M9

Genome Comparison Revealed Region-Enriched Genes of African Swine Fever Virus(ASFV)

Seungyeon Choi, Jongbin Park, and Eun Bae Kim*

Department of Animal Life Science, College of Animal Life Sciences, Kangwon National University, Chuncheon, Kangwon-do, Republic of Korea

* Corresponding author : Eun Bae Kim, itanimal@kangwon.ac.kr

Abstract

Since the introduction of African swine fever virus (ASFV) into Georgia in 2007 from Africa, the disease has spread from European continent to Asia, which is causing severe economic losses in the swine industry due to the lack of knowledge of any good vaccines. Several studies already compared its genome, but they did not consider regional characteristics. Here, we analyzed 89 ASFV genomes originated from Africa, America, Asia and Europe using comparative genomics to identify the regional characteristics. We downloaded the genomes from NCBI GenBank and classified them based on country and continent. From the genome comparison, we found 605 orthologs from 89 genomes, which includes 54 core and 551 accessory genes. Of the accessory genes, a total of 15 genes were identified to be region- enriched genes (1, 2, and 12 genes from America, Europe, and Asia, respectively). However, further studies are required to find biological and geographical meaning of such region-enriched genes. This work will present an improved genetic basis to develop new strategies to reduce the risk of ASFV transmission to domestic pigs.

Introduction

African swine fever (ASF) is a highly lethal contagious disease of domestic pigs, which is caused by the African swine fever virus (ASFV). ASF can cause a very high mortality reaching up to 100% when introduced into a new area. Thus, ASFV is causing severe economic losses in the swine industry due to the lack of knowledge of any good vaccines. After the first report of ASF in 1921 in Kenya, it was introduced European countries in 1957. And then, a new outbreak of ASF was reported in the Liaoning province of China in August 2018. The disease rapidly spread to 12 provinces in China causing 46 outbreaks in China. The pan-genomic that leads to concept of core gene and accessory gene is determined to characterize the virulence. Especially, if the number of unique genes in the pan-genome of a specific species has been increasing with the increasing number of strains, this species is considered to have an ‘open’ pan-genome. However, the pan-genome of ASFV and its property have not been clarified. Increased in genetic diversity as ASFV spreads by region, it is urgent to understand the genetic characteristics of ASFV according to its region or regional characteristics. Several studies already compared its genome, but they did not consider regional characteristics. Here, we analyzed 89 ASFV genomes originated from Africa, America, Asia and Europe using comparative genomics to identify the regional characteristics. Our results will provide insight into the genetic evolution of ASFV according to regional characteristics of ASFV and basic information for vaccine development.

Materials and Methods

▷ **Pan-genome inference**

We downloaded the genomes from NCBI GenBank and classified them based on country and continent. From the genome comparison, we found 605 orthologs from 89 genomes, which includes 54 core and 551 accessory genes. If a gene is present in all analysed ASFV genomes, this gene is considered to be a core gene. In order to estimate the size of pan-genome of ASFV, we sampled a varied number of strains (from 1 to 88, with increments of 1) from our data set, and for each number of strains, we randomly resampled 10 times. All of the calculations were done using python packages including pandas (v 0.24.2), matplotlib (v 2.0.0), numpy (v 1.16.6), scipy (v 1.2.2).

▷ **Phylogenomic analysis**

To analysis the phylogenetic relationship according to regional comparison of ASFV, we inferred phylogenies using maximum likelihood from downloaded NCBI ortholog data by ASFV strain. A maximum likelihood tree was constructed using python packages including matplotlib (v 2.0.0), scipy (v 1.2.2). And then, the phylogenetic tree was added label branches that was matched with origin region of ASFV strain.

▷**Extraction of region-enriched genes**

To detect region-enriched genes of ASFV genes, we filtered accessory genes that found only specific continent. Of the accessory genes, a total of 15 genes were identified to be region- enriched genes (1, 2, and 12 genes from America, Europe, and Asia, respectively) and then we compared the function of the region-enriched genes using BLASTn.

