The final tree was midpoint rooted by FigTree v1.4.2. (https://tree.bio.ed.ac.uk/software/figtree/). Bayesian inference of phylogeny was performed using the Bayesian Markov Chain Monte Carlo (MCMC) method in the BEAST package 1.8.2 [47]. The most suitable substitution models were selected in JModelTestv2.1.10 [48]. The dataset was analyzed using the GTR + G substitution model. The Markov Chain Monte Carlo chains were run for 2×10^8 generations and sampled every 2000 generations. TreeAnnotator v1.10.4 was used to generate maximum-clade credibility (MCC) trees with a burn-in rate of 10%. Summary phylogenies were visualized in FigTreev1.4.2. Phylogenetic network analysis was performed with the Network software (version 10.2), which is available as shareware on the Fluxus Technology website (https://www.fluxus-engineering.com/) [49]. Mutations of single nucleotide variations (SNVs) in the coding region sequences of ASFV were used. The data was run by median joining network algorithm and Steiner tree algorithm, respectively, with the epsilon parameter set to zero. We also performed an exploratory run by setting the epsilon parameter to 10. It is revealed that both algorithms yielded an identical structure of network.

2.4. The Analysis of Evolutionary Rates and Estimation of Time to the Most Recent Common Ancestry. Temporal dynamics of ASFV were analyzed with time-resolved phylogenies using the Bayesian MCMC method in the BEAST package 1.8.2. The datasets were analyzed using the GTR + G substitution model under a lognormal uncorrelated relaxed clock model with a coalescent constant size model. The MCMC analyses were run in duplicate for 2×10^8 generations, sampling every 2000 generations. Convergence was assessed from the effective sample size (ESS) with a 10% burn-in using Tracer v1.6. ESS values above 200 were accepted.

2.5. Phylogeographic Analysis. The spatial transmission patterns of ASFV were reconstructed using a phylogeographic analysis in BEAST 1.8.2. The geographic location and sampling time of all ASFV genome sequences were obtained as described previously. The best-supported pairwise diffusions were identified using SPREAD3 v0.9.6 [50]. Bayesian Stochastic Search Variable Selection (BSSVS) was used to estimate the significance of pairwise switches between trait states. Migration pathways with a Bayes factor greater than 5 and a mean posterior value greater than 0.5 were considered to be important.

2.6. Detection of Selection Pressures. Positive selection was detected using Datamonkey (https://www.datamonkey.org/). Single Likelihood Ancestor Counting (SLAC), Fixed Effects Likelihood (FEL), Mixed Effects Model of Evolution (MEME), and Fast Unconstrained Bayesian AppRoximation (FUBAR) were used to infer selection [51]. Sites identified by at least two algorithms were considered to be undergoing conservative positive selection.

2.7. Protein Structure Prediction. The three-dimensional structure of protein I267L was modelled using the Alphafold2 program [52] and PyMOL 1.4.1 [53].

3. Results

3.1. Genome Sequences Diversity of Genotype II ASFV in China during 2018 to 2020. We obtained the full-length genome sequence of the ASFV strain for the first outbreak of ASF in China in Shenyang, Liaoning province on August 3rd, 2018. The spleen sample collected from an ASF-confirmed pig on the infected farm was used for next-generation sequencing. The final assemble of the ASFV China/LN/2018/1 genome was achieved from a reference-based alignment consisting of 149,069 mapped reads (0.27% of 54,956,232 total reads) with an average depth of 85 reads per nt. The ASFV China/LN/2018/1 whole-genome sequence has a length of 189,397 bp, not including terminal inverted repeats and cross links. The sequence data has been submitted to the GenBank database under accession number OP856591.

The ASFV China/LN/2018/1 genome was added to the other 9 genotype II ASFV genomes in China in the curated dataset described previously, making a total of 10 ASFV genome sequences collected in China from August 3rd, 2018 to October 1st, 2020. The genomic diversity of the genotype II ASFV genomes in China was further investigated. Comparing to the ASFV China/LN/2018/1 genome, mutations in each genome sequence were checked manually. A total of 128 variable sites were found (Figure 1). Further investigation of these sites revealed 2 structural variations (in

TABLE 1: List of genotype II ASFV genome sequences used in this study.

NCBI Acc nos.	Strains	Hosts	Lengths	Collection dates	Countries	Genotypes	Ref
LR536725.1	Belgium 2018/1	Wild boar	190,599	2018/9	Belgium	II	[28]
MK543947.1	Belgium/Etalle/wb/2018	Wild boar	190,202	2018/9/10	Belgium	II	[29]
	China/LN/2018/1	Pig	189,397	2018/8/3	China	II	This study
MK128995.1	China/2018/AnhuiXCGQ	Pig	189,393	2018/9/2	China	II	[27]
MK333180.1	Pig/HLJ/2018	Pig	189,404	2018/9/5	China	II	[30]
MK645909.1	ASFV-wbBS01	Wild boar	189,394	2018/11/1	China	II	NA
MK940252.1	CN/2019/InnerMongolia-AES01	Pig	189,403	2019/2/19	China	II	NA
MN172368.1	ASFV/pig/China/CAS19-01/2019	Pig	189,405	2019/1/2	China	II	[31]
MN393476.1	Wuhan 2019-1	Pig	190,576	2019/8/19	China	II	NA
MT496893.1	GZ201801	Pig	189,393	2018/12/22	China	II	NA
MW521382.1	HuB20	Pig	188,643	2020/10/1	China	II	NA
MW656282.1	Pig/Heilongjiang/HRB1/2020	Pig	189,355	2020/9/12	China	II	[32]
LR722600.1	CzechRepublic 2017/1	Wild boar	190,594	2017/6	Czech	II	[33]
LS478113.1	Estonia 2014	Wild boar	182,446	2014/9	Estonia	II	[34]
FR682468.2	Georgia 2007/1	Pig	190,584	2007/1/4	Georgia	II	[35]
LR899193.1	Germany 2020/1	Wild boar	190,592	2020/9/10	Germany	II	NA
MN715134.1	ASFV_HU_2018	Wild boar	190,601	2018/4/24	Hungary	II	[36]
MK628478.1	ASFV/LT14/1490	Wild boar	189,399	2014/1	Lithuania	II	[37]
MW856068.1	MAL/19/Karonga	Pig	183,325	2019	Malawi	II	[38]
LR722599.1	Moldova 2017/1	Pig	190,598	2017	Moldova	II	[20]
MG939583.1	Pol16_20186_o7	Pig	189,401	2016	Poland	II	[39]
MG939584.1	Pol16_20538_o9	Pig	189,399	2016	Poland	II	[39]
MG939585.1	Pol16_20540_o10	Pig	189,405	2016	Poland	II	[39]
MG939586.1	Pol16_29413_o23	Pig	189,393	2016	Poland	II	[39]
MG939587.1	Pol17_03029_C201	Wild boar	189,405	2017	Poland	II	[39]
MG939588.1	Pol17_04461_C210	Wild boar	189,401	2017	Poland	II	[39]
MG939589.1	Pol17_05838_C220	Wild boar	189,393	2017	Poland	II	[39]
MH681419.1	ASFV/POL/2015/Podlaskie	Wild boar	189,394	2015/2	Poland	II	[40]
MT847620.1	Pol17_55892_C754	Pig	189,414	2017	Poland	II	[14]
MT847621.1	Pol18_28298_O111	Pig	189,409	2018	Poland	II	[14]
MT847622.1	Pol17_31177_O81	Pig	189,422	2017	Poland	II	[14]
MT847623.2	Pol19_53050_C1959/19	Pig	189,413	2019	Poland	II	[14]
KJ747406.1	Kashino 04/13	Wild boar	189,387	2013/3/1	Russia	II	NA
KP843857.1	Odintsovo_02/14	Wild boar	189,333	2014/2	Russia	II	NA
MT459800.1	ASFV/Kabardino-Balkaria 19/WB-964	Wild boar	189,252	2019/3/26	Russia	II	NA
MW306190.1	ASFV/Amur 19/WB-6905	Wild boar	189,248	2019/8/29	Russia	II	[41]
MW306191.1	ASFV/Primorsky 19/WB-6723	Wild boar	189,256	2019/8/28	Russia	II	[41]
MW306192.1	ASFV/Ulyanovsk 19/WB-5699	Wild boar	189,263	2019/8/21	Russia	II	[41]
MT748042.1	ASFV/Korea/pig/PaJu1/2019	Pig	190,597	2019/9/16	South Korea	II	NA
LR813622.1	Tanzania/Rukwa/2017/1	Pig	183,186	2017	Tanzania	II	[42]
MW396979.1	ASFV/Timor-leste/2019/1	Pig	192,237	2019/9/27	Timor-Leste	II	[43]
MN194591.1	ASFV/Kyiv/2016/131	Pig	191,911	2016/4/11	Ukraine	II	[44]

