Complete genome sequence of an African swine fever virus (ASFV POL/2015/Podlaskie) determined directly from pig erythrocyte-associated nucleic acid

Olesen, Ann Sofie; Lohse, Louise; Dalgaard, Marlene Danner; Woniatkowski, Grzegorz; Belsham, Graham J.; Bøtner, Anette; Rasmussen, Thomas Bruun

Published in:
Journal of Virological Methods

Link to article, DOI:
10.1016/j.jviromet.2018.07.015

Publication date:
2018

Document Version
Peer reviewed version

Link back to DTU Orbit

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Accepted Manuscript

Title: Complete genome sequence of an African swine fever virus (ASFV POL/2015/Podlaskie) determined directly from pig erythrocyte-associated nucleic acid

Authors: Ann Sofie Olesen, Louise Lohse, Marlene Danner Dalgaard, Grzegorz Woźniakowski, Graham J. Belsham, Anette Bøtner, Thomas Bruun Rasmussen

PII: S0166-0934(18)30195-2
DOI: https://doi.org/10.1016/j.jviromet.2018.07.015
Reference: VIRMET 13509

To appear in: Journal of Virological Methods

Received date: 9-4-2018
Revised date: 27-7-2018
Accepted date: 27-7-2018


This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Complete genome sequence of an African swine fever virus (ASFV POL/2015/Podlaskie) determined directly from pig erythrocyte-associated nucleic acid

Ann Sofie Olesen a, Louise Lohse a, Marlene Danner Dalgaard b, Grzegorz Woźniakowski c, Graham J. Belsham a, Anette Bøtner a, Thomas Bruun Rasmussen a*

aDTU National Veterinary Institute, Technical University of Denmark, Lindholm, Kalvehave, Denmark
bDTU Bioinformatics, Technical University of Denmark, Kgs. Lyngby, Denmark
cDepartment of Swine Diseases, National Veterinary Research Institute, Pulawy, Poland

* Corresponding author: Thomas Bruun Rasmussen, DTU National Veterinary Institute, Technical University of Denmark, Lindholm, DK-4771 Kalvehave, Denmark, phone number +45 35887850, e-mail address tbrur@vet.dtu.dk

Highlights:
- The complete genome sequence of an ASFV from Poland (2015) was determined.
- This sequence was obtained directly from pig erythrocyte-associated nucleic acids.
- The work should assist studies on the genetic diversity of the European ASFVs.
ABSTRACT

African swine fever (ASF) is an important disease of domestic pigs and wild boar. The disease is caused by African swine fever virus (ASFV). In 2014, ASFV was introduced into Eastern Europe, and it has since then continued to spread within various Eastern European countries. Investigating differences in sequences between ASFV isolates may be a valuable tool to understand differences in virulence among them, however currently, no complete genome sequences of the viruses responsible for the Eastern European outbreaks have been reported. In this study, the complete genome sequence of a highly virulent ASFV was determined directly from erythrocyte-associated nucleic acids obtained from a pig experimentally infected with an isolate from Poland (ASFV POL/2015/Podlaskie). The sequence (ca. 189kb) of this recent European ASFV showed 95 nt differences (99.95 % identity) from the ASFV Georgia 2007/1 genome. The complete sequence of ASFV/POL/2015/Podlaskie should assist further studies on the genetic diversity and evolution of the European ASFVs.

Keywords: ASF, complete genome, NGS, Poland, WGS

African swine fever (ASF) is an important viral disease of domestic pigs and wild boar (EFSA Panel on Animal Health and Welfare, 2014). The African swine fever virus (ASFV), the only member of the Asfarviridae family, has a large (170-193 kbp) double-stranded DNA genome (Dixon et al., 2013). In 2007, ASFV was introduced into Georgia, and it has then spread within the Transcaucasian countries, the Russian Federation, Belarus and Ukraine (EFSA Panel on Animal Health and Welfare, 2014). More recently, since 2014, the virus has spread into various Eastern European countries (EFSA Panel on Animal Health and Welfare, 2015). Currently, virulence factors of the virus are not fully defined and
filling these knowledge gaps is required for understanding differences in virulence among ASFV strains (Blome et al., 2013). Comparison of genome sequences between strains of different virulence may be a valuable tool to understand such issues. At present, only the complete genome sequence of the ASFV Georgia 2007/1 isolate is available (Chapman et al., 2011). No complete genome sequences of viruses responsible for the Eastern European ASF outbreaks have been reported. In this study, we describe the complete genome sequence of the recent Eastern European ASFV POL/2015/Podlaskie (the isolate is described in Olesen et al. (2017)).

