



Coding-Complete Genome Sequence of an African Swine Fever Virus from an Outbreak in 2021 among Domestic Pigs in Pangasinan, Philippines

 Andrew D. Montecillo,^a Zyne Baybay,^b Ralph Carolyn Cabug,^b Wreahlen Carias,^b John Paulo Jose,^c Sheila Mantaring,^c Annabelle Briones,^c Amanda Warr,^d  Christine Burkard,^d Lucille Villegas,^a Homer Pantua^{a,b}

^aMicrobiology Division, Institute of Biological Sciences, University of the Philippines Los Baños, Los Baños, Philippines

^bBioAssets Corporation, City of Santo Tomas, Batangas, Philippines

^cIndustrial Technology Development Institute, Department of Science and Technology (DOST-ITDI), Bicutan, Taguig City, Philippines

^dThe Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, United Kingdom

ABSTRACT We report a coding-complete genome sequence of an African swine fever virus from an outbreak in 2021 among domestic pigs in Pangasinan, Philippines using Oxford Nanopore Technologies minION. The linear genome assembly is a single contig with 192,377 bp.

The first African Swine Fever (ASF) outbreak in the Philippines was recorded in July 2019, and it continues to spread, leading to devastating impacts in the livestock industry. ASF, caused by the sole member of double-stranded DNA virus of the genus *Asfivirus*, family *Asfarviridae*, with a genome size of 170 to 194 kb, is a highly contagious disease with up to 100% fatality among infected domestic and wild pigs (1). Full-genome sequencing of the ASF virus is essential in understanding the biology and epidemiology of the pathogen. Here, we report an ASF virus coding-complete genome locally sequenced using Nanopore sequencing technology.

Extraction, library preparation, and sequencing kits were used following the respective manufacturer's protocol unless otherwise specified. Whole-blood samples were collected from pigs by a licensed veterinary consultant following the guidelines of the Bureau of Animal Industry, Department of Agriculture (Philippines) during an ASF outbreak in Pangasinan, Philippines in 2021. The total DNA was extracted in duplicate from a frozen ASF diagnostic PCR-positive EDTA-stabilized sample (from a grower pig) using MagMax DNA Multi-Sample Kit (Thermo Fisher Scientific). Then, NEBNext microbiome DNA enrichment kit (New England BioLabs) was used to deplete host DNA.

At least 5 ng of DNA was used as input to the rapid PCR barcoding kit (SQK-RPB004; ONT). The library was loaded on MinION SpotON R9.4.1 FLO-MIN106 flow cell (ONT) and sequenced using the 24-h script in MinKNOW Software (v.22.03.5; ONT) on MinION MK1b device. Reads were basecalled using super accurate model (dna_r9.4.1_450bps_sup.cfg) with default parameters in Guppy v6.0.6 (ONT).

Basecalled reads were combined in one fastq file and were quality-controlled using Porechop v0.2.4 (<https://github.com/rwick/Porechop>) with default parameters (-discard_middle). Reads were then mapped using Minimap2 v2.17-r941 (2) (-ax map-ont) against Sscrofa11.1 reference (GCF_000003025.6) (3) to remove host-derived reads. The unmapped reads, extracted with SAMtools v1.9 (4), were used for *de novo* assembly using Flye v2.4.2 (5) (parameters: nano-hq, genome size = 200 k, metagenome mode). The assembly was polished in Medaka v1.4.4 (<https://github.com/nanoporetech/medaka>) and was aligned against ASFV whole genomes from RefSeq database using MAFFT v7.490 (6). Maximum likelihood trees were constructed in IQTREE v2.2.0 (7).

A total of 105,660 reads (N50 = 4.23 kb) were generated of which 93,582 were retained after adapter trimming. About 1,422 of these did not map to Sscrofa11.1 reference genome.

Editor Jelle Matthijnsens, KU Leuven

Copyright © 2022 Montecillo et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Andrew D. Montecillo, admontecillo@up.edu.ph, or Homer Pantua, homer.pantua@bioassets.com.ph. The authors declare no conflict of interest.

