GMR
The Original

THE CENTRAL REGION OF THE B646L GENE SUPPORTS AFRICAN SWINE FEVER VIRUS GENOTYPE II IDENTIFICATION

Jacob Anderson C. Sanchez[1]*
[1]College of Arts and Sciences, Pampanga State Agricultural University, Magalang, Pampanga, Philippines
[1]Animal Disease Diagnostics Research Facility, Pampanga State Agricultural University, Magalang, Pampanga, Philippines
*Corresponding author: Jacob Anderson C. Sanchez, jacobsanchezphl@gmail.com

ABSTRACT
The B646L gene encoding the major capsid protein p72 is the standard genetic marker for African swine fever virus (ASFV) genotyping, with classification conventionally based on sequencing the C-terminal region. In contrast, the central region is routinely amplified during diagnostic PCR but has not been systematically evaluated for genotype assignment. Here, we assessed the genotyping utility of the central B646L region using laboratory-generated ASFV-positive sequences together with representative complete ASFV genomes. Phylogenetic analyses of the complete B646L coding sequence (CDS), central region, and C-terminal region consistently assigned the laboratory-generated sequences to genotype II. Comparative sequence analyses further identified seven parsimony-informative sites within the central region that distinguished genotype II from its closest neighboring genotype, despite the fragment containing fewer informative sites than the complete CDS and the conventional C-terminal region. These findings demonstrate that the routinely amplified central B646L fragment retains sufficient sequence variation for ASFV genotype II identification and represents a practical source of molecular information that can complement routine molecular surveillance.

KEYWORDS: African swine fever virus; B646L; genotype identification; p72; phylogenetic analysis

INTRODUCTION

African swine fever (ASF) is a highly contagious hemorrhagic disease of domestic and wild pigs caused by African swine fever virus (ASFV), the sole member of the family Asfarviridae (Alonso et al., 2018). There are 24 known genotypes of this virus (Bastos et al., 2003). Among them, the rapid and widespread distribution of ASFV genotype II across Europe and Asia, caused substantial economic losses to the swine industry (Dixon et al., 2020; Mighell & Ward, 2021). Molecular characterization has therefore become an important component of ASF surveillance and outbreak investigations by supporting genotype identification and molecular epidemiological analyses.

The B646L gene, which encodes the major capsid protein p72, is the most widely used genetic marker for ASFV genotype classification. The current p72 genotyping system is based on a partial C-terminal fragment of B646L described by Bastos et al. (2003), which is the frequently used marker for genotype assignment. Another region of the same gene, located at the central region of B646L, described herein as central region, is amplified by the conventional diagnostic PCR assay described by Wilkinson (2000) and subsequently prescribed by the World Organization for Animal Health (WOAH) for routine diagnostic use (Wilkinson, 2000; Aguero et al., 2003). Consequently, different regions of B646L have become established for different applications in ASFV molecular investigations.

Although whole-genome sequencing and SNP analyses provide high-resolution genetic information for ASFV characterization (Chernyshev et al., 2024), partial sequencing of the **B646L** gene remains widely used in routine diagnostic laboratories, especially in settings with limited access to whole-genome sequencing. Consequently, the genotyping potential of the routinely sequenced central region remains to be established.

In brief, we evaluated the central region of the B646L gene using laboratory-generated sequences together with representative complete ASFV genome sequences retrieved from GenBank. Complete B646L coding sequences were extracted from annotated genomes, and the corresponding central and C-terminal regions were compared for sequence variation and genotype discrimination. This study provides new insight into the potential application of the central region of B646L for ASFV genotyping.

II. MATERIALS AND METHODS

Sequence datasets

Two sequence datasets were analyzed in this study. The first comprised laboratory-generated partial B646L gene sequences obtained from African swine fever virus (ASFV)-positive pork samples processed at the Animal Disease Diagnostics Research Facility (ADDRF), Pampanga State Agricultural University. These sequences were generated using the conventional PCR assay of Wilkinson et al. (2000). Briefly, PCR amplification was performed using the primer pair ASFV_Primer1 (5′-ATGGATACCGAGGGAATAGC-3′) and ASFV_Primer2 (5′-CTTACCGATGAAAATGATAC-3′), generating a 278-bp fragment from the central region of the B646L gene. PCR products were sequenced bidirectionally by Macrogen Inc. (Seoul, South Korea).

The second dataset comprised representative complete ASFV genome sequences retrieved from the National Center for Biotechnology Information (NCBI) GenBank database. Complete B646L coding sequences (CDS) were extracted from annotated genome records for comparative analyses.