MW656282.1/Pig/Heilongjiang/HRB1/2020), 2 Tandem repeat sequences variations, 99 single nucleotide variations (SNVs), 19 homopolymer variations (1-2 bases of insertion and deletion in the homopolymeric tracts), and 6 highly variable poly G or poly C tract variations.

3.2. Phylogenetic Analysis of Genotype II ASFV and Estimation of Time to Most Recent Common Ancestry. The ASFV China/LN/2018/1 genome was added to other genotype II ASFV genomes in the curated dataset, making a total of 42 ASFV genome sequences for further analysis. Signatures of recombination were detected in 3 sequences (LS478113.1/Estonia 2014, MG939584.1/Pol16_20538_o9, MW656282.1/Pig/Heilongjiang/HRB1/2020) by at least five of the seven different detection algorithms ($P < 0.01$). These 3 genome sequences were not included in the further phylogenetic analysis. Considering the uncertainty of the terminal inverted repeat sequence on both genomic ends, the coding region sequence corresponding to the nucleotide site 2113-190,166 of the reference genome Georgia 2007/1 was used for phylogenetic analysis of the final dataset of 39 genotype II ASFV strains. These genomes were collected from domestic pigs (23 genomes) and wild boars (16) from 15 countries including Belgium (2), China (9), the Czech (1), Estonia (1), Georgia (1), Germany (1), Hungary (1), Lithuania (1), Malawi (1), Moldova (1), Poland (10), Russia (6), South Korea (1), Tanzania (1), Timor-Leste (1), and Ukraine (1) during 2007 to 2020 (Table 1).

The maximum clade credibility (MCC) phylogenetic dendrogram of ASFV was constructed by the Bayesian-based coalescent approach. The MCC tree showed 3 clades with
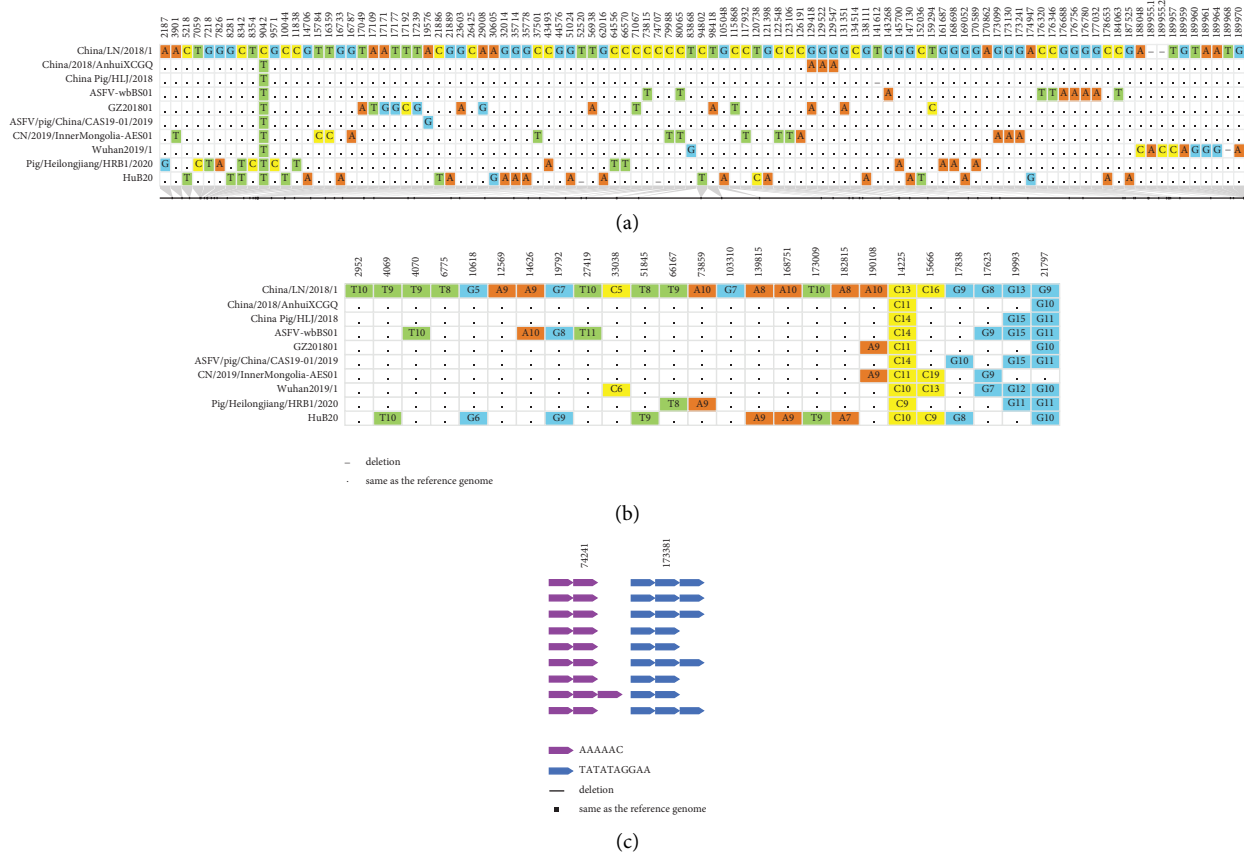
(a)

(b)

(c)

Figure 1: Genome sequence diversity of genotype II ASFV in China from August 3rd to October 1st, 2020. Nucleotide position in the reference genome Georgia 2007/1 was shown. (a) The number and detail of mutations at SNPs and indel. (b) The number and detail of mutations of homopolymer and variable polyG/C sequence. (c) Sequence variations of the tandem repeat sequences (TRS). Each arrow with different color represents a type of TRS. The sequence of the repeat unit in each TRS was also shown.

temporal and spatial correlation (Figure 2). Clade A included two strains collected from Africa. Clade B included 7 strains which were collected from Georgia, Russia (3 strains), Lithuania, and Poland from 2007 to 2019. Clade C included strains collected from Europe and Aisa from 2014 to 2020, grouping into one singleton (subcalde C1) and two subclades (C2 and C3). Subcalde C1 was a singleton of Odintsovo_02/14, a strain collected from Russia in 2014. Subcalde C2 included 13 strains: 11 strains collected from Poland during 2016 to 2019, one strain from Germany in 2020, and one strain from Ukraine in 2016. Subclade C3 was composed of 2 clusters, cluster C3.1 and Cluster C3.2. Cluster C3.1 included 4 strains from the Czech Republic, Moldova, and Belgium from 2017 to 2018. Cluster C3.2 was composed of one strain from Hungary in 2018, 9 strains collected from China from 2018 to 2020, one strain from South Korea in 2019, one strain from Timor-Leste in 2019 and two strains from eastern Russia in 2019. The tree topology indicated that genotype II ASFV has evolved into different genetic groups since being introduced into Europe and then Asia.