Results

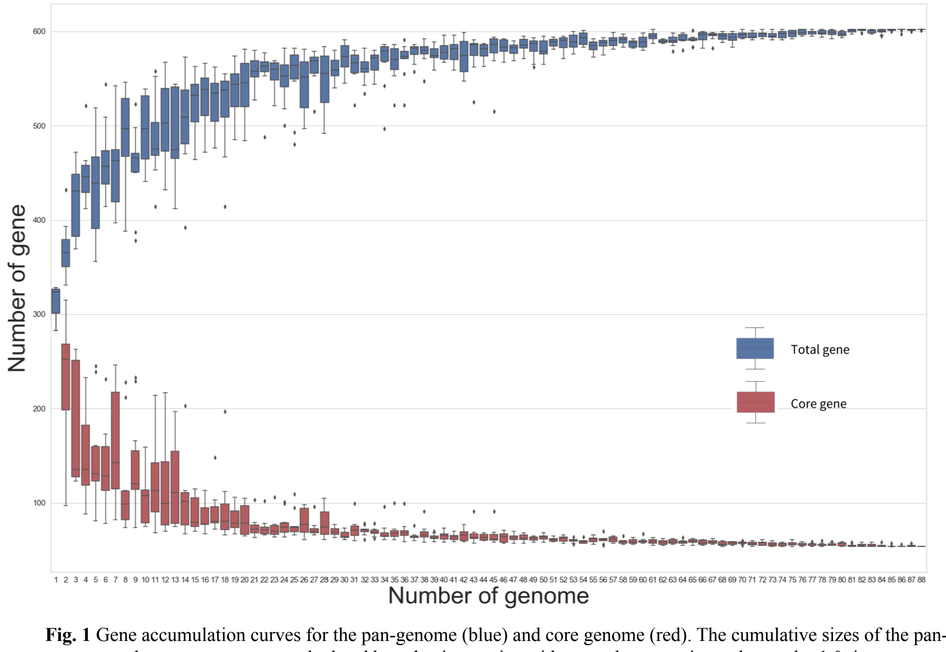

Fig. 1 Gene accumulation curves for the pan-genome (blue) and core genome (red). The cumulative sizes of the pan-genome and core genome were calculated by selecting strains without replacement in random order 1,0 times

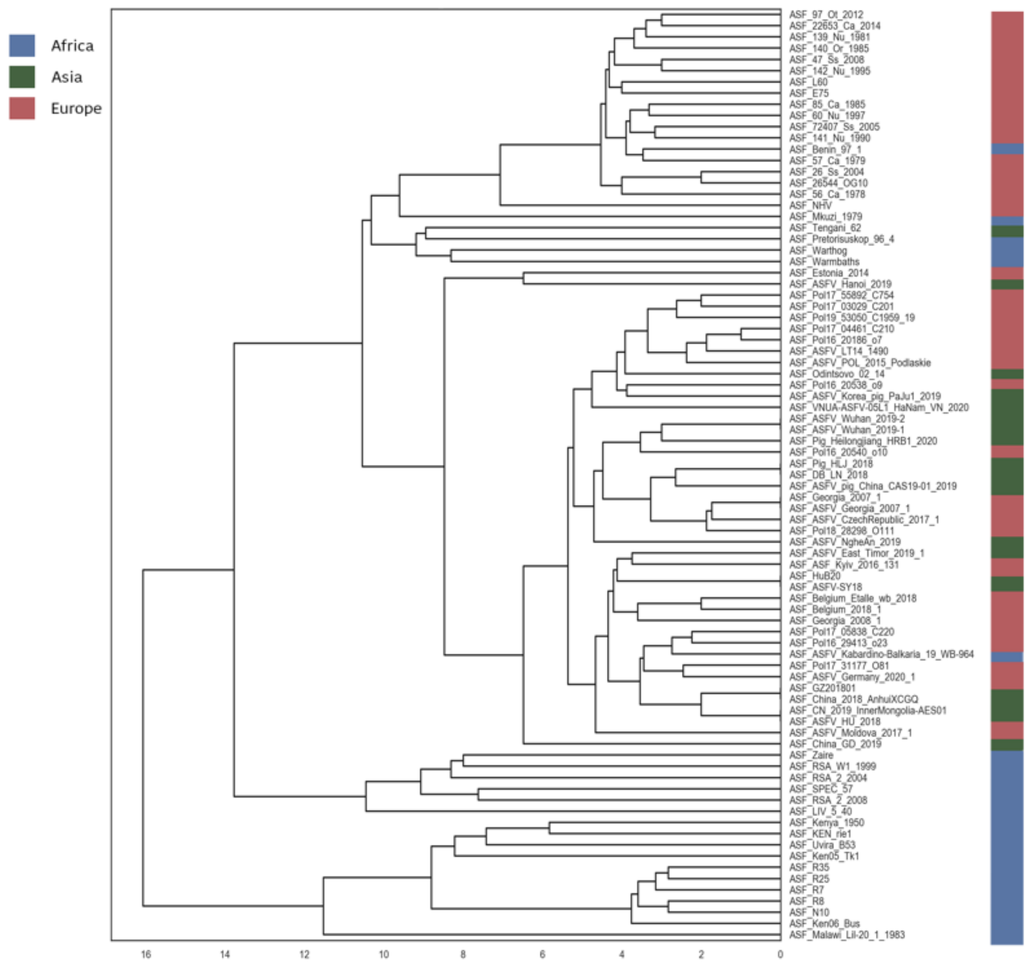

Fig. 2. The phylogeny of ASFV based on its genome. Label branches indicates origin of the ASFV genome.

Table 1. Origin region and function of region-enriched genes

Region	Genes	Function
Europe	Ortholog_536	Interacting selectively and non-covalently with the oxidized form, FAD, of flavin-adenine dinucleotide, the coenzyme or the prosthetic group of various flavoprotein oxidoreductase enzymes
	Ortholog_141	Plays a role in virus cell tropism, and may be required for efficient virus replication in macrophages
America	Ortholog_156	Uncharacterized membrane protein X69R
Asia	Ortholog_121	Be essential for the core assembly. Its myristoyl moiety may function as a membrane-anchoring signal to bind the developing core shell to the inner viral envelope .
	Ortholog_158	Be essential for the core assembly. Its myristoyl moiety may function as a membrane-anchoring signal to bind the developing core shell to the inner viral envelope.
	Ortholog_277	Helicase ATP-binding domain in association with helicase C-terminal domain catalyze the separation of double-stranded nucleic acids
	Ortholog_450	May participate in a redox cascade for the formation of disulfide bonds in viral proteins
	Ortholog_58	Plays a role in virus cell tropism, and may be required for efficient virus replication in macrophages
	Ortholog_89	Plays a role in virus cell tropism, and may be required for efficient virus replication in macrophages.
	Ortholog_261	Plays a role in virus cell tropism, and may be required for efficient virus replication in macrophages
	Ortholog_379	Plays a role in virus cell tropism, and may be required for efficient virus replication in macrophages
	Ortholog_385	Plays a role in virus cell tropism, and may be required for efficient virus replication in macrophages
	Ortholog_573	Plays a role in virus cell tropism, and may be required for efficient virus replication in macrophages
	Ortholog_498	Transmembrane protein B66L
	Ortholog_19	Uncharacterized protein I73R

Discussion

We reported the regional genome characterization of the ASFV strain and detected region-enriched genes of ASFV genes, searched function of the genes. This analysis expect to make a contribution to trace the evolution of ASFV during its spread. Analysis of these data provides valuable information for improvement of ASF diagnosis and vaccine development.

References

1. Bao, J., Wang, Q., Lin, P., Liu, C., Li, L., Wu, X., Wang, et al. (2019). Genome comparison of African swine fever virus China/2018/AnhuiXCGQ strain and related European p72 Genotype II strains. Transboundary and Emerging Diseases, 66, 1167– 1176.
2. Wang L, Luo Y, Zhao Y, Gao GF, Bi Y, Qiu H-J. (2020) Comparative genomic analysis reveals an ‘open’ pan-genome of African swine fever virus. Transbound. Emerg. Dis. 67, 1553-1562.
3. Bao YJ, Qiu J, Luo Y, Rodríguez F, Qiu HJ. The genetic variation landscape of African swine fever virus reveals frequent positive selection and adaptive flexibility. Transbound Emerg Dis. 2021 Mar 9.
4. Wang, Z., Jia, L., Li, J. et al. Pan-Genomic Analysis of African Swine Fever Virus. Virol. Sin. 35, 662–665 (2020)
5. Davies, K., Goatley, L. C., Guinat, C., Netherton, C. L., Gubbins, S., Dixon, L. K., & Reis, A. L. (2017). Survival of African swine fever virus in excretions from pigs experimentally infected with the Georgia 2007/1 isolate. Transboundary and Emerging Diseases, 64, 425– 431.