The First Complete Genome Sequence of the African Swine Fever Virus Genotype X and Serogroup 7 Isolated in Domestic Pigs From the Democratic Republic of Congo

Patrick N. Bisimwa  
Evangelical University in Africa, Department of Animal Science and Production

Juliette R. Ongus  
Jomo Kenyatta University of Agriculture and Technology, Department of Molecular Biology and Biotechnology

Lucilla Steinaa  
International Livestock Research Institute

Espoir B. Bisimwa  
Evangelical University in Africa, Department of Animal Science and Production

Bochere Edwina  
International Livestock Research Institute

Eunice M. Machuka  
International Livestock Research Institute

Jean-Baka Domelevo Entfellner  
International Livestock Research Institute

Edward Okoth  
International Livestock Research Institute

Roger Pelle (✉ r.pelle@cgiar.org)  
Biosciences eastern and central Africa-International Livestock Research Institute (BecA-ILRI) Hub, Nairobi, Kenya  
https://orcid.org/0000-0003-1053-085X

Research

Keywords: African swine fever virus, complete genome sequence, initial characterization, genotype X, South Kivu

DOI: https://doi.org/10.21203/rs.3.rs-85005/v1

License: ☄️ ☀️ This work is licensed under a Creative Commons Attribution 4.0 International License.  Read Full License
Abstract

Background: African swine fever (ASF), a highly contagious hemorrhagic disease, affects domestic pigs in the Democratic Republic of Congo (DRC) where regular outbreaks are reported leading to high mortality rates approaching 100% in the affected regions. No study on the characteristics of the complete genome of strains responsible for ASF outbreaks in the South Kivu province of DRC is available, limited a better understanding of molecular evolution and spread of this virus within the country. The present study aimed at determining the complete genome sequence of ASFV strains genotype X involved in 2018-2019 ASF disease outbreaks in South Kivu province of DRC.

Materials and Methods: Total genomic of a spleen sample from an ASFV genotype X-positive domestic pig in Uvira, during the 2018-2019 outbreaks in South Kivu, was sequenced using the Illumina HiSeq X platform. Obtained trimmed reads using Geneious Prime 2020.0.4 were blasted against a pig reference genome then contigs were generated from the unmapped reads enriched in ASFV DNA using Spades implemented in Geneious 2020.0.4. The assembly of the complete genome sequence of ASFV was achieved from the longest overlapping contigs. The new genome was annotated with the genome annotation transfer utility (GATU) software and the CLC Genomics Workbench 8 software was further used to search for any ORFs that failed to be identified by GATU. Subsequent analyses of the newly determined Uvira ASFV genotype X genome were done using BLAST for databases search, CLUSTAL W for multiple sequences alignments and MEGA X for phylogeny.

Results: 42 Gbp paired-end reads of 150 bp long were obtained containing about 0.1% of ASFV DNA. The assembled Uvira ASFV genome, termed Uvira B53, was 180,916 bp long that could be assembled in 2 contigs. The Uvira B53 genome had a GC content of 38.5%, encoded 168 open reading frames (ORFs) and had 98.8% nucleotide identity with the reference ASFV genotype X Kenya 1950. The phylogenetic relationship with selected representative genomes clustered the Uvira B53 strain together with ASFV genotype X reported to date (Kenya 1950 and Ken05/Tk1). Multiple genome sequences comparison with the two reference ASFV genotype X strains showed that 131 of the 168 ORFs were fully conserved in the Uvira B53. The other 37 ORFs were divergent mainly due to SNPs and indels (deletions and insertions). Most of 46 multigene family (MGF) genes identified were affected by various genetic variations. However, 8 MGF ORFs present in Kenya 1950 and Ken05/Tk1 were absent from the Uvira B53 genome including three members of MGF 360, four of MGF 110 and one of MGF 100 while one MGF ORF (MGF 360-1L) at the left end of the genome was truncated in Uvira B53. Moreover, ORFs DP96R and p285L were also absent in the Uvira B53 genome. In contrast, the ORF MGF 110-5L present in Uvira B53 and Ken05/Tk1 was missing in Kenya 1950. The analysis of the intergenic region between the I73R and I329L genes also revealed sequence variations between the three genotype X strains mainly characterized by a deletion of 69 bp in Uvira B53 and 36 bp in Kenya 1950, compared to Ken05/Tk1. Assessment of the CD2v (EP402R) antigen unveiled the presence of SNPs and indels particularly in the PPPKPY tandem repeat region between selected variants representing the eight serogroups reported to date. Uvira B53 had identical CD2v variable region to the Uganda (KM609361) strain, the other only ASFV serogroup 7 reported to date.

Conclusion: We report the first complete genome sequence of an African swine fever virus (ASFV) p72 genotype X and CD2v serogroup 7, termed Uvira B53. This study provides vital data on genetic characteristics...
and evolution of ASFV useful for tracing the geographical spread of ASF and essential for improved design of control and management strategies against ASF.