Sequence assembly and similarity analysis

Forward and reverse chromatograms of the laboratory-generated sequences were manually inspected and assembled into consensus sequences using Geneious Prime 2025.1.3. Consensus sequences were examined using ORFfinder (NCBI) to verify that the amplified fragment corresponded to the B646L gene encoding the major capsid protein p72. Sequence similarity was assessed using BLASTn against the GenBank nucleotide database (Altschul et al., 1990).

Preparation of B646L sequence datasets

Representative complete ASFV genome sequences were retrieved from the National Center for Biotechnology Information (NCBI) GenBank database. Sequence selection included representative genotypes described by Quembo et al. (2018), Qu et al. (2022), representative isolates recognized by the International Committee on Taxonomy of Viruses (ICTV), and complete genome sequences of ASFV genotype II from the Philippines. Complete B646L coding sequences were extracted from annotated genome records. For genomes lacking B646L annotations, the coding sequence was identified by pairwise sequence alignment with the ASFV Georgia 2007/1 reference genome (GenBank accession FR682468). The central and C-terminal regions were subsequently extracted from each B646L coding sequence for comparative analyses.

Sequence variation and phylogenetic analyses

Independent multiple sequence alignments were generated for the laboratory-generated sequences, complete B646L coding sequences (CDS), and the central and C-terminal regions using Clustal Omega (Sievers et al., 2011) implemented in Geneious Prime.

The numbers of conserved, variable, and parsimony-informative sites were determined using MEGA version 12 (Tamura et al., 2021). Maximum-likelihood phylogenetic trees were reconstructed using the Tamura 3-parameter model with gamma-distributed rate variation (T92+G) for the complete B646L CDS and C-terminal region, and the Jukes–Cantor model with gamma-distributed rate variation (JC+G) for the central region. Branch support was assessed using 1,000 bootstrap replicates.

Laboratory-generated sequences were assigned to ASFV genotypes based on their phylogenetic clustering with representative reference sequences. Phylogenetic trees derived from the complete B646L CDS, central region, and C-terminal region were compared to evaluate the consistency of genotype assignment.

III. RESULTS

B646L regions evaluated

Figure 1 shows a schematic illustration for the three B646L regions evaluated in this study. Comparative analyses were performed using the complete B646L coding sequence (1,941 bp), the 278-bp central region amplified by the conventional diagnostic PCR assay (Wilkinson et al., 2000), and the 415-bp C-terminal region used for p72 genotyping (Bastos et al., 2003). These regions formed the basis of the subsequent sequence variation and phylogenetic analyses.

Laboratory-generated ASFV sequences were assigned to genotype II

BLASTn analysis of the 13 laboratory-generated partial B646L sequences identified African swine fever virus (ASFV) genotype II as the closest match for all isolates. Each sequence showed 100% nucleotide identity, 94% query coverage, and an E-value of 0 against the highest-scoring GenBank entries. Because the top 100 BLASTn hits exhibited identical alignment statistics, only representative matches from Asian countries are presented in Table 1 to provide geographical context relevant to the Philippines.

The central B646L region retains genotype II assignment

Maximum-likelihood analysis of the laboratory-generated partial B646L sequences placed all 13 laboratory-generated ASFV sequences within the genotype II clade together with representative reference sequences (Figure 2). The laboratory-generated sequences clustered with genotype II isolates from the Philippines, Vietnam, China, Singapore, Hong Kong, and South Korea, whereas representative sequences from genotypes I, IV, V, VIII, IX, X, and XII formed distinct clades. The genotype II cluster was supported by a bootstrap value of 99%.

Because the analyses were based on partial B646L sequences, the phylogenetic trees were interpreted with respect to genotype assignment and clustering rather than broader evolutionary relationships among ASFV genotypes. Representative complete B646L coding sequences (CDS), together with their corresponding central and C-terminal regions, consistently placed the Philippine genotype II reference sequences within the genotype II clade (Figure 3), indicating that genotype II assignment was retained regardless of the B646L region analyzed.

Comparative nucleotide variation across the B646L gene

Comparison of representative ASFV genotype I and genotype II sequences identified seven parsimony-informative sites within the 278-bp central region of the B646L gene (Table 2). These sites were located at genome positions 104,490, 104,517, 104,534, 104,637, 104,697, 104,706, and 104,727 relative to the ASFV Georgia 2007/1 reference genome (FR682468). The laboratory-generated sequences shared identical nucleotide patterns with representative genotype II reference sequences at all seven positions, whereas genotype I reference sequences exhibited alternative nucleotides.

