AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

# Complete Genome Sequence of a Virulent African Swine Fever Virus from a Domestic Pig in Ukraine

**Ganna Kovalenko,[a] Anne-Lise Ducluzeau,[b] Liudmyla Ishchenko,[a] Mykola Sushko,[c] Maryna Sapachova,[c] Nataliia Rudova,[d] Oleksii Solodiankin,[d] Anton Gerilovych,[d] Ralf Dagdag,[e] Matthew Redlinger,[e] Maksym Bezymennyi,[a] Maciej Frant,[f] Christian E. Lange,[g] Inna Dubchak,[h] Andrii A. Mezhenskyi,[c] Serhiy Nychyk,[a] Eric Bortz,[a,e] Devin M. Drown[b,i]**

[a]Institute of Veterinary Medicine (IVM), National Academy of Agrarian Sciences of Ukraine, Kiev, Ukraine
[b]Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, Alaska, USA
[c]State Scientific Research Institute of Laboratory Diagnostics and Veterinary and Sanitary Expertise (SSRILDVSE), Kiev, Ukraine
[d]National Scientific Center Institute of Experimental and Clinical Veterinary Medicine (NSC IECVM), Kharkiv, Ukraine
[e]Department of Biological Sciences, University of Alaska Anchorage, Anchorage, Alaska, USA
[f]National Veterinary Research Institute (NVRI), Pulawy, Poland
[g]Metabiota, San Francisco, California, USA
[h]Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, California, USA
[i]Department of Biology and Wildlife, University of Alaska Fairbanks, Fairbanks, Alaska, USA

**ABSTRACT** Here, we report the complete genome sequence of an African swine fever (ASF) virus (ASFV/Kyiv/2016/131) isolated from the spleen of a domestic pig in Ukraine with a lethal case of African swine fever. Using only long-read Nanopore sequences, we assembled a full-length genome of 191,911 base pairs in a single contig.

African swine fever (ASF) is a hemorrhagic viral disease of pigs that is characterized by high mortality and significant economic losses in domestic pigs. The virulent p72 genotype II lineage of African swine fever virus (ASFV; *Asfivirus*, family *Asfarviridae*) has been spreading rapidly from Georgia to Eastern and Central Europe since 2007 (1, 2). The global situation dramatically escalated, with confirmation that ASFV reached China in 2018 and had spread quickly across East Asia in 2019 (2, 3). In Ukraine, 485 laboratory-confirmed outbreaks of ASFV were recorded between 2012 and 2019 (4). The continued detection of outbreaks and the spread over the entire country raises concerns of ASFV becoming endemic in Ukraine.

Tissue samples were collected from a domestic pig from ASF outbreak number 131 in Kyiv Oblast, Ukraine, in 2016. The samples were frozen, and total DNA was extracted in duplicate from spleen tissue using the PowerMicrobiome RNA isolation kit (Mo Bio) following the manufacturer's protocol. The extraction kit retains both RNA and DNA, and a diagnostic conventional real-time PCR (RT-PCR) (LSI VetMax p72 PCR kit; Thermo Fisher Scientific) confirmed ASFV in the sample (http://www.asf.vet.ua/index.php/asfinukraine).

For full-genome sequencing, we purified the extracted DNA using Agencourt AMPure XP beads (Beckman Coulter) with three different bead-to-DNA (vol/vol) ratios (0.4×, 0.7×, and 1.0×) to retain a range of fragment lengths. We collected DNA sequence data across two sequencing runs using the Oxford Nanopore Technologies (ONT) MinION platform (Table 1). For sequencing run 1, we pooled 522 ng of DNA, including 300 ng of DNA purified from the 0.4× cleanup and 222 ng from the 0.7× cleanup. For sequencing run 2, we used 490 ng of DNA from the 1.0× cleanup. For each run, we prepared a rapid sequencing library (SQK-RAD004; ONT) and sequenced the prepared library on an R9.4.1 flow cell (FLO-MIN106) for 48 h

**Citation** Kovalenko G, Ducluzeau A-L, Ishchenko L, Sushko M, Sapachova M, Rudova N, Solodiankin O, Gerilovych A, Dagdag R, Redlinger M, Bezymennyi M, Frant M, Lange CE, Dubchak I, Mezhenskyi AA, Nychyk S, Bortz E, Drown DM. 2019. Complete genome sequence of a virulent African swine fever virus from a domestic pig in Ukraine. Microbiol Resour Announc 8:e00883-19. https://doi.org/10.1128/MRA.00883-19.