Background

African swine fever (ASF), a hemorrhagic contagious viral disease affecting domestic pigs and wild boars, is characterized by high mortality and has become a significant threat to the global pig industry [1]. ASF is caused by African swine fever virus (ASFV), a large DNA virus and the only known member of the genus Asfivirus within the Asfarviridae virus family [2]. The genome is a double-stranded DNA varying between 170 to 193 kbp in size containing 150 to 167 proteins coding open reading frames (depending on the virus strains) and a conserved central region of about 125 kb, while the ends are variable in size [3]. The epidemiology is complex and embraces various patterns in endemic regions [4, 5]. To date, neither therapy nor vaccine are available against the disease and therefore, the primary prevention methods rely on the application of strict biosecurity measures, early diagnosis and stamping out the infected pig herds [6]. Based on the sequencing of the B646L gene, which encodes the capsid protein p72, 24 genotypes have been currently reported worldwide and all of them are known to circulate in Africa [7, 8].

Previous studies have identified p72 genotype I, IX and XIV to be present in the Democratic Republic of Congo (DRC) [9, 10]. In addition, our recent study has identified for the first time the presence of p72 genotype X in DRC [11] in symptomatic domestic pigs during outbreaks. This emphasize the need for continued characterization of ASFV strains responsible for outbreaks to better understand the spread of the disease in DRC and to develop preventive measures to control the spread. To date, only two complete and fully annotated genomes of ASFV genotype X are available in the GenBank, both of which are from Kenyan virus strains [12]. However, no ASFV complete genome sequencing study has been reported in DRC, an ASFV endemic country, limited a better understanding of molecular evolution and spread of this virus within the country. In this study, we report the first complete genome sequence of ASFV genotype X and serogroup 7 associated with ASF outbreaks in domestic pigs in the South Kivu province in eastern DRC in 2019. Analysis of this ASFV genome in comparison with the previously sequenced and available ASFV genomes provides new insights into genetic diversity of African strains and knowledge that may advance our understanding of host-virus interaction and pathogenicity.

Materials And Methods

Ethics statement

Ethical approval for the study reported here and the permission for the collection of samples was provided by the Interdisciplinary Centre for Ethical Research (CIRE) established by the Evangelical University in Africa, Bukavu, DRC, with reference (UEA/SGAC/KM 132/2016). A consent form which described the aim of the study was signed by farmers willing to participate in the study after translation into local languages.

DNA preparation and ASFV detection
A spleen tissue sample was collected during an ASF outbreak from an ASF symptomatic domestic pig in 2019 from the Uvira district of South Kivu province, DRC. Total genomic DNA was extracted directly from 30 mg of tissue using the DNeasy Blood and Tissue Kit (Qiagen, USA) following the manufacturer’s recommendations. The quality and integrity of the extracted DNA were assessed by 0.8% agarose gel electrophoresis in the presence of 0.5 µg/ml Gel Red (Biotium, Fisher Biotech, Australia) for DNA visualization. The quantity and purity of the DNA were estimated by NanoDrop™ 2000 (Thermo Fisher, USA) spectrophotometer, ensuring that sample with 260/280 OD ratio between 1.8 and 2 were included for downstream analyses. To confirm the presence of ASFV DNA in the sample, polymerase chain reaction (PCR) amplification assay was performed using the ASFV diagnosis primers PPA1/PPA2 (Peste Porcine Africaine) that target the virus VP73 (p72) coding region to generate an amplicon of 257 bp [13]. For p72 genotype classification, the C-terminal region of the p72 protein gene (B646L) was amplified and sequenced using primers p72-U/D [5].

**Complete genome sequencing of ASFV**

The whole genome sequencing of the total genomic DNA purified directly from ASFV-infected pig spleen was carried out at the Genohub Inc. (Austin, Texas USA) using the Illumina HiSeq X platform (Illumina, USA). DNA Libraries were prepared using KAPA HyperPlus Kit (KAPA Biosystems, Wilmington, MA) according to the manufacturer’s protocol with 10 ng DNA as input. The final library quality and quantity were analyzed in a Bioanalyzer 2100 (Agilent Technologies, USA) and Qubit™ Fluorometer (Thermo Scientific, USA), respectively. Whole genome sequencing (WGS) of 150 bp paired-end reads were sequenced on Illumina HiSeq X (Illumina Inc., San Diego, CA). A total of 42 Gbp paired-end reads of 150 bp long were obtained.