The estimates of evolutionary rates and time to most recent common ancestry (TMRCA) of genotype II ASFV were calculated by the Bayesian-based coalescent approach. The mean evolutionary rate of the ASFV GII genome was estimated at 1.312×10^{-5} nucleotide substitutions per site per year (95% HPD, 7.7432×10^{-6}–1.8733×10^{-5}). The divergence times and the countries of origins of genotype II ASFV were estimated by using the Bayesian phylogenetic analysis. For clade B, which is expanding in Europe and Asia, the median TMRCA was estimated to be August 2005 (95% HPD: April 2000–January 2007), and its country of origin was inferred as Georgia, with strong support (100% posterior probability) (Figure 3).

3.3. Subgroup Divergence of Genotype II ASFV in China and Estimation of Time to Most Recent Common Ancestry. Sequences collected from different provinces in China during 2018 to 2020 were widely interspersed, with no particular grouping according to distinct temporal and spatial correlation. Sequences from other neighboring countries, including South Korea, Timor-Leste, and Eastern Russia, were embedded within the cluster of 9 genomes obtained from China. There was strong support for monophyletic grouping of all the sequences from Asia and a strain from Hungary which was sampled on April 24th, 2018 (100% posterior probability). Similar phylogenetic relationships for these genomes were also seen using maximum likelihood (ML) phylogenetic reconstruction (Figure S1). There was also strong support for grouping
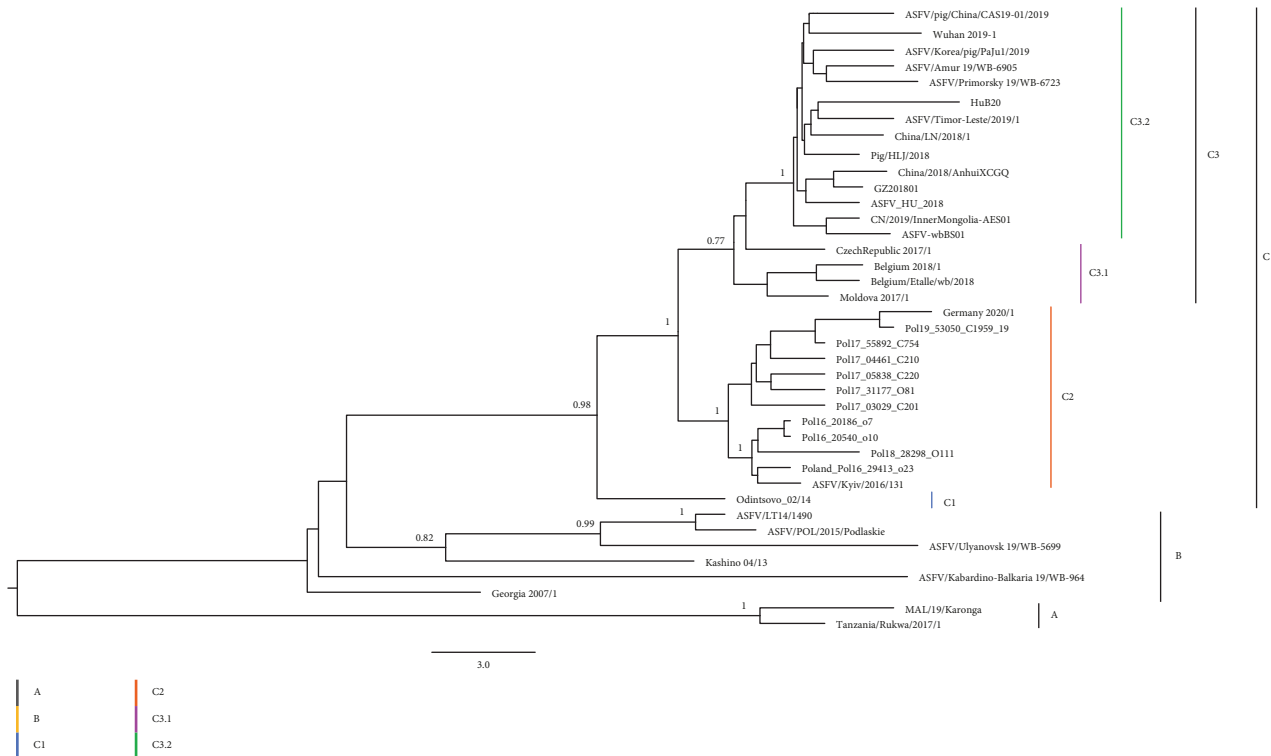
FIGURE 2: Maximum-clade credibility phylogenetic tree of genotype II ASFV genome sequences. The tree is midpoint-rooted. The scale bar is given in numbers of substitutions per site. The posterior probabilities higher than 0.7 are shown at the corresponding nodes. Cluster assignment is shown by color: black for clade A yellow for clade B blue for clade C1, orange for clade C2, purple for clade C3.1, and green for clade C3.2.

sequences from Asia, Hungary, Belgium, Czech, Moldova, Germany, Poland, and Ukraine (100% posterior probability). It is suggested that all the ASFV genotype II genomes in China share a single origin and have a common ancestor with those from Central or Eastern Europe.

TMRCA estimation showed that the emergence of the ancestor strain of ASFV in Asia dated back to April 2017 (95% HPD, February 2016; March 2018) (Figure 3). The ancestor at the node linking all the Asian ASFV sequences with the most closely related sequence in Hungary was dated to October 2016. It is consistent with the 5-month time difference between the first outbreak in Hungary in April 2018 and the first outbreak in China in August 2018. The median TMRCA of the genomes in the other countries in Asia was July 2018 (95% HPD, June 2017; April 2019). It is consistent with the 13-month time difference between the first outbreak in China in August 2018 and the first outbreaks in South Korea and Timer-Leste in September 2019.

3.4. Phylogeographic Analysis of Genotype II ASFV.

To further investigate the origin of genotype II ASFV in China, a discrete phylogeographic analysis was conducted. The Bayesian stochastic search variable selection (BSSVS) method was used to infer the geographic spread (Figure 4). It was revealed that ASFV spread from Central or Eastern Europe into China was strongly supported with a Bayes factor (BF) of 64.86 and an associated posterior probability (PP) of 0.82. Well-supported rates were also identified for

ASFV spread from Poland to Ukraine (BF = 81.39; PP = 0.85), from Poland to Germany (BF = 38.44; PP = 0.73), and from Central or Eastern Europe to Belgium (BF = 40.14; PP = 0.74). It is suggested that phylogeographic analysis based on high-quality genome sequences could be used for tracing the transmission of ASFV during an epidemic.