Received 8 August 2022

Accepted 20 October 2022

Published 9 November 2022

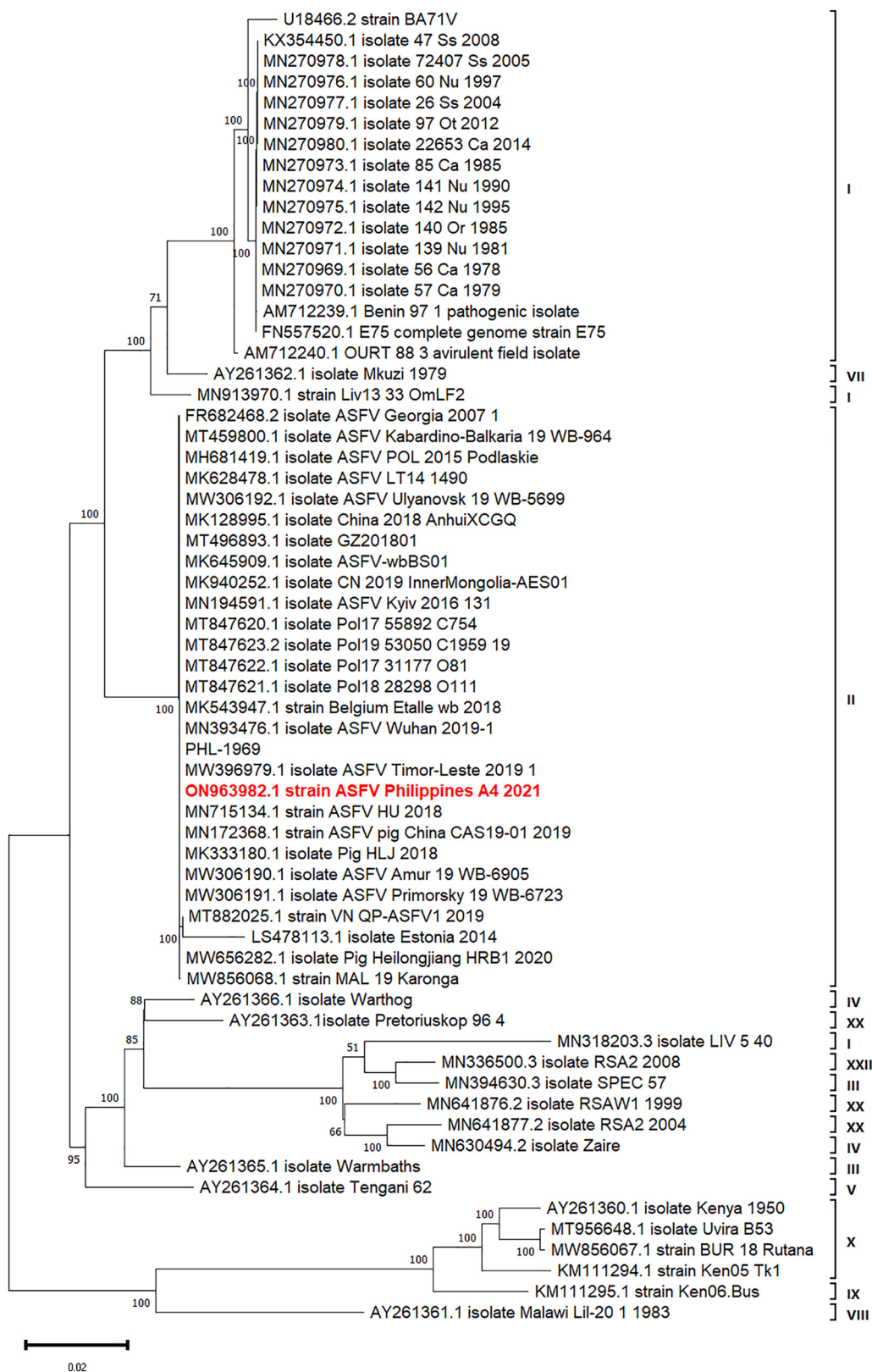


FIG 1 Maximum-likelihood consensus tree of select ASFV reference genomes and Philippines A4 2021 strain (**ON963982.1**) inferred using the ultrafast bootstrap (9) implemented in the IQ-TREE software (substitution model: GTR+F+R3). The tree is midpoint rooted. The scale bar is given in numbers of substitutions per site, and bootstrap resampling (1,000 replications) support values are shown at the nodes.