The ASFV POL/2015/Podlaskie sequence was determined using blood from a domestic pig experimentally infected with this virus isolate; this was derived from spleen material collected from an ASFV-infected wild boar in Podlaskie voivodeship (province) in February 2015 (Olesen et al., 2017). For the experimental infection, virus was isolated from spleen material by two passages in porcine pulmonary alveolar macrophages. Following virus inoculation, EDTA-stabilized blood was obtained from the pig (no. 23) at euthanasia, eight days later, following an acute rapid disease progression characterized by fever, incoordination, diarrhea and viremia (Olesen et al., 2017). Within 12 hours of collection of the blood sample, erythrocytes were separated from mononuclear cells and the plasma fraction of the blood using LymphoPrep™ density gradient medium (STEMCELL™ Technologies). Viral DNA was extracted from the erythrocyte-enriched sample (9.5 log_{10} virus genome copies/mL sample) on a MagNA Pure 96 system (Olesen et al., 2017). DNA libraries were generated using the extracted DNA with the Nextera XT DNA library preparation kit and sequenced on a MiSeq with Reagent Kit v3 600 bp (Illumina Inc.). The complete ASFV genome was determined by de novo assembly and iterative mapping using CLC Genomics Workbench (QIAGEN). PCR products were prepared and sequenced using Sanger and MiSeq methodologies to confirm parts of the sequence representing ca. 35kb of the genome. The PCR products were amplified from DNA extracted from the blood of pig 23, but also
from spleen material collected from a pig experimentally infected with a 2007 Georgian isolate (Nielsen and Uttenthal, 2013). Finally, the consensus sequence was compared to the ASFV Georgia 2007/1 genome (accession number FR682468 (Chapman et al., 2011)) using CLC Genomics Workbench and Geneious version 10.2.3 (Biomatters).

The complete genome sequence of ASFV POL/2015/Podlaskie (189393 bp, GenBank accession number MH681419) was obtained from a reference-based alignment consisting of 113246 mapped reads (0.84 % of 13489946 total reads) with an average coverage of 103 reads per nt using blood from pig 23. In addition, an identical consensus sequence, with a coverage of 159 reads per nt, was assembled from 253045 mapped reads combined from this pig and ten other pigs infected with the same inoculum (as described in Olesen et al. (2017)) using the same approach.

Comparison to the ASFV Georgia 2007/1 sequence revealed just 95 nt differences (thus 99.95 % identity) (Table 1). Of note, were a 3 bp deletion leading to loss of the terminal glycine codon in the MFG_110-14L open reading frame (ORF), and a deletion of a stretch of 6 Gs within the ORF of the ASFV_G_ACD-00350 gene resulting in the loss of two glycine codons. The effect on protein functions resulting from these deletions is unknown. Furthermore, four non-synonymous mutations were observed at: nt 62197 in the gene F1055L (Glu to Val), nt 155895 in the H240R gene (Arg to His), nt 166066 in the gene E199L (Gly to Arg), and nt 181965 in the gene I9R (Lys to Glu). Eleven other SNPs were observed, these were synonymous or situated in non-coding regions, including heterogeneity at nt 189223 (G/A), nt 189226 (A/G) and nt 189253 (C/G) (Table 1).

The remaining differences were 1-2 nt indels in homopolymeric regions, predominantly within intergenic regions, but 17 of these indels were identified within coding regions (Table 1). Some indels (17) were confirmed by sequencing of PCR products to rule out possible homopolymer artefacts. Most of these indels (14 out of 17) were also observed in DNA extracted from the spleen material collected from the
pig experimentally infected with a 2007 Georgian isolate (Nielsen and Uttenthal, 2013) (Table 1). This indicates that different variants of the ASFV Georgia 2007/1 sequence exist; these should be investigated further.

The complete genome sequence of the ASFV/Pol/2015/Podlaskie obtained from an experimentally infected pig should assist further studies on the genetic diversity and the evolution of the European ASFVs. Importantly, the use of whole-genome sequencing of viruses directly from blood-derived nucleic acid samples should minimize bias introduced by steps such as virus isolation and PCR amplification prior to sequencing.

Conflict of interest

The authors declare no conflicts of interest.

Acknowledgements

We are grateful to laboratory technicians Fie Fisker Brønnum, Hanne Egelund Hansen, Hanne Petersen, Heidi Elbrink and Neslihan Bicen, and to animal caretakers Heidi Lehman, Janni Oxfeldt, Josefine Cecilie Hammer and Marion Baltzer Petersen for their excellent work.

We gratefully acknowledge financial support from the CoVetLab, the Danish Ministry of Environment and Food and the Technical University of Denmark. The National Science Centre supported the scientific activities of Dr. Grzegorz Woźniakowski (grant no: UMO-2016/21/D/NZ6/00974).

References


Table 1

Variation detected between ASFV POL/2015/Podlaskie and ASFV Georgia 2007/1.

<table>
<thead>
<tr>
<th>Type</th>
<th>Genome position(*)(**) and modification</th>
<th>Annotation (ORF position(*))</th>
<th>Effect on ORF(***)</th>
<th>PCR (POL/2015)</th>
<th>PCR (Geo/07)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indel</td>
<td>412_413insA</td>
<td>Terminal inverted repeat</td>
<td>InsA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>432_433insCC</td>
<td></td>
<td>InsCC</td>
<td></td>
<td>No change</td>
</tr>
<tr>
<td>Indel</td>
<td>433_434_insT</td>
<td></td>
<td>InsT</td>
<td></td>
<td>InsT</td>
</tr>
<tr>
<td>Indel</td>
<td>440_441insA</td>
<td></td>
<td>InsA</td>
<td></td>
<td>InsA</td>
</tr>
<tr>
<td>Indel</td>
<td>1604_1607delTT</td>
<td>MGF_360-1L (852..1934)</td>
<td>Frameshift</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>1622_1623insT</td>
<td>MGF_360-1L (852..1934)</td>
<td>Frameshift</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>3246_3247insA</td>
<td>KP177R (3212..3781)</td>
<td>Frameshift</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>3861_3862insA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>5670_5671_insA</td>
<td>ASFV_G_ACD_00070 (&lt;5604..5732)</td>
<td>Frameshift</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>5820_5821insT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>5869_5870insA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>7823_7824insA</td>
<td>ASFV_G_ACD_00120 (7795..7920)</td>
<td>Frameshift</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>7940_7941_insA</td>
<td>ASFV_G_ACD_00120 (7795..7920)</td>
<td>Frameshift</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>8505_8506insA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>10465_10466insA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>13266_13270delCCC</td>
<td>MGF_110-14L (13253..13603)</td>
<td>Glycine codon deleted(****)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-------------------</td>
<td>-------------------------------</td>
<td>-----------------------------</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>13665_13666insA</td>
<td>ASFV_G_ACD_00240 (13602..13685)</td>
<td>Frameshift</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>16879_16880insGG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>18736_18737insT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>18830_18831insGG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>19032_19039delGGGGGG</td>
<td>ASFV_G_ACD_00350 (18994..19128)</td>
<td>Two Glycine codons deleted delGGGGGG</td>
<td>No change</td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>19461_19462insT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>20669_20670insA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>20834_20836delG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>21588_21589insA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>23027_27028insA</td>
<td></td>
<td>InsA</td>
<td>InsA</td>
<td></td>
</tr>
<tr>
<td>SNP</td>
<td>A26037G</td>
<td>MGF_360-10L (25401..26438)</td>
<td>Synonymous mutation</td>
<td>G</td>
<td>No change</td>
</tr>
<tr>
<td>Indel</td>
<td>26455_26456insT</td>
<td></td>
<td>InsT</td>
<td>InsT</td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>31678_31679insA</td>
<td></td>
<td>InsA</td>
<td>InsA</td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>39024_39025insTT</td>
<td></td>
<td>InsTT</td>
<td>InsT</td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>40730_40731insT</td>
<td></td>
<td>InsT</td>
<td>InsT</td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>42616_42617insT</td>
<td></td>
<td>InsT</td>
<td>InsT</td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>44547_44548insT</td>
<td></td>
<td>InsT</td>
<td>InsT</td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>47321_47322insA</td>
<td></td>
<td>InsA</td>
<td>InsA</td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>48663_48664insA</td>
<td>A151R (48651..49127)</td>
<td>Frameshift</td>
<td>InsA</td>
<td>InsA</td>
</tr>
<tr>
<td>Indel</td>
<td>55974_55975insA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNP</td>
<td>T62197A</td>
<td>F1055L (59853..62999)</td>
<td>Non synonymous mutation (E to V)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>Location</td>