3.5. Identification of Anchor Mutation in Genotype II ASFV Genomes.

To trace the occurrence and possible fixation of mutations during evolution of genotype II ASFV, we investigated the SNVs in ASFV genotype II genomes. All the SNVs were extracted and listed in a matrix (Figure 5(a)). Eight SNVs were identified as anchor mutations, which were fixed in all the sequences in corresponding clusters. The gene location and codon position of each anchor mutation was also displayed (Figure 5(b)). The distribution of amino acid changes of each anchor mutation was also investigated (Figure 5(c)). Sequences in clade C were distinguished from those in clade B by three nonsynonymous mutations: C7059T (G590A/Trp197∗ in MGF110-1L), A44576G (A967G/Lys323Gluin MGF505-9R) and T134514C (A1241G/Asn414Ser in NP419L). Sequences in subclade C2 and C3 shared one nonsynonymous mutation, T170862A (A583T/Ile195Phe in I267L). Subclade C2, including genomes from Poland, Ukraine, and Germany, was defined by one synonymous and two nonsynonymous mutations: G10668A (C60T/Ser20 in MGF110-7L), G39306A (G988A/Val330Ile in MGF505-5R), and C66152A (C434A/Ser145Tyr
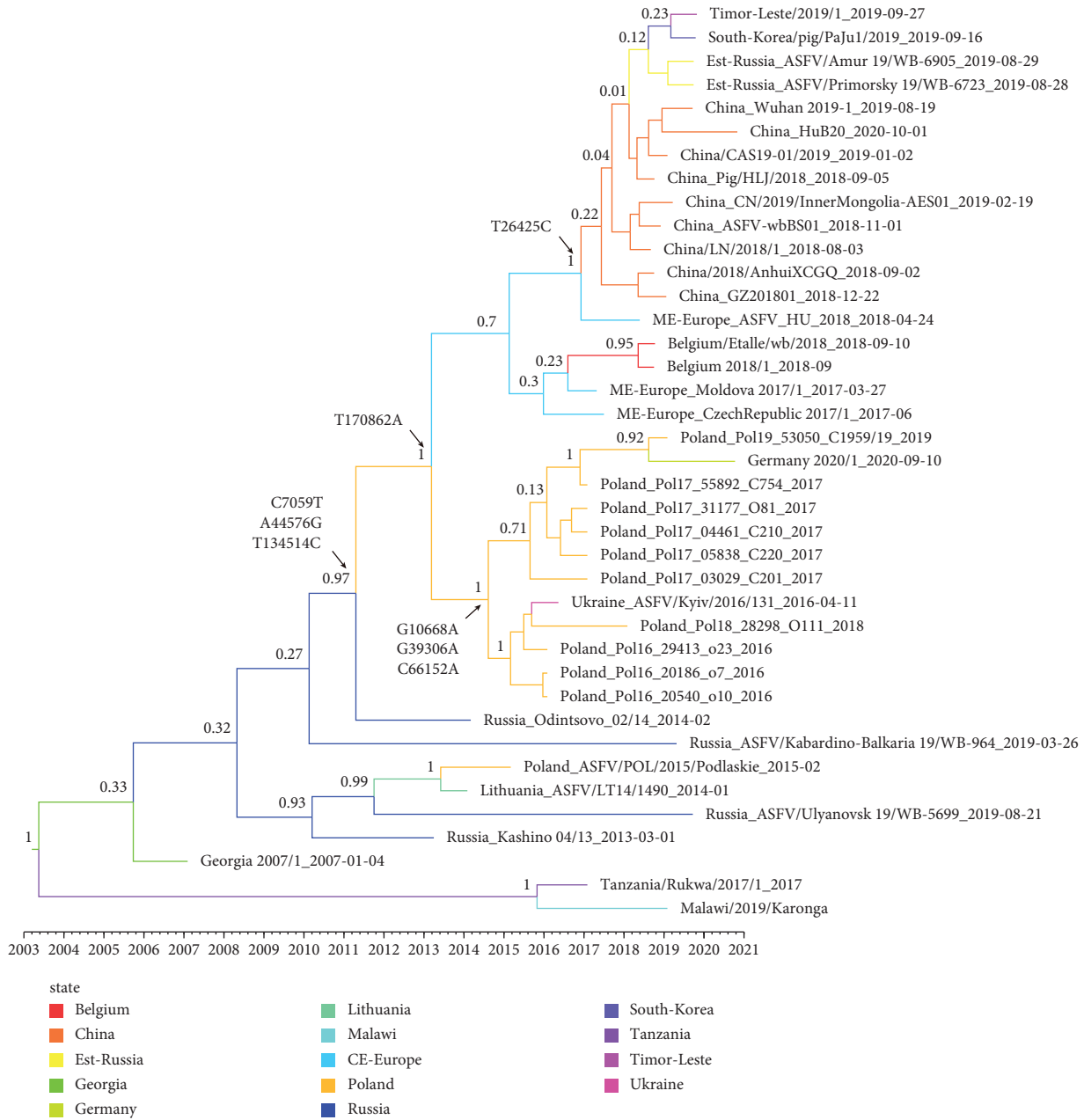
FIGURE 3: Time-revolved phylogenetic tree based on genotype II ASFV genome sequences using Bayesian MCMC analysis. The tree was estimated using the GTR + G substitution model under a lognormal uncorrelated relaxed clock model with a coalescent constant size model. Branches are color-coded by country or region. The scale bar indicates the time in year. The posterior probability is shown at each node.

in K145R). A nonsynonymous mutation in MGF360-10L (A986G/Asn329Ser), T26425C, characterized cluster C3.2 containing genomes from Hungary, eastern Russia, China, and other Asian countries, as shown in Figure 3.

A median joining phylogenetic network was constructed based on the SNVs for exploring evolutionary relationships among ASFV genotype II genomes. Three clades with temporal and spatial correlation were identified (Figure 6). The putative ancestor of clade C was derived from that of clade B by 4 anchor mutations. The putative ancestor of subclade C2 was derived from the common node by 3 anchor mutations,