**Genome assembly and annotation**

Cleaning of raw reads was achieved using Geneious Prime 2020.0.4 and a Phred quality score (Q) of 30 and above. The trimmed reads were blasted against the pig reference genome Sscrofa11.1 from Ensembl (Acc. No. GCA_000003025.6) then the unmapped reads obtained significant hits to ASFV. The original 42 Gbp paired reads contained about 0.1% of ASFV DNA. Contigs were generated using Spades implemented in Geneious 2020.0.4 using default parameters in order to obtain the full genome sequence. Assembly of the complete genome sequence of Uvira B53 was achieved from the longest overlapping contigs. The genome annotation transfer utility (GATU) software at the Viral Bioinformatics Resource Center [14] was used to annotate the newly constructed genome using Ken05/Tk1 ASFV genome as a reference (GenBank Acc. No. NC044945) and CLC Genomics Workbench 8 software (Qiagen, USA) was further used to search for any ORFs that failed to be identified by GATU. A BLAST search was carried out against the NCBI database to confirm the identified ORFs. Moreover, the two ASFV genome sequences were aligned with MAFFT v.7.427 implemented in Geneious software 2020.0.4 to identify substitutions sites. Multiple sequence alignments of retrieved whole genome sequences from the GenBank with Uvira B53 genome were generated using CLUSTAL W [15]. The evolutionary analyses were inferred using unrooted neighbor-joining (NJ) method with MEGA X software [16] and 1000 bootstrap replications were estimated to assess the robustness of individual clades. The same software was used to perform the phylogenetic construction based on the most divergent genes using Maximum-likelihood method. The genome sequence generated from this study has been submitted to GenBank under the accession number MT956648.
Results

General features of the genome sequence of ASFV Uvira B53 strain

The ASFV Uvira B53 strain is a p72 genotype X isolated from the spleen of an ASFV-positive and symptomatic domestic pig in South Kivu province, East of DRC, during outbreaks reported in the Uvira district in 2019 [11]. Genotype X confirmation of the ASFV Uvira B53 was done through BLAST result showing the highest percentage identity with existing GenBank ASFV sequences of the genotype X (Kenya 1950 GenBank Acc. No. AY261360.1 and Ken05/Tk1 GenBank Acc. No. KM111294.1). The draft assembly of the Uvira B53 genome comprises 180,916 bp that could be assembled in 2 contigs with an N50 of 112,709 bp. Terminal inverted repeats were missing at both ends of the genome sequence, probably due to the limited amount of ASFV material (about 0.1% of total reads) in the spleen DNA sample sequenced. The base composition of the genome sequence showed a GC content of 38.5%, which is comparable to that of other genotype X strains, i.e. Ken05/Tk1 (38.3%) and Kenya 1950 (38.4%). Sequence annotation using GATU software [14] revealed a total of 168 protein-coding genes (Table 1). In total, 46 multigene family (MGF) genes were identified including MGF 100 (3 members), MGF 110 (13 members), MGF 300 (3 members), MGF 360 (18 members) and MGF 505 (9 members). However, there was a deletion of one MGF 360 member (MGF 360-2L) and 2 other MGF 360 members (MGF 360-21R and MGF 360-1L) in the right variable region; another MGF 360 member in the left variable region (MGF 360-1L) was truncated, being only 95 bp long starting from the 5’ end (compared to the 1071 bp of the complete ORF) and likely non-functional. Also, four MGF 110 members were missing in the Uvira B53 genome (MGF 110-4L, MGF 110-7L, MGF 110-8L and MGF 110-9L) as well as one MGF 100 member (MGF 100-1R).

Table 1 Summary of the Uvira B53 ASFV genomic sequencing data

| Genome assembly | Number of contigs | Largest contig (bp) | Total length (bp) | % GC content | N50 (bp) | L50 | ORFs |
|-----------------|------------------|--------------------|------------------|--------------|---------|-----|------|
| ASFV Uvira B53  | 9                | 112,709            | 180,916          | 38.5         | 112,709 | 1   | 168  |

Initial comparative analysis of the Uvira B53 strain genome with other ASFV strains

A total of seventeen ASFV complete genome sequences representing different strains reported in several ASF endemic countries were retrieved from the GenBank and were used in this study for comparison and phylogeny. Genome’s GenBank accession numbers, country and year of isolation, virus genotype, host, global alignment percentage identity to the genome of ASFV strain Uvira B53, and the reference are shown in Table 2. The pair-wise alignment between the Uvira B53 strain and other ASFV genomes showed the highest maximum percentage identity with strains of the genotype X, specifically Kenya 1950 (98.85%) and Ken05/Tk1 (95.52%) (Table 2).