<td>Description</td>
<td>Class</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>----------</td>
<td>-------------</td>
<td>-------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>63845_63846insT</td>
<td>EP1242L (66514..70242)</td>
<td>Synonymous mutation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>63860_63861insA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>63867_63868insT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>63876_63877insT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNP</td>
<td>G68524A</td>
<td>EP1242L (66514..70242)</td>
<td>Synonymous mutation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNP</td>
<td>G99635A</td>
<td>B34L (99614..100678)</td>
<td>Synonymous mutation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>81525_81526insT</td>
<td>C84L (81353..81592)</td>
<td>Frameshift</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>81619_81620insA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNP</td>
<td>G99635A</td>
<td>B34L (99614..100678)</td>
<td>Synonymous mutation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNP</td>
<td>G99635A</td>
<td>B34L (99614..100678)</td>
<td>Synonymous mutation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNP</td>
<td>G155895A</td>
<td>H240R (155339..156064)</td>
<td>Non synonymous mutation (R to H)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>161251_161253delA</td>
<td>QP383R (160525..161676)</td>
<td>Frameshift (change of 12 contiguous residues)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>Description</td>
<td>Reference</td>
<td>Notes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>--------------------------------------------------</td>
<td>-----------------</td>
<td>--------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>161292_161293insA</td>
<td>QP383R (160525..161676)</td>
<td>Frameshift (return to initial ORF)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNP</td>
<td>C165718T</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNP</td>
<td>C165762T</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNP</td>
<td>C165780T</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNP</td>
<td>C165782T</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNP</td>
<td>C166066T</td>
<td>E199L (165845..166444)</td>
<td>Non synonymous mutation (G to R)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>167763_167764insA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>169454_169455insT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>169462_169463insA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>169741_169742insT</td>
<td>I267L (169605..170444)</td>
<td>Frameshift</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>172017_172018insT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>173565_173566insA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>174954_174955insA</td>
<td>ASFV_G_ACD_01760 (174920..175003)</td>
<td>Frameshift (fusion with downstream ORF I177L)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>175632_175633insA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>177493_177494insA</td>
<td>MGF_360-16R (176590..177519)</td>
<td>Frameshift (fusion with downstream ORF DP63R)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>179379_179380insA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNP</td>
<td>A181965G</td>
<td>I9R (181698..181988)</td>
<td>Non synonymous mutation (K to E) near C-terminus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>183305_183306insT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>183314_183315insT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>184967_184969delT</td>
<td>ASFV_G_ACD_01940 (184949..185080)</td>
<td>Frameshift</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>186352_186353insT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>189022_189024delA</td>
<td>DP60R (188997..189161)</td>
<td>Frameshift</td>
<td>delA</td>
<td>No change</td>
</tr>
<tr>
<td>SNP</td>
<td>G189223A</td>
<td>Quasispecies</td>
<td>G/A</td>
<td>G/A</td>
<td></td>
</tr>
<tr>
<td>SNP</td>
<td>A189226G</td>
<td>Quasispecies</td>
<td>No change</td>
<td>No change</td>
<td></td>
</tr>
<tr>
<td>SNP</td>
<td>C189253G</td>
<td>Quasispecies</td>
<td>C/G</td>
<td>C/G</td>
<td></td>
</tr>
<tr>
<td>Indel</td>
<td>189282_189283insT</td>
<td></td>
<td>InsT</td>
<td>InsT</td>
<td></td>
</tr>
</tbody>
</table>

(*) The positions of detected variations and open reading frame (ORF) positions are based on the reference genome (ASFV Georgia 2007/1, FR682468).

(**) The position of the indels refer to the position of the last nucleotide within each homopolymer region. Genome positions for indels are given by the flanking nt positions.

(***) The effect of the frameshifts can increase or decrease the size of the protein or can just alter the amino acid sequence.

(****) Read in right to left direction.