Alignment of the C-terminal region similarly identified four parsimony-informative sites distinguishing representative genotype I and genotype II sequences (Table 3). These sites were located at genome positions 103,821, 103,875, 103,962, and 104,031 relative to the ASFV Georgia 2007/1 reference genome. Representative genotype II sequences shared identical nucleotide patterns across all four positions, whereas genotype I reference sequences contained alternative nucleotides.

Comparison of sequence variation across the complete B646L coding sequence (CDS), central region, and C-terminal region revealed differences in the numbers of conserved, variable, and parsimony-informative sites (Table 4). The complete B646L CDS contained 1,679 conserved sites, 262 variable sites, and 211 parsimony-informative sites. Among the two partial regions, the C-terminal region contained 76 variable and 61 parsimony-informative sites, whereas the central region contained 32 variable and 28 parsimony-informative sites.

IV. DISCUSSION

The present study evaluated whether the central region of the B646L gene, routinely amplified during conventional African swine fever virus (ASFV) diagnosis, contains sufficient sequence variation for genotype assignment. Although ASFV genotyping has traditionally relied on sequencing the C-terminal region of B646L (Bastos et al., 2003), the routinely amplified central fragment consistently classified the laboratory-generated sequences as genotype II and produced the same genotype assignment as both the complete B646L coding sequence (CDS) and the conventional C-terminal genotyping region. These findings indicate that sequence information generated during routine diagnostic PCR can provide molecular information beyond confirmation of virus presence.

The B646L gene encodes the major capsid protein p72 and remains the principal genetic marker for ASFV molecular characterization because sequence variation within the gene forms the basis of the current classification of ASFV into 24 recognized genotypes (Bastos et al., 2003; Galindo & Alonso, 2017; Qu et al., 2022). Consequently, sequencing of the C-terminal region has become the standard approach for genotype assignment in molecular epidemiological investigations. In contrast, routine laboratory diagnosis relies on amplification of the central region using the conventional PCR assay of Wilkinson et al. (2000), subsequently optimized by Aguero et al. (2003) and recommended by WOAH (2024). Despite its widespread use in diagnostic laboratories, the potential of this routinely generated amplicon for genotype assignment has remained largely unexplored.

Comparison of phylogenetic trees reconstructed from the complete B646L CDS, central region, and conventional C-terminal region demonstrated that Philippine genotype II reference sequences remained assigned to the genotype II clade regardless of the B646L region analysed. Because the phylogenetic analyses were based on partial gene regions, the trees were interpreted with respect to genotype assignment rather than broader evolutionary relationships among ASFV genotypes. Notably, genotype I was consistently recovered as the nearest neighboring clade to genotype II in all B646L phylogenies. Sequence variation analyses therefore focused on these two genotypes, representing the most stringent comparison for evaluating the discriminatory capacity of the central region. Seven parsimony-informative sites were identified within the central region, compared with four in the conventional C-terminal fragment, demonstrating that the diagnostic amplicon retained sufficient nucleotide variation to distinguish genotype II from its closest neighboring genotype.

Recent advances in whole-genome sequencing have substantially improved ASFV molecular epidemiology (Warr et al., 2021; Koltsov., 2022, and Chernyshev et al., 2024). These approaches provide considerably greater genetic resolution than single-gene analyses and remain indispensable for investigating virus evolution, transmission pathways, and outbreak dynamics. However, conventional PCR followed by Sanger sequencing continues to be widely implemented in routine veterinary diagnostic laboratories because of its accessibility, lower cost, and established diagnostic performance (Aguero et al., 2003; WOAH, 2024). Rather than replacing whole-genome sequencing or the conventional C-terminal genotyping strategy, the present findings demonstrate that sequencing of the routinely amplified central B646L fragment can provide complementary genotype information without additional laboratory procedures.

Overall, the central region of B646L, routinely amplified during ASFV diagnosis, retained sufficient sequence variation for consistent genotype II assignment in the dataset analysed. Sequencing of this diagnostic amplicon may therefore provide a practical source of complementary genotype information, particularly where conventional PCR and Sanger sequencing remain the primary molecular tools.

REFERENCES

1. Aguero, M., Fernández, J., Romero, L., Sánchez Mascaraque, C., Arias, M., & Sánchez-Vizcaíno, J. M. (2003). Highly sensitive PCR assay for routine diagnosis of African swine fever virus in clinical samples. Journal of Clinical Microbiology, 41(9), 4431–4434. https://doi.org/10.1128/JCM.41.9.4431-4434.2003

2. Alonso, C., Borca, M., Dixon, L., Revilla, Y., Rodriguez, F., Escribano, J. M., & ICTV Report Consortium. (2018). ICTV virus taxonomy profile: Asfarviridae. Journal of General Virology, 99(5), 613–614. https://doi.org/10.1099/jgv.0.001049
3. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. Journal of Molecular Biology, 215(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2
4. Bastos, A. D. S., Penrith, M.-L., Crucière, C., Edrich, J. L., Hutchings, G., Roger, F., Couacy-Hymann, E., & Thomson, G. R. (2003). Genotyping field strains of African swine fever virus by partial p72 gene characterisation. Archives of Virology, 148(4), 693–706. https://doi.org/10.1007/s00705-002-0946-8
5. Chernyshev, R., Mazloum, A., Zinyakov, N., Kolbin, I., Shotin, A., Korennoy, F. I., Sprygin, A. V., Chvala, I. A., & Igolkin, A. (2024). Unique nucleotide polymorphism of African swine fever virus circulating in East Asia and Central Russia. Viruses, 16(12), 1907. https://doi.org/10.3390/v16121907
6. Department of Agriculture. (2020). Administrative Order No. 12, Series of 2020: National zoning and movement plan for the prevention and control of African swine fever. Department of Agriculture, Republic of the Philippines.
7. Dixon, L. K., Sun, H., & Roberts, H. (2020). African swine fever. Antiviral Research, 174, 104702. https://doi.org/10.1016/j.antiviral.2019.104702
8. Galindo, I., & Alonso, C. (2017). African swine fever virus: A review. Viruses, 9(5), 103. https://doi.org/10.3390/v9050103
9. Geneious Prime. (2025). Geneious Prime 2025.1.3. Biomatters Ltd. https://www.geneious.com
10. Hsu, C. H., Schambow, R., Montenegro, M., Miclat-Sonaco, R., & Perez, A. M. (2023). Factors affecting the spread, diagnosis, and control of African swine fever in the Philippines. Pathogens, 12(8), 1068. https://doi.org/10.3390/pathogens12081068
11. Koltsov, A., Tulman, E. R., Namsrayn, S., Kutish, G. F., & Koltsova, G. (2022). Complete genome sequence of virulent genotype I African swine fever virus strain K49 from the Democratic Republic of the Congo, isolated from a domestic pig (Sus scrofa domesticus). Archives of Virology, 167(10), 2377–2380. https://doi.org/10.1007/s00705-022-05543-2
12. Mighell, E., & Ward, M. P. (2021). African swine fever spread across Asia, 2018–2019. Transboundary and Emerging Diseases, 68(5), 2722–2732. https://doi.org/10.1111/tbed.14039
13. Qu, H., Ge, S., Zhang, Y., Wu, X., & Wang, Z. (2022). A systematic review of genotypes and serogroups of African swine fever virus. Virus Genes, 58(2), 77–87. https://doi.org/10.1007/s11262-021-01879-0
14. Quembo, C. J., Jori, F., Heath, L., Pérez-Sánchez, R., & Vosloo, W. (2018). Genetic characterization of African swine fever virus isolates from soft ticks at the wildlife/domestic interface in Mozambique and identification of a novel genotype. Transboundary and Emerging Diseases, 65(2), 420–431. https://doi.org/10.1111/tbed.12700
15. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., & Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular Systems Biology, 7, 539. https://doi.org/10.1038/msb.2011.75
16. Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: Molecular Evolutionary Genetics Analysis version 11. Molecular Biology and Evolution, 38(7), 3022–3027. https://doi.org/10.1093/molbev/msab120
17. Warr, A., Newman, C., Craig, N., Vendele, I., Pilare, R., Cariazo Cruz, L., Galase Barangan, T., Morales, R. G., Opriessnig, T., Venturina, V. M., Mananggit, M. R., Lycett, S., Domingo, C. Y. J., & Tait-Burkard, C. (2021). No part gets left behind: Tiled nanopore sequencing of whole ASFV genomes stitched together using Lilo. bioRxiv. https://doi.org/10.1101/2021.12.01.470769
18. Wilkinson, P. J. (2000). The laboratory diagnosis of African swine fever. In Manual of standards for diagnostic tests and vaccines (4th ed., Chapter 2.1.12). Office International des Epizooties.
19. World Organisation for Animal Health. (2024). Addressing African swine fever: Protocols and guidelines for laboratory diagnosis. https://doi.org/10.20506/asf.3475
20.