Table 2 Comparison of complete genome sequences of Uvira B53 ASFV with selected genomes from the GenBank
| Strain name          | GenBank Acc. No. | Country     | Year | p72 genotype | Host  | % identity to Uvira B53 | Reference |
|----------------------|------------------|-------------|------|--------------|-------|-------------------------|-----------|
| Tengani 62           | AY261364         | Malawi      | 1962 | V/I          | Pig   | 86.6                    | [17]      |
| Georgia 2007/1       | FR682468         | Georgia     | 2007 | II           | Pig   | 85.4                    | [18]      |
| Kenya 1950           | AY261360         | Kenya       | 1950 | X            | Pig   | 98.8                    | [19]      |
| Kenya05/Tk1          | NC044945         | Kenya       | 2005 | X            | Tick  | 95.3                    | [20]      |
| Benin 97/1           | NC044956         | Benin       | 1997 | I            | Pig   | 86.2                    | [21]      |
| Ken06.Bus            | KM111295         | Kenya       | 2006 | IX           | Pig   | 92.2                    | [20]      |
| R35                  | MH025920         | Uganda      | 2015 | IX           | Pig   | 91.9                    | Unpublished |
| N10                  | MH025919         | Uganda      | 2015 | IX           | Pig   | 91.9                    | Unpublished |
| Pretoriskop/96/4     | AY261363         | South Africa| 1996 | XX/I        | Tick  | 86.1                    | [19]      |
| Warthog              | AY261366         | Namibia     | 1980 | IV           | Warthog | 86.5                  | [19]      |
| R8                   | MH025916         | Uganda      | 2015 | IX           | Pig   | 91.9                    | Unpublished |
| Mkuzi 1979           | NC044953         | South Africa| 1979 | I/VII       | Tick  | 85.7                    | [19]      |
| Belgium 2018/1       | LR536725         | Belgium     | 2018 | II           | Wild boar | 85.4                | [22]      |
| RSA_2_2004           | MN641877         | South Africa| 2004 | XX          | Wild boar | 84.6                | Unpublished |
| Zaire                | MN630494         | Zaire       | 2020 | XX           | Pig   | 85.4                    | Unpublished |
| Pig/China/Cas19-01/19 | MN172368        | China       | 2019 | II           | Pig   | 85.4                    | [23]      |
| 85/Ca/1985           | MN270973         | Italy       | 1985 | I            | Pig   | 86.3                    | [24]      |

**Genome comparison of ASFV Uvira B53 with reported ASFV genotype X strains**

Comparatively, with a length of 180,916 bp, the newly determined Uvira B53 genome is about 10-13 kbp shorter than the two reference genotype X strains from the GenBank, i.e. Kenya 1950 (193,886 bp) and Ken05/Tk1 (191,058 bp). However, Uvira B53 was genetically closer to Kenya 1950, a pig-derived strain exhibiting 98.8% DNA identity, than to Ken05/Tk1, a tick-derived strain with 95.3% identity (Table 2). A visual representation of the whole genome alignment of homologous genes between the ASFV genotype X generated using the Viral Orthologous Cluster V.2.0 [14] is shown in Figure 1. Sequence alignment showed that the length difference observed between these genomes is due mainly to the absence of some genes in Uvira B53 particularly the members of the 5 multigene families such as MGF 100, MGF 110, MGF 300, MGF...
360, and MGF 505. In summary, 10 genes were not present in the Uvira B53 genome including 8 multigene families (MGF 360-2L, MGF 110-4L, MGF 110-7L, MGF 110-8L, MGF 100-1R, MGF 110-9L, MGF 360-21R and MGF 360-1L), DP96R encoding the UK protein [25] and p285L of unknown functions (Table 3). More specifically, MGF 360-1L and MGF 360-21R were absent in the right terminal in the Uvira B53 strain while they were present in the other genotype X strains. Moreover, MGF 110-4L, MGF 110-7L, MGF 110-8L, MGF 110-9L and MGF 100-1R were absent in Uvira B53 strain but present in the two reference genotype X strains. In contrast, MGF-110-5L was absent in the Kenya 1950 isolate but present in the Uvira B53 and Ken05/Tk1 (Table 3).

**Table 3 ORFs present in Kenya 1950 and Ken05/Tk1 but absent in Uvira B53**

| ORF name     | Uvira B53 | Kenya 1950 | Ken05/Tk1 |
|--------------|-----------|------------|-----------|
| DP96R        | -         | +          | +         |
| MGF 100-1R   | -         | +          | +         |
| MGF 110-4L   | -         | +          | +         |
| MGF 110-7L   | -         | +          | +         |
| MGF 110-8L   | -         | +          | +         |
| MGF 110-9L   | -         | +          | +         |
| MGF 360-1L (right) | - | +          | +         |
| MGF 360-21R (right) | - | +          | +         |
| MGF 360-2L   | -         | +          | +         |
| p285L        | -         | +          | +         |
| MGF 110-5La  | +         | -          | +         |

Abbreviation: (-), ORF not present. aMGF 110-5L is present in Uvira B53 but absent in Kenya 1950.