and cluster C3.2 one anchor mutation. Mutational branches in cluster C3.2 formed a star-like radiation topology, indicating a single origin and then expansion of strains in this cluster. Two sequences were identified as the ancestral node in cluster C3.2, one from Heilongjiang province in China on September 5th, 2018 and the other from South Korea on September 16, 2019. A total of 10 mutational branches radiated from the ancestral node of cluster C3.2, representing 9 genomes collected from China during 2018 to 2020, one genome from Hungary, one genome from Timor-Leste, and two genomes from eastern Russia.
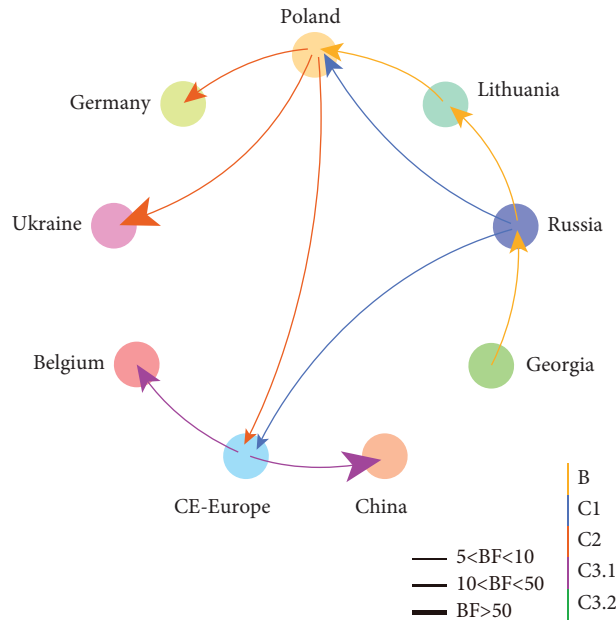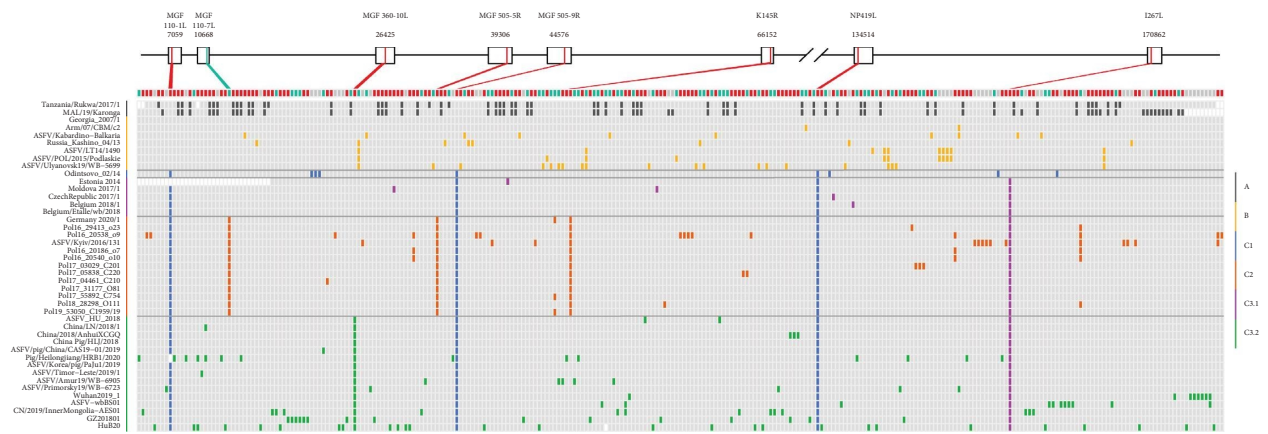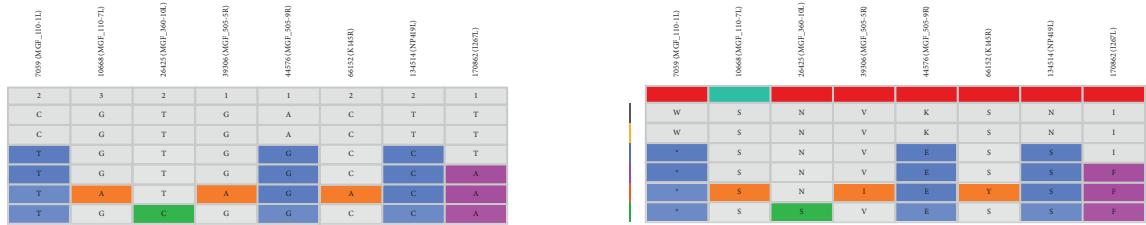
FIGURE 4: Spatial diffusion pathway of genotype II ASFV in Europe and Asia. Arrows indicate the direction of the switch; only switches supported by BF values >5 are shown. Line thickness is proportional to the dispersal route significance level. The line is colored according to the cluster assignment.

FIGURE 5: Pattern of genotype II ASFV single nucleotide polymorphisms (SNPs) distribution. (a) Distribution of SNPs in genotype II ASFV genomes. A total of 277 SNPs show a pattern of genotype II ASFV mutation accumulation, showing one mutation site per column. The strain name is shown at the left, indicating one strain per row. The grey blocks indicate sequence identity with the reference genome Georgia 2007/1 sequence. The top row shows the type of mutation (dark grey, intergenic; green, synonymous; red, nonsynonymous), with the gene location indicated above. Cluster assignment is shown at the left by color. (b) Distribution of 8 anchor SNPs in different cluster of genotype II ASFV genomes. The top row shows the codon position of SNPs with the gene location indicated above. (c) Distribution of amino acid changes of anchor SNPs in different cluster of genotype II ASFV genomes. The top row shows the type of mutation (green, synonymous; red, nonsynonymous), with the gene location indicated above.

FIGURE 6: Phylogenetic network of genotype II ASFV genomes. The network was constructed from the SNPs in the coding region of the genome sequences. The circle nodes represent ASFV strains, which are proportional to the number of taxa. Each notch on the links represents a mutation event. Red dots indicate putative ancestor nodes. Cluster assignment is shown by color. The median-joining network algorithm and the Steiner algorithm were used.

3.6. Measure of Selection Pressures in Genotype II ASFV Genomes. Site-specific selection pressures were measured as the ratio of nonsynonymous (dN) to synonymous (dS) nucleotide substitutions per site to infer sites subjected to positive selection in the coding sequences of MGF110-1L, MGF110-7L, MGF360-10L, MGF505-5R, MGF505-9R, K145R, NP419L, and I267L. Using a posterior probability of 0.88 for FUBAR, 12 sites were identified as under positive selection. Using the default threshold of significance of $P < 0.1$, two sites were found by SLAC method. The threshold was reevaluated manually into $P < 0.05$ to check for the accuracy of the default parameters. The same result was achieved by the SLAC method. Amino acid 414 in

NP419L and amino acid 195 in I267L were identified as under positive selection by both the FUBAR and SLAC methods (Table 2).

The three-dimensional structure of I267L was modelled using the Alphafold2 program (Figure 7). The function of I267L was predicted to be binding with nucleic acids.

4. Discussion

To our knowledge, this is the first in-depth analysis on genotype II ASFV genomic evolution. Our estimation of the evolutionary rate of 39 ASFV genotype II genomes collected in Eurasia and Africa between 2007 and 2020 is 1.312×10^{-5}

TABLE 2: List of amino acid positions identified as evolving under positive selection.

Protein	MEME	FUBAR	FEL	SLAC
I267L	—	195 (0.84) 129 (0.83) 229 (0.81)	—	195
K145R	—	145 (0.87)	—	—
MGF110_1L	—	—	—	—
MGF110_7L	—	39 (0.88)	—	—
MGF360_10L	—	329 (0.87)	—	—
MGF505_5R	—	330 (0.86) 338 (0.86) 462 (0.87)	—	—
MGF505_9R	—	323 (0.87) 465 (0.87)	—	—
NP419L	—	414 (0.87)	—	414

The default threshold of significance ($P < 0.1$) was used for MEME, SLAC, and FEL. Posterior probability of 0.81~0.88 was used for FUBAR. Posterior probability was given in parentheses. "—" indicates no positively selected sites was identified.
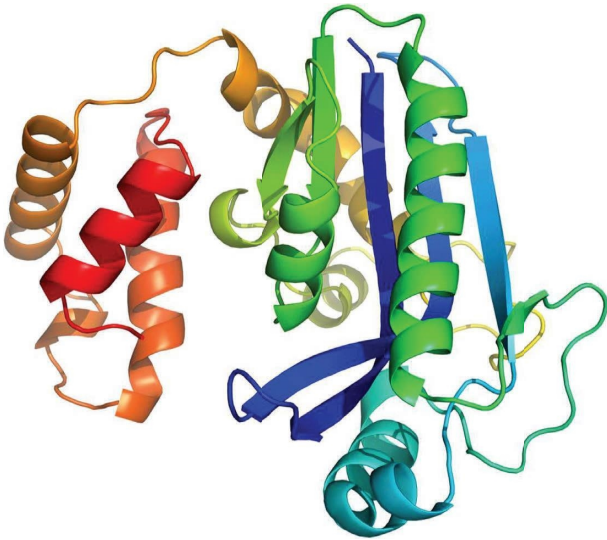

FIGURE 7: The three-dimensional structure modelling of I267L by the Alphafold2 program.

nucleotide substitutions per site per year (95% HPD, 7.7432×10^{-6}; 1.8733×10^{-5}), which is around 3.9 times faster than the value calculated for ASFV genotype I genomes collected in Sardinia [24, 26]. The widespread presence of ASFV in Eurasia, the large population of infected domestic pigs and wild boars, and the inadequate biosecurity of the pig industry in Asia may have sped the expansion and evolution of the virus.