Mean assembly coverage was 21x (200 kb target genome size) and consisted of three linear contigs (203,823 bp total length). Only the longest fragment (192,377 bp, 38.3% GC content) was confirmed to be ASFV DNA using NCBI blastn search against RefSeq ASFV genomes (taxid:10497) on NCBI Virus database (accessed April 2022). The ASFV contig was aligned against PHL-1969 genome (8) and showed 99.88% identity in the longest contiguous alignment (192,120 bp) with 15 mismatches and 221 gaps but was missing the first 2,300 bp of the PHL-1969 genome 5'-end. Our strain belongs to the clade with PHL-1969, Timor-Leste 2019 strain (MW396979.1), Wuhan 2019 (MN393476.1), and other genotype II strains from China, Ukraine, Belgium, and Poland (Fig. 1).

Data availability. The genome sequence has been deposited in NCBI GenBank under the accession no. ON963982, and the raw sequence data can be found in the GenBank SRA under BioProject accession no. PRJNA857442. The version described in this paper is the first version, ON963982.1.

ACKNOWLEDGMENTS

A.D.M. is supported by the Philippine Council for Agriculture, Aquatic, and Natural Resources Research and Development of the Department of Science and Technology (DOST-PCAARRD) through its Graduate Research and Education Assistantship for Technology (GREAT) Program and by the University of the Philippines Los Baños Faculty Fellowship grant. This project is supported by the BioAssets Corporation, the DOST Science for Change Program (S4CP) Business Innovation through S&T for Industry (BIST) Grant no. 8738, and by the DOST-ITDI Grant no. LFP-EBD-2021-05. The authors would like to thank Dr. Elmer Vitug for the samples.

REFERENCES

1. Forth J, Forth L, King L, Groza O, Hübner A, Olesen A, Höper D, Dixon L, Netherton C, Rasmussen T, Blome S, Pohlman A, Beer M. 2019. A deep sequencing workflow for the fast and efficient generation of high-quality African swine fever virus whole genome sequences. *Viruses* 11:846. <https://doi.org/10.3390/v11090846>.
2. Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34:3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>.
3. Warr A, Affara N, Aken B, Beiki H, Bickhart DM, Billis K, Chow W, Eory L, Finlayson HA, Flicek P, Girón CG, Griffin DK, Hall R, Hannum G, Hourlier T, Howe K, Hume DA, Izuogu O, Kim K, Koren S, Liu H, Manchanda N, Martin FJ, Nonneman DJ, O'Connor RE, Phillippy AM, Rohrer GA, Rosen BD, Rund LA, Sargent CA, Schook LB, Schroeder SG, Schwartz AS, Skinner BM, Talbot R, Tseng E, Tuggle CK, Watson M, Smith TPL, Archibald AL. 2020. An improved pig reference genome sequence to enable pig genetics and genomics research. *Gigascience* 9:giaa051. <https://doi.org/10.1093/gigascience/giaa051>.
4. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25: 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
5. Kolmogorov M, Yuan J, Lin Y, Pevzner P. 2019. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 37:540–546. <https://doi.org/10.1038/s41587-019-0072-8>.
6. Katoh K, Standley D. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30: 772–780. <https://doi.org/10.1093/molbev/mst010>.
7. Nguyen L, Schmidt H, Von Haeseler A, Minh B. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 32:268–274. <https://doi.org/10.1093/molbev/msu300>.
8. Warr A, Newman C, Craig N, Vendelè I, Pilare R, Cruz L, Barangan T, Morales R, Opriessnig T, Venturina V, Mananggit M, Lycett S, Domingo C, Tait-Burkard C. 2021. No part gets left behind: Tiled nanopore sequencing of whole ASFV genomes stitched together using Lilo. *bioRxiv*.
9. Hoang D, Chernomor O, Von Haeseler A, Minh BQ, Vinh LS. 2018. UFBoot2: improving the ultrafast bootstrap approximation. *Mol Biol Evol* 35:518–522. <https://doi.org/10.1093/molbev/msx281>.