Overall, of the 168 ORFs identified in the Uvira B53 genome, 134 shared 100% identity with homologs in Ken05/Tk1 whereas the 34 others were polymorphic (56.6 to 99.6% identity). On the other hand, Uvira B53 and Kenya 1950 also shared 167 ORFs including 136 with 100% sequence identity and 31 that were divergent (74.4% to 99.7% sequence identity). Therefore, one Uvira B53 ORFs (MGF 110-5L) was absent in Kenya 1950. Altogether, the 168 ORFs in the Uvira B53 genome could be clustered into two main groups: 131 conserved and 37 non-conserved ORFs (Table 4).

**Conserved ORFs**

The conserved category included 131 Uvira B53 ORFs for proteins showing 100% amino acid identity with the two reference ASFV genotype X analyzed. Some of them encode for structural proteins, transcription, replication and processing factors, enzymes and proteins involved in nucleotide metabolism, DNA repair. Whereas several other ORFs were classified as coding for membranes proteins from which 16 belong to the
members of MGF. Also clustered in the category of conserved ORFs were protein-coding genes A238 (an IκB-like protein), H339R (the viral protein involved in host-virus interaction), E301R (proliferating cell nuclear antigen), B263R (the TATA box binding protein), Bcl-2 A179L (the apoptosis regulating protein) and E120R (the DNA-binding structure). Furthermore, most conserved proteins included several uncharacterized ORFs such as F317L, H171R (with 100% identity between the strains) (Data not shown).

Non-conserved, variable ORFs

Sequence comparison revealed that 37 Uvira B53 ORFs were polymorphic in either Ken05/Tk1 or Kenya 1950, or both strains, with 50.6% – 99.6% sequence identity (Table 4). In comparison with Ken05/Tk1, non-conserved ORFs included seven proteins involved in the putative signal peptide and transmembrane region (B169L, C84L, CP123L, E146L, I177L, I196L, and X69R), one belonging to helicase superfamily II (A859L), the structural protein p54 (E183L), the lectin-like protein (EP153R), the CD2 homolog (EP402R), a ERCC4 predicted nuclease and potential death domain (EP364R), NifS-like PLP-dependent transferase (QP383R), twelve members of the MGF including ve of MGF 110, (2L, 5L, 6L, 11L, and 13L), two MGF 300 (2R and 4L), four MGF 360 (6L, 8L, 13L and 18R), one MGF 505 (1R) and eight ORFs of unknown functions.

With respect to Kenya 1950 strain, we identified 31 ORFs with 74.4% to 99.7% sequence identity (Table 4). One ORF (MGF 110-5L) was missing in Kenya 1950 while present in the other strains. Similarly, two ORFs (I10L and MGF 505-4R) were also absent in Ken05/Tk1 while present in the two other genotype X strains.

Table 4 Uvira B53 polymorphic ORFs with respect to Ken05/Tk1 and Kenya 1950
| Gene | Function | Ken05/Tk1 | Kenya 1950 |
|------|----------|-----------|------------|
| A859L | Helicase superfamily II | 99.4 | 100 |
| B117L | Transmembrane region containing protein | 90 | 77.1 |
| B169L | Putative signal peptide | 91.5 | 98.8 |
| B407L | Unknown | 97.3 | 99.5 |
| B475L | Unknown | 99.6 | 95.8 |
| C84L | Putative signal peptide | 98.5 | 97 |
| CP123L | Putative signal peptide | 100 | 99.2 |
| D129L | Unknown | 86 | 92.7 |
| E146L (j16L) | Putative signal peptide | 98.6 | 100 |
| E183L (p54, j13L) | Structural protein p54 | 98.4 | 91.4 |
| EP153R | Lectin-like protein | 84.3 | 79.5 |
| EP364R | Predicted nuclease and potential DEATH domain | 95.5 | 97.1 |
| EP402R (CD2v) | CD2-like protein | 76 | 96.2 |
| I10L | Unknown | 100 | 97.6 |
| I10L_2 | Unknown | 61.8 | 96 |
| I12R | Unknown | 88.3 | 78.7 |
| I177L (k14L) | Putative signaling peptide | 90.1 | 100 |
| I196L (k15L) | Putative signaling peptide | 79.4 | 94.9 |
| I8L | Unknown | 62.1 | 100 |
| I9R | Unknown | 97.9 | 99 |
| L60L | Unknown | 52.1 | 74.4 |
| O61R (p12) | Structural protein p12 | 98.4 | 96.8 |
| QP383R (j11R) | NifS-like PLP-dependent transferase | 90.3 | 99.7 |
| X69R | Putative signal peptide | 95.8 | 84.9 |
| MGF 110-11L | 110 multigene | 76.1 | 92.1 |
| MGF 110-13L-14L | 110 multigene | 96.7 | 95.3 |
| MGF 110-2L | 110 multigene | 83.7 | 95.6 |
| MGF 110-5L | 110 multigene | 98.4 | (-) |
| MGF 110-6L | 110 multigene | 65.3 | 95.1 |
Comparison of the region between I73R and I329L genes in the Uvira B53 and other genotype X strains

Previous studies have demonstrated that ASFV genotype X strains reported to date are closely related and are known to be widespread in Kenya [20]. Some small length variations among these strains’ genomes are mostly due to the number of tandem repeat sequences (TRS) either within genes or within intergenic regions. The intergenic region between I173R and I329L genes is essential for discriminating between closely related ASFV strains. In that regard, we assessed the tandem repeat sequence in the intergenic region between those two genes (region 173,611–173,760) in the Uvira B53 and the two Kenyan genotype X strains. The multiple sequence alignment revealed a significant size variation due to indels. With respect to Ken05/Tk1, Uvira B53 showed a 69 bp deletion, whereas Kenya 1950 featured a 36 bp deletion (Figure 2).