The presence of a single putative common ancestor for all ASFV genome sequences in China provides clear evidence for a single introduction of the virus. According to the analysis of currently available genome sequences, ASFV in China shares the same origin with those in Hungary, Belgium, the Czech, Moldova, Germany, Poland, and Ukraine. However, it is not possible for us to ascribe the countries of origin for genotype II ASFV in China to a country or an area in a country. There is a data gap for those countries which were not represented by ASFV genome sequences or had a limited number of ASFV genome sequences published.

Between 2007 and August 2018, a total of 11 European countries reported ASFV outbreaks. Among these countries, genome sequences from Romania and Belarus are not publicly available. Genome sequences of ASFV in Russia, especially the Far East, Ukraine, Lithuania, Latvia, and Estonia between 2014 and 2018 are limited. Previous studies demonstrated that the sequencing of more ASFV samples dramatically improved the resolution of ASFV phylogenetic structure and permitted finer molecular dating [24]. More genome sequences of ASFV strains circulating in European countries will likely help improve our knowledge on the molecular epidemiology of genotype II ASFV.

In this study, we construct the transmission pathway of the spread of ASFV genotype II basing on the genome-wide SNVs. The putative-ancestor-virus-centered radial shape of the Asian cluster suggests that samples collected from different provinces in China and other Asian countries may come from the same virus origin. According to the field epidemiology investigation, long-distance transmission of animal or animal products played an important role in the spread of ASFV in Asia [54, 55]. Not only it is true that more than half of the world's domestic pigs are raised in Asia, but also it has the highest population density of farmed pigs in the world. The pig production in Asia, in general, was not well industrialized had a inadequate biosecurity [9]. It explains how the ASF virus rapidly swept Asia in a short time after the introduction and produced a variety of mutational variants.

The anchor SNVs which defined clusters of ASFV GII in Europe and Asia were located in MGF110-1L, MGF110-7L, MGF360-10L, MGF505-5R, MGF505-9R, K145R, NP419L, and I267L. K145R and MGF505-5R have been used as molecular markers to differentiate ASF viruses circulating in Poland [14]. MGF-360-10L, MGF-505-9R, and I267L have also been used to differentiate ASF viruses collected in Russia [41]. However, this is the first time to identify 8 anchor mutations for differentiating genotype II ASFV. These anchor mutations can serve as candidate molecular markers for genetic characterization of genotype II ASF viruses circulating in Europe and Asia and help trace the origin and transmission of this virus.

TABLE 3: List of function of the proteins harboring these anchor mutations.

Genes	Functions	References
MGF 110-1L	Nonessential for in vitro propagation and nonaffect viral virulence	[56]
MGF 110-7L	Induces host cell translation suppression and stress granule formation by activating the PERK/PKR-eIF2 alpha pathway	[57]
MGF 360-10L	Implicated in the modulation of the type I interferon (IFN) response and associated to the virulence of the virus	[58]
MGF 505-5R	NA	NA
MGF 505-9R	NA	NA
K145R	A highly abundant viral protein accumulated diffusely in the cytoplasm of infected cells but nonessential for in vitro propagation of ASFV	[59]
NP419L	DNA ligase	[60]
I267L	An important virulence factor by inhibiting RNA polymerase III-RIG-I-mediated innate immunity	[61]

The function of some of the genes harboring these anchor mutations has been determined. As shown in Table 3, 3 out of the 8 genes have been previously identified as involving immune modulation of host responses. MGF 110-7L induces host cell translation suppression and stress granule formation by activating the PERK/PKR-eIF2 alpha pathway [57]. MGF360-10L has been implicated in the modulation of the type I interferon (IFN) response and associated with the virulence of the virus [58]. I267L acts as an important virulence factor by inhibiting RNA polymerase III-RIG-I-mediated innate immunity [61]. Although the function of MGF505-5R and MGF505-9R remains unknown, it is well known that MGF530/505 genes have roles in suppressing the induction of type I IFN [58]. Previous studies revealed that pK145R is a highly abundant viral protein that accumulates diffusely in the cytoplasm of infected cells but is not essential for the in vitro propagation of ASFV [59]. MGF110-1L is also nonessential for in vitro propagation [56]. The mutations in these genes might be necessary for the adaptation of ASFV to the large population of domestic pigs in Europe and Asia.

ASFV is replicated and assembled in swine macrophage cells, which are rich in free oxygen radicals and cause constant damage to the virus genome. To efficiently overcome these DNA damages, ASFV virus has evolved its own repair system, including repair DNA polymerase (AsfvPolX) and ligase (NP419L, also known as AsfvLIG). AsfvLIG is one of the most error-prone ligases identified to date, which plays important roles in the strategic mutagenesis and genotype formation of ASFV. AsfvLIG contains one N-terminal region, one adenylation domain in the center, and one OB-fold domain (OB) at the C-terminus [60]. The site residue subjected to positive selection (Asn414Ser) locates in the OB-fold domain, in close proximity to two active site residues (Leu402 and Gln403 of the OB domain) which are crucial for the catalytic efficiency of AsfvLIG. Therefore, it is most likely that the positive selection of DNA ligase may facilitate the mutagenesis of the ASFV genome and further adaption of the virus to pandemic spread in Eurasia.

Our study suggests that the structure of the evolution topology for genotype II ASFV strains is determined by SNVs. The role of other mutation events in ASFV evolution needs to be further investigated. ASFV strains with structural variations including recombination, fragment deletion, or insertion are of specific interest to researchers because of the potential for virulence attenuation. Previous research on genome sequence comparisons of different ASFV genotypes suggested that recombination might play an important role in ASFV evolution [62–64]. According to our analysis of genotype II ASFV strains, two structural variations are identified as clade A specific. The role of structural variations in the evolution of genotype II ASFV needs to be further investigated.

5. Conclusion

ASFV is presently causing a global animal health emergency because of its devastating effects on pig populations. The lack of vaccines or effective therapies against ASFV urges the need to address the existing relevant gaps in knowledge of the virus genome evolution. Genomic epidemiology has played an irreplaceable role in investigating the origin, tracing the spread, and monitoring the evolution of ASFV. It will play a key role in helping contain the ASFV pandemic's spread.

Data Availability

Disclosure

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Authors' Contributions

Yong Zhang, Qinghua Wang and Zhongyi Zhu contributed equally to this work.

Acknowledgments

Supplementary Materials

Figure S1 to this article can be found as supplementary material. (*Supplementary Materials*)

References

[1] L. K. Dixon, H. Sun, and H. Roberts, "African swine fever," *Antiviral Research*, vol. 165, pp. 34–41, 2019.

[2] S. Ge, J. Li, X. Fan et al., "Molecular characterization of african swine fever virus, China, 2018," *Emerging Infectious Diseases*, vol. 24, no. 11, pp. 2131–2133, 2018.

[3] D. Kolbasov, I. Titov, S. Tsybanov, A. Gogin, and A. Malogolovkin, "African swine fever virus, siberia, Russia, 2017," *Emerging Infectious Diseases*, vol. 24, no. 4, pp. 796–798, 2018.