**Editor** Jelle Matthijnssens, KU Leuven



Address correspondence to Devin M. Drown, dmdrown@alaska.edu.

**Received** 13 August 2019
**Accepted** 17 September 2019
**Published** 17 October 2019

**TABLE 1** Summary of sequencing data statistics

| Data set | No. of reads | Yield (bp) | Avg length (bp) | Avg quality (Q score) |
|---|---|---|---|---|
| Run 1 raw | 2,213,244 | 5,213,092,675 | 2,355.4 | 10.5 |
| Run 1 quality controlled | 1,649,561 | 3,994,767,329 | 2,421.7 | 13.1 |
| Run 2 raw | 3,290,571 | 6,235,197,650 | 1,894.9 | 10.5 |
| Run 2 quality controlled | 2,430,267 | 4,631,455,885 | 1,905.7 | 13.1 |
| Total quality controlled | 4,079,828 | 8,626,223,214 | 2,114.4 | 13.1 |
| After *S. scrofa* removed | 98,078 | 27,426,602 | 279.6 | 12.5 |

using a MinION Mk1B device. We base called raw data using Guppy v3.1.5 (ONT) with a high-accuracy model (dna_r9.4.1_450bps_hac.cfg) and default parameters. Before assembly, we created a quality-controlled data set using Porechop v0.2.4 (https://github.com/rrwick/Porechop) with default parameters to trim adaptors and discard sequences with middle adapters (-discard_middle) and Filtlong v0.2.0 (https://github.com/rrwick/Filtlong) to filter by a quality (Q) score of ≥10 (-min_mean_q 90).

To remove reads likely originating from the host, we used Minimap2 v2.17-r941 (4) with default parameters for Nanopore reads (-ax map-ont) to align our quality-controlled data to the *Sus scrofa* 11 reference genome (GenBank assembly accession no. GCA_000003025). We extracted the unmapped reads for *de novo* assembly using SAMtools v1.9 (5). We assembled the genome using Flye v2.4.2 (6) with default parameters specifying the estimated genome size (-genome-size = 200k) and Nanopore reads (-nano-raw). Our raw assembly (coverage, 32×) consists of 227,741 bp in 9 linear contigs ($N_{50}$, 193,207 bp). We confirmed via an NCBI blastn search (7) that only the longest contig, 193,207 bp, was ASFV. The top hits for this contig had a greater than 99% identity to recently published ASFV genomes (e.g., GenBank accession no. LR536725 and MK128995). For this longest contig, we used a polishing pipeline specific to the error profile of Nanopore reads. We completed two rounds of polishing with the graphics processing unit (GPU)-enabled version of Racon v1.3.3 (https://github.com/clara-genomics/racon-gpu) (8) with the following parameters: score for matching bases (-match 8), score for mismatching bases (-mismatch -6), threshold for average base quality of windows (-quality-threshold -1), default gap penalty (-gap -8), default window (-window-length 500), and number of Compute Unified Device Architecture (CUDA) batches (-c 2). We completed a final round of Nanopore-specific polishing with Medaka v0.8.0 (https://github.com/nanoporetech/medaka).

Our polished assembly (coverage, 27×) consists of 191,911 bp in a single linear contig (GC content, 38.3%). We confirmed that this is a Georgia lineage p72 genotype II strain by using blastn (7) to compare the sequence variation at the p72 (B646L) gene and found 99.9% identity to Georgia 2007/1 (GenBank accession no. FR682468), with the only difference associated with two different homopolymer regions. We used MAFFT v7.388 (9, 10) to align our polished assembly with the Georgia 2007/1 assembly. A visual inspection of this alignment with Geneious Prime v 2019.1.1 confirmed overlap of the complete genome, including ends. We found a 10-nucleotide insertion (GGAA TATATA) between the I73R and the I329L genes, as previously reported (11–13). According to those previous reports, this insertion is not linked to attenuation or virulence. This insertion is present in the China/2018/AnhuiXCGQ genome (GenBank accession no. MK128995) but absent from ASFV/POL/2015/Podlaskie (accession no. MH681419) and Georgia 2007/1. We did not detect the left-end genome deletion found in an attenuated strain from Estonia (14). This is the first full-genome report of a virulent Georgia lineage p72 genotype II strain from the multiyear zoonotic outbreaks of ASF in Ukraine, highlighting the accuracy and deployability of Nanopore sequencing on a MinION platform for analysis of an emerging disease that has spread across Eurasia.

**Data availability.** This genome sequence has been deposited in GenBank under the accession no. MN194591. The version described in this paper is the first version, MN194591.1. Raw data for this project can be found in the GenBank SRA under accession no. PRJNA555080.

## ACKNOWLEDGMENTS

Many thanks to Natalia Usachenko, Maryna Kit, Xiao Bai, and Jeremy Buttler for technical support; Grzegorz Woźniakowski and Jürgen Richt for helpful discussions; and to Natalia Gudz, Nataliya Mykhaylovska, Olga Fedorenko, Karen Hite, David Mustra, and Mary Guttieri for project support.

This work was funded by the U.S. Defense Threat Reduction Agency (DTRA) through the Biological Threat Reduction Program in Ukraine (Ukraine Project-9). The contents of this publication are the responsibility of the author and do not necessarily reflect the views of DTRA or the United States Government. Research reported in this publication was in part supported by an NIH NIGMS Institutional Development Award (IDeA), grant no. P20GM103395 (Alaska INBRE). We also acknowledge the generous support of the Institute of Arctic Biology at UAF and that of the UAA College of Arts and Sciences.

## REFERENCES

1. Chapman DA, Darby AC, Da Silva M, Upton C, Radford AD, Dixon L. 2011. Genomic analysis of highly virulent Georgia 2007/1 isolate of African swine fever virus. Emerg Infect Dis 17:599–605. https://doi.org/10.3201/eid1704.101283.
2. OIE World Organisation for Animal Health. 2019. African swine fever (ASF) report no 21: June 21–July 04, 2019. OIE World Organisation for Animal Health, Paris, France. http://www.oie.int/fileadmin/Home/eng/Animal_Health_in_the_World/docs/pdf/Disease_cards/ASF/Report_21_Current_situation_of_ASF.pdf.
3. Ge S, Li J, Fan X, Liu F, Li L, Wang Q, Ren W, Bao J, Liu C, Wang H, Liu Y, Zhang Y, Xu T, Wu X, Wang Z. 2018. Molecular characterization of african swine fever virus, China, 2018. Emerg Infect Dis 24:2131–2133. https://doi.org/10.3201/eid2411.181274.
4. Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34:3094–3100. https://doi.org/10.1093/bioinformatics/bty191.
5. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079. https://doi.org/10.1093/bioinformatics/btp352.
6. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol 37:540. https://doi.org/10.1038/s41587-019-0072-8.
7. Zhang A, Schwartz S, Wagner L, Miller W. 2000. A greedy algorithm for aligning DNA sequences. J Comput Biol 7:203–214. https://doi.org/10.1089/10665270050081478.
8. Vaser R, Sovic I, Nagarajan N, Sikic M. 2017. Fast and accurate *de novo* genome assembly from long uncorrected reads. Genome Res 27:737–746. https://doi.org/10.1101/gr.214270.116.
9. Katoh K, Misawa K, Kuma KI, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 30:3059–3066. https://doi.org/10.1093/nar/gkf436.
10. Katoh K, Standley DM. 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.
11. Gallardo C, Fernández-Pinero J, Pelayo V, Gazaev I, Markowska-Daniel I, Pridotkas G, Nieto R, Fernández-Pacheco P, Bokhan S, Nevolko O, Drozhzhe Z, Pérez C, Soler A, Kolvasov D, Arias M. 2014. Genetic variation among African swine fever genotype II viruses, eastern and central Europe. Emerg Infect Dis 20:1544. https://doi.org/10.3201/eid2009.140554.
12. Mazur-Panasiuk N, Woźniakowski G, Niemczuk K. 2019. The first complete genomic sequences of African swine fever virus isolated in Poland. Sci Rep 9:4556. https://doi.org/10.1038/s41598-018-36823-0.
13. Goller KV, Malogolovkin AS, Katorkin S, Kolbasov D, Titov I, Höper D, Beer M, Keil GM, Portugal R, Blome S. 2015. Tandem repeat insertion in African swine fever virus, Russia, 2012. Emerg Infect Dis 21:731–732. https://doi.org/10.3201/eid2104.141792.
14. Zani L, Forth JH, Forth L, Nurmoja I, Leidenberger S, Henke J, Carlson J, Breidenstein C, Viltrop A, Höper D, Sauter-Louis C, Beer M, Blome S. 2018. Deletion at the 5′-end of Estonian ASFV strains associated with an attenuated phenotype. Sci Rep 8:6510. https://doi.org/10.1038/s41598-018-24740-1.