Phylogenetic analysis of the complete genomes of ASFV strains and the polymorphic genes

The genetic relationship between the ASFV strains was assessed through multiple sequence alignments of the whole complete genome sequences from 17 representative ASFV strains retrieved from the GenBank. Phylogenetic analysis grouped the viruses into different clusters corresponding to their genotypes as expected. Thus, Uvira B53 clustered with the two other ASFV p72 genotype X, Kenya 1950 and Ken05/Tk1 strains (Figure 3). The closest but distinct cluster to this genotype X group was the cluster composed of genotype IX ASFV strains (Ken06.Bus, R8, R35 and N10) whereas the most distantly related clusters concerning Uvira B53 included genotype XX with virus strains from DRC (Zaire) and South Africa (Pretoriuskop/96 and RSA_2_2004) as well as genotype IV containing the Namibian warthog strain.

Furthermore, we looked at polymorphic genes among all the 18 ASFV strains and carried out a phylogenetic analysis of the four most divergent of these genes, which included I196L, KP177L, EP153R and I177L. The Uvira B53 protein variants for I196L and I177L genes clustered with the Kenyan strains Kenya 1950 and Ken05/Tk1 from the 15 other strains analyzed (Figure 4 A and D). In contrast, protein variants encoded by the two other genes, KP177L and EP153R, showed hypervariable regions among the strains and separated the three ASFV genotype X into different clusters (Figure 4 B and C). The KP177L gene product clustered Uvira B53 (genotype X) together with Benin 97/1 and 85/Ca/1985 (genotype I). The KP177L gene was absent in the Ugandan strains (R8, R35 and N10) genotype IX. In contrast, given the low bootstrap percentage (38%) of the
node value, the EP153R protein grouped Uvira B53 with Ugandan strains (R8, R35, N10) and Kenyan strain Ken06.Bus (Figure 4C).

**Amino acid sequence comparison of the EP402R (CD2V) and serotyping**

To determine the hemadsorption inhibition (HAI) and serogroup characteristics, the protein sequence of Uvira B53 EP402R gene was compared with the ones of 13 other ASFV strains retrieved from the GenBank and representing the 8 serogroups known to date. The results revealed a high sequence variation in the CD2v protein among all the strains. The Uganda (KM609361) strain of serogroup 7 was the most closely related to Uvira B53 displaying 99% amino acid identity (the Uvira B53 CD2v is 373 amino acid long, three amino acid residues longer than its Uganda counterpart; data not shown), suggesting that Uvira B53 reported in this study belongs to serogroup 7, representing the second ASFV serogroup 7 reported to date.

The C-terminal end of CD2v is characterized by a tandem repeat sequence (TRS) of six amino acids PPPKPC. Comparative analysis of the partial TRS region showed sequence diversity due to amino acid substitutions and deletions (Figure 5). However, O-77 and STP-1 strains, both of serogroup 4, did not contain indels in the sequences thus displaced the longest TRS. For Uvira B53 and Uganda (KM609361) strains, this partial TRS was identical.

**Discussion**

In this study, we generated the first complete genome sequence of an ASFV strain, the Uvira B53, directly from a spleen tissue sample of ASF symptomatic domestic pig sourced from Uvira district in South Kivu province, eastern DRC in 2019. Overall, the assembled 180,916 bp long genome sequence obtained was of a high-quality BLAST-match with the previously reported p72 genotype X sequences available in the GenBank. The Uvira B53 ASFV strain is genetically more closely related to the Kenyan strains genotype X than to other ASF viruses of different genotypes.

Compared to the two published genotype X which are all from Kenya, the Uvira B53 genome sequence contained a group of fully conserved genes and a group of divergent genes with SNPs and/or indels (insertions/deletions). In addition, a couple of ORFs found in the two Kenyan genotype X were absent in the Uvira B53 strain: DP96R coding for the UK protein [25] and p285L with an unknown function. The absence of the DP96R gene has also been reported in the virulent Ken06.Bus strain genotype IX [20] and the p285L gene was reported in the virulent Benin 97/1 strain [21]. DP96R is among the known ASFV virulence factors [26]. Based on this information, the absence of DP96R in Uvira B53 may suggest that the virulence of the DRC strain is different from that of the Kenyan strains genotype X. However, DP96R and p285L were found in the non-virulent strain OURT 88/3 [21] as well as in the Malawian virulent strain Tegani 62 [17]. Therefore, further study and information are needed to determine the role of these genes in Uvira B53 and improve our understanding of variable levels of pathogenicity between different ASFV strains.

Like in the two published genotype X, Kenya 1950 and Ken05/Tk1, the terminal inverted repeats were missing at both ends of the Uvira B53 genome sequence. This could be due to lack of enough ASFV DNA in the spleen DNA sample sequenced. In addition, the Illumina sequencing platform used generates only short reads, which makes it difficult to effectively sequence and assemble repetitive regions as short repeats could collapse. This
limitation could be addressed by working with sufficient amount of virus DNA through tissue culture and virus isolation and as well as by using a platform that generate by long-read sequencing.

Several non-conserved divergent genes were identified to be affected by mutations including three proteins involved in putative signal peptide and transmembrane region (C262R, CP123L and E248R), one helicase superfamily II (F1055L), one AP endonuclease class II (E296R) and one RNA polymerase subunit 3 (H359L). However, the effects of these substitutions on protein function are unknown.

Several genetic variations including SNPs, indels (insertions/deletions) and complete loss of ORFs were observed within the MGF member genes between Uvira B53 and known genotype X strains, specifically in genes previously reported to be implicated in determining host range and virulence such as MGF 360, MGF 505 and MGF 110. Indels and loss of ORFs were among the major causes of the differences in length between these genomes as previously reported [27]. They were observed in all the three strains genotype X analyzed but at higher magnitude in Uvira B53 than the two other strains, making the newly completed 180,916 bp long Uvira B53 genome 10 and 13 kbp shorter than the Kenya 1950 and Ken05/Tk1 genomes, respectively. The biological implications of these variations remain to be determined.

The high level of genetic conservation was observed in some ORFs of MGF 360 (9L, 10L, 13L, 15R and 16) and MGF 505 (3R, 4R and 10R) among Uvira B53 and Kenya 1950, a virulent strain [20] whereas they were absent in the Ken05/Tk1, an avirulent strain [20]. This feature may suggest difference and similarity in the Uvira B53 pathogenicity with Ken05/Tk1 and Kenya 1950, respectively. This hypothesis is in agreement with a previous study that demonstrated that some members of MGF 360 and MGF 505 such as MGF360-12L, -13L, and -14L and MGF505-1R, -2R, and -3R were associated with ASFV host range specificity, blocking of the host innate response, and virus virulence, as they were present in several virulent strains of ASFV while absent in the avirulent strains like OURT 88/3 and NHV [28, 29, 30]. Further analysis of these differences should provide more insights into the molecular mechanisms underlying the virulence of Uvira B53 genotype X in pigs in South Kivu.

The comparative study based of the intergenic region between I73R and I329L genes showed the genome sequence of the Uvira B53 strain was closer to Kenya 1950 than Ken05/Tk1, as both the Uvira B53 and Kenya 1950 contained deletions compared to Ken05/Tk1. However, the Uvira B53 deletion was 33 bp longer than that of Kenya 1950. This observation corroborates our previous comparative analysis of this intergenic region which revealed this deletion in genotype X strains from symptomatic pigs in South Kivu, DRC [11].

Analysis of phylogenetic relationship was performed using some representative sequence data sets retrieved from the GenBank. The analysis clustered the ASFV Uvira B53 together with genotype X strains reported in Kenya. This result is in accordance with our previous study reporting circulation of ASFV genotype X in symptomatic ASFV infected pigs from South Kivu province [11]. Uvira B53 is different from genotype IX, I and XIV that have recently been reported in some provinces of DRC such as Bandundu, Equateur, Katanga, Kinshasa, Maniema and Province Orientale [10]. They are sylvatic cycle-associated genotype that has been identified from ticks, warthog and domestic pigs in Kenya, Tanzania and Burundi [5, 31, 32]. The phylogenetic analysis of some of the divergent protein-coding gene products, particularly the I196L and I177L, matched with the phylogenetic relationship of the complete genome sequence that grouped the Uvira B53 with Kenya
1950 and Ken05/Tk1 genotype X strains. Two other gene products (EP153R and KP177L) were highly divergent in all the strains. A similar observation was reported by Chapman et al. [18] where phylogenetic analyses of several divergent gene products did not match with one of the complete genome sequences. The divergence may probably be due to the occurrence of recombination events and antigenic genome diversity. Because of this, we suggest to be cautious when performing phylogenetic relationships between ASFV strains based on a small number of genes.

Several variations were identified in the C-type lectin-like protein (EP153R) in all the strains, making it one of the most divergent genes. Variations in this protein may alter the expression of the functional protein (C-type lectin) that may influence the immuno-modulatory functions as previously demonstrated [33].