[4] V. P. Le, D. G. Jeong, S. W. Yoon et al., "Outbreak of african swine fever, vietnam, 2019," *Emerging Infectious Diseases*, vol. 25, no. 7, pp. 1433–1435, 2019.

[5] R. J. Rowlands, V. Michaud, L. Heath et al., "African swine fever virus isolate, Georgia, 2007," *Emerging Infectious Diseases*, vol. 14, no. 12, pp. 1870–1874, 2008.

[6] K. Smietanka, G. Wozniakowski, E. Kozak et al., "African swine fever epidemic, Poland, 2014-2015," *Emerging Infectious Diseases*, vol. 22, no. 7, pp. 1201–1207, 2016.

[7] S. M. Lakin, V. K. O'Donnell, L. Xu et al., "Whole genome sequencing and molecular epidemiology of the 2021 African swine fever virus outbreak in the Dominican Republic," *Transboundary and Emerging Diseases*, vol. 2, 2022.

[8] C. P. McCarthy, M. Vaduganathan, E. Solomon et al., "Running thin: implications of a heparin shortage," *The Lancet*, vol. 395, no. 10223, pp. 534–536, 2020.

[9] Y. Woonwong, D. Do Tien, and R. Thanawongnuwech, "The future of the pig industry after the introduction of african swine fever into Asia," *Animal Frontiers*, vol. 10, no. 4, pp. 30–37, 2020.

[10] M. Moniruzzaman, C. A. Martinez-Gutierrez, A. R. Weinheimer, and F. O. Aylward, "Dynamic genome evolution and complex virocell metabolism of globally-distributed giant viruses," *Nature Communications*, vol. 11, no. 1, p. 1710, 2020.

[11] S. Blome, C. Gabriel, and M. Beer, "Pathogenesis of African swine fever in domestic pigs and European wild boar," *Virus Research*, vol. 173, no. 1, pp. 122–130, 2013.

[12] D. Beltrán-Alcrud, M. Aria, C. Gallard, S. A. Kramer, and M.-L. Penrit, *African Swine Fever: Detection and Diagnosis – A Manual for Veterinarians*, Food and Agriculture Organization of the United Nations (FAO), Rome, Italy, 2017.

[13] C. Gallardo, J. Fernandez-Pinero, V. Pelayo et al., "Genetic variation among African swine fever genotype II viruses, eastern and central Europe," *Emerging Infectious Diseases*, vol. 20, no. 9, pp. 1544–1547, 2014.

[14] N. Mazur-Panasiuk, M. Walczak, M. Juszkiewicz, and G. Wozniakowski, "The spillover of african swine fever in western Poland revealed its estimated origin on the basis of O174L, K145R, MGF 505-5R and IGR I73R/I329L genomic sequences," *Viruses*, vol. 12, no. 10, p. 1094, 2020.

[15] H. T. T. Tran, A. D. Truong, A. K. Dang et al., "Novel method for sub-grouping of genotype II African swine fever viruses based on the intergenic region between the A179L and A137R genes," *Veterinary Medical Science*, vol. 8, no. 2, pp. 607–609, 2022.

[16] M. W. Carroll, D. A. Matthews, J. A. Hiscox et al., "Temporal and spatial analysis of the 2014-2015 Ebola virus outbreak in West Africa," *Nature*, vol. 524, no. 7563, pp. 97–101, 2015.

[17] N. R. Faria, R. Azevedo, M. U. G. Kraemer et al., "Zika virus in the Americas: early epidemiological and genetic findings," *Science*, vol. 352, no. 6283, pp. 345–349, 2016.

[18] R. Lu, X. Zhao, J. Li et al., "Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding," *The Lancet*, vol. 395, no. 10224, pp. 565–574, 2020.

[19] M. A. Martin, D. VanInsberghe, and K. Koelle, "Insights from SARS-CoV-2 sequences," *Science*, vol. 371, no. 6528, pp. 466-467, 2021.

[20] J. H. Forth, L. F. Forth, J. King et al., "A deep-sequencing workflow for the fast and efficient generation of high-quality african swine fever virus whole-genome sequences," *Viruses*, vol. 11, no. 9, p. 846, 2019.

[21] L. Jia, M. Jiang, K. Wu et al., "Nanopore sequencing of African swine fever virus," *Science China Life Sciences*, vol. 63, no. 1, pp. 160–164, 2020.

[22] J. J. Esposito, S. A. Sammons, A. M. Frace et al., "Genome sequence diversity and clues to the evolution of variola (smallpox) virus," *Science*, vol. 313, no. 5788, pp. 807–812, 2006.

[23] J. Bao, Y. Zhang, C. Shi et al., "Genome-wide diversity analysis of african swine fever virus based on a curated dataset," *Animals*, vol. 12, 2022.

[24] M. S. Fiori, D. Sanna, F. Scarpa et al., "A deeper insight into evolutionary patterns and phylogenetic history of ASFV epidemics in Sardinia (Italy) through extensive genomic sequencing," *Viruses*, vol. 13, 2021.

[25] Z. J. Shen, H. Jia, C. D. Xie et al., "Bayesian phylodynamic analysis reveals the dispersal patterns of african swine fever virus," *Viruses*, vol. 14, 2022.

[26] C. Torresi, M. Fiori, L. Bertolotti et al., "The evolution of African swine fever virus in Sardinia (1978-2014) as revealed by whole-genome sequencing and comparative analysis," *Transboundary and Emerging Diseases*, vol. 40, Article ID tbed.13540, 2020.

[27] J. Bao, Q. Wang, P. Lin et al., "Genome comparison of African swine fever virus China/2018/AnhuiXCGQ strain and related European p72 Genotype II strains," *Transboundary and Emerging Diseases*, vol. 66, no. 3, pp. 1167–1176, 2019.

[28] J. H. Forth, M. Tignon, A. B. Cay et al., "Comparative analysis of whole-genome sequence of african swine fever virus Belgium 2018/1," *Emerging Infectious Diseases*, vol. 25, no. 6, pp. 1249–1252, 2019.

[29] G. Gilliaux, M. Garigliany, A. Licoppe et al., "Newly emerged African swine fever virus strain Belgium/Etalle/wb/2018: complete genomic sequence and comparative analysis with reference p72 genotype II strains," *Transboundary and Emerging Diseases*, vol. 66, no. 6, pp. 2566–2591, 2019.

[30] X. Wen, X. He, X. Zhang et al., "Genome sequences derived from pig and dried blood pig feed samples provide important insights into the transmission of African swine fever virus in China in 2018," *Emerging Microbes & Infections*, vol. 8, no. 1, pp. 303–306, 2019.

[31] L. Jia, J. Chen, H. Liu et al., "Potential m6A and m5C methylations within the genome of A Chinese african swine

fever virus strain," *Virologica Sinica*, vol. 36, no. 2, pp. 321–324, 2020.

[32] E. Sun, Z. Zhang, Z. Wang et al., "Emergence and prevalence of naturally occurring lower virulent African swine fever viruses in domestic pigs in China in 2020," *Science China Life Sciences*, vol. 64, no. 5, pp. 752–765, 2021.

[33] J. H. Forth, L. F. Forth, P. Vaclavek et al., "Whole-genome sequence of an african swine fever virus isolate from the Czech republic," *Microbiol Resour Announc*, vol. 9, no. 44, Article ID e00948-20, 2020.

[34] L. Zani, J. H. Forth, L. Forth et al., "Deletion at the 5'-end of Estonian ASFV strains associated with an attenuated phenotype," *Scientific Reports*, vol. 8, no. 1, p. 6510, 2018.

[35] D. A. Chapman, A. C. Darby, M. Da Silva, C. Upton, A. D. Radford, and L. K. Dixon, "Genomic analysis of highly virulent Georgia 2007/1 isolate of African swine fever virus," *Emerging Infectious Diseases*, vol. 17, no. 4, pp. 599–605, 2011.

[36] F. Olasz, I. Meszaros, S. Marton et al., "A simple method for sample preparation to facilitate efficient whole-genome sequencing of african swine fever virus," *Viruses*, vol. 11, 2019.

[37] C. Gallardo, A. Soler, R. Nieto et al., "Experimental infection of domestic pigs with african swine fever virus Lithuania 2014 genotype II field isolate," *Transboundary and Emerging Diseases*, vol. 64, no. 1, pp. 300–304, 2017.

[38] J. N. Hakizimana, J. B. Ntirandekura, C. Yona et al., "Complete genome analysis of African swine fever virus responsible for outbreaks in domestic pigs in 2018 in Burundi and 2019 in Malawi," *Tropical Animal Health and Production*, vol. 53, no. 4, p. 438, 2021.

[39] N. Mazur-Panasiuk, G. Wozniakowski, and K. Niemczuk, "The first complete genomic sequences of African swine fever virus isolated in Poland," *Scientific Reports*, vol. 9, no. 1, p. 4556, 2019.

[40] A. S. Olesen, L. Lohse, M. D. Dalgaard et al., "Complete genome sequence of an African swine fever virus (ASFV POL/2015/Podlaskie) determined directly from pig erythrocyte-associated nucleic acid," *Journal of Virological Methods*, vol. 261, pp. 14–16, 2018.

[41] A. Mazloum, A. van Schalkwyk, A. Shotin et al., "Comparative analysis of full genome sequences of african swine fever virus isolates taken from wild boars in Russia in 2019," *Pathogens*, vol. 10, no. 5, p. 521, 2021.

[42] E. P. Njau, J. B. Domelevo Entfellner, E. M. Machuka et al., "The first genotype II African swine fever virus isolated in Africa provides insight into the current Eurasian pandemic," *Scientific Reports*, vol. 11, no. 1, p. 13081, 2021.

[43] P. Mileto, F. da Conceicao, V. Stevens et al., "Complete genome sequence of african swine fever virus isolated from a domestic pig in timor-leste, 2019," *Microbiology Resource Announcements*, vol. 10, no. 26, Article ID e0026321, 2021.

[44] G. Kovalenko, A. L. Ducluzeau, L. Ishchenko et al., "Complete genome sequence of a virulent african swine fever virus from a domestic pig in Ukraine," *Microbiology Resource Announcement*, vol. 8, no. 42, e00819 pages, 2019.

[45] K. Katoh, K. Misawa, K. Kuma, and T. Miyata, "MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform," *Nucleic Acids Research*, vol. 30, no. 14, pp. 3059–3066, 2002.

[46] A. Stamatakis, "RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies," *Bioinformatics*, vol. 30, no. 9, pp. 1312-1313, 2014.

[47] A. J. Drummond, M. A. Suchard, D. Xie, and A. Rambaut, "Bayesian phylogenetics with BEAUti and the BEAST 1.7," *Molecular Biology and Evolution*, vol. 29, no. 8, pp. 1969–1973, 2012.

[48] D. Darriba, G. L. Taboada, R. Doallo, and D. Posada, "jModelTest 2: more models, new heuristics and parallel computing," *Nature Methods*, vol. 9, no. 8, p. 772, 2012.

[49] H. J. Bandelt, P. Forster, and A. Rohl, "Median-joining networks for inferring intraspecific phylogenies," *Molecular Biology and Evolution*, vol. 16, no. 1, pp. 37–48, 1999.

[50] F. Bielejec, G. Baele, B. Vrancken, M. A. Suchard, A. Rambaut, and P. Lemey, "SpreaD3: interactive visualization of spatiotemporal history and trait evolutionary processes," *Molecular Biology and Evolution*, vol. 33, no. 8, pp. 2167–2169, 2016.

[51] S. L. Kosakovsky Pond and S. D. W. Frost, "Not so different after all: a comparison of methods for detecting amino acid sites under selection," *Molecular Biology and Evolution*, vol. 22, no. 5, pp. 1208–1222, 2005.

[52] J. Jumper, R. Evans, A. Pritzel et al., "Highly accurate protein structure prediction with AlphaFold," *Nature*, vol. 596, no. 7873, pp. 583–589, 2021.

[53] E. Bramucci, A. Paiardini, F. Bossa, and S. Pascarella, "PyMod: sequence similarity searches, multiple sequence-structure alignments, and homology modeling within PyMOL," *BMC Bioinformatics*, vol. 13, no. S4, p. S2, 2012.

[54] L. Gao, X. Sun, H. Yang et al., "Epidemic situation and control measures of african swine fever outbreaks in China 2018-2020," *Transboundary and Emerging Diseases*, vol. 68, no. 5, pp. 2676–2686, 2021.

[55] J. Yang, K. Tang, Z. Cao et al., "Demand-driven spreading patterns of African swine fever in China," *Chaos*, vol. 31, no. 6, Article ID 061102, 2021.

[56] E. Ramirez-Medina, E. Vuono, S. Pruitt et al., "Development and in vivo evaluation of a mgf110-1L deletion mutant in african swine fever strain Georgia," *Viruses*, vol. 13, 2021.

[57] H. Zhong, S. Fan, Y. Du et al., "African swine fever virus mgf110-7L induces host cell translation suppression and stress granule formation by activating the PERK/PKR-eIF2α pathway," *Microbiology Spectrum*, vol. 10, no. 6, Article ID e0328222, 2022.

[58] A. L. Reis, C. C. Abrams, L. C. Goatley et al., "Deletion of African swine fever virus interferon inhibitors from the genome of a virulent isolate reduces virulence in domestic pigs and induces a protective response," *Vaccine*, vol. 34, no. 39, pp. 4698–4705, 2016.

[59] A. Hubner, C. Keßler, K. Pannhorst et al., "Identification and characterization of the 285L and K145R proteins of African swine fever virus," *Journal of General Virology*, vol. 100, no. 9, pp. 1303–1314, 2019.

[60] Y. Chen, H. Liu, C. Yang et al., "Structure of the error-prone DNA ligase of African swine fever virus identifies critical active site residues," *Nature Communications*, vol. 10, no. 1, p. 387, 2019.

[61] Y. Ran, D. Li, M. G. Xiong et al., "African swine fever virus I267L acts as an important virulence factor by inhibiting RNA

polymerase III-RIG-I-mediated innate immunity," *PLoS Pathogens*, vol. 18, no. 1, Article ID e1010270, 2022.

[62] X. Li, K. Xiao, Z. Zhang et al., "The recombination hot spots and genetic diversity of the genomes of African swine fever viruses," *Journal of Infection*, vol. 80, no. 1, pp. 121–142, 2020.

[63] M. Nefedeva, I. Titov, S. Tsybanov, and A. Malogolovkin, "Recombination shapes African swine fever virus serotype-specific locus evolution," *Scientific Reports*, vol. 10, no. 1, Article ID 18474, 2020.

[64] Z. Zhu, C. T. Xiao, Y. Fan et al., "Homologous recombination shapes the genetic diversity of African swine fever viruses," *Veterinary Microbiology*, vol. 236, Article ID 108380, 2019.