In addition to SNPs, variations in this region of CD2v (EP402R) were characterized by the presence of the tandem repeat amino acid sequence PPPKPY [34] that was repeated 3-11 times among the strains analyzed. This CD2v variable region is identical in the Uganda (KM609361) strain of serogroup 7 and Uvira B53 with 5 tandem repeats. BLAST search also revealed that the Uganda (KM609361) strain’s CD2v antigen was the most closely related antigen to Uvira B53’s CD2v, displaying 99% amino acid identity. This finding implied that Uvira B53 strain belongs to the ASFV serogroup 7. This is the first report of a complete genome sequence of ASFV of the serogroup 7. In addition to the serogroups 1 and 2 already reported in DRC [35], ASFV Uvira B53 genotype X and serogroup 7 represents the third ASFV serogroup to be found in DRC.

**Conclusion**

African swine fever outbreaks were reported in South Kivu in 2019 causing mortality in domestic pigs. We report for the first time the complete genome sequence of the ASFV Uvira B53, a genotype X and serogroup 7 strain involved in outbreaks in South Kivu. Its similarity with Kenya 1950 strain based on MGF members suggests that it is a virulent strain, corroborated by the symptomatic phenotype of the pig from where the sequence originated. The new information reported in this study provides valuable insights for tracing geographical spread and biological evolution of the virus, which are essential for the design of prevention and control strategies of this disease.

**List Of Abbreviations**

ASF: African swine fever; ASFV: African swine fever virus; BLAST: Basic Local Alignment Search Tool; DRC: Democratic Republic of Congo; DNA: deoxyribonucleic acid; GATU: The genome annotation transfer utility; Kbp: kilo base pair; MGF: Multigene family; NCBI: National Center for Biotechnology Information; ORF: open reading frame; PAUISTI: Pan African University Institute of Science Technology and Innovation; TRS: tandem repeat sequence; UEA: Université Evangélique en Afrique.

**Declarations**

**Availability of data and material**
The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

**Ethics approval and consent to participate**

See Materials and Methods section.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

**Funding**

This work was funded by the Biosciences Eastern and Central Africa - International Livestock Research Institute (BecA-ILRI) Hub, through Africa Biosciences Challenge Fund (ABCF).

**Authors’ contributions**

BNP collected samples for DNA isolation, performed laboratory work and wrote the manuscript; JRO supervised the study, conceived and designed the study, and edited the manuscript; LS supervised the study and edited the manuscript; EBB conceived and designed the study and edited the manuscript; EB performed the laboratory work and sequence analysis; EMM performed the laboratory work and sequence analysis; JBE performed sequence analysis, wrote and edited the manuscript; EO performed sequence analysis, wrote and edited the manuscript, RP supervised the study, performed sequence analysis, wrote and edited the manuscript. All authors read and approved the manuscript.

**Acknowledgements**

We acknowledge the Biosciences eastern and central Africa - International Livestock Research Institute (BecA-ILRI) Hub, through the Africa Biosciences Challenge Fund (ABCF) program for funding this work. The ABCF Program is funded by the Australian Department for Foreign Affairs and Trade (DFAT) through the BecA-CSIRO partnership; the Syngenta Foundation for Sustainable Agriculture (SFSA); the Bill & Melinda Gates Foundation (BMGF); the UK Department for International Development (DFID) and; the Swedish International Development Cooperation Agency (Sida). The work was also supported by the Wellcome Trust (Grant 105684/Z/14/Z). We are grateful to the African Union (AU) through Pan African University Institute of Basic Sciences, Technology and Innovation (PAUSTI) for the PhD scholarship. We extend our gratitude to Université Evangélique en Afrique (UEA) for funding our fieldwork.

**References**

1. Costard S, Wieland B, De Glanville W, Jori F, Rowlands R, Vosloo W, Roger F, Pfeiffer DU, Dixon LK: African swine fever: how can global spread be prevented? Philosophical transactions of the Royal Society of
2. Alonso, C., Borca, M., Dixon, L., Revilla, Y., Rodriguez, F., Escribano, J.M., Ictv Report, C., 2018. ICTV virus taxonomy profile: asfarviridae. Journal of General Virology, 99:613–614.

3. Dixon, L.K., Chapman, D.A.G., Netherton, C.L., Upton, C., 2013. African swine fever virus replication and genomics. Virus Res. 173, 3–14. https://doi.org/10.1016/j.virusres.2012.10.020

4. Anderson, E.C., Hutchings, G.H., Mukarati, N., Wilkinson, P.J., 1998. African swine fever virus infection of the bushpig (Potamochoerus porcus) and its significance in the epidemiology of the disease. Vet. Microbiol. 62, 1–15.

5. Bastos, A.D., Penrith, M.L., Crucière, C., Edrich, J.L., Hutchings, G., Roger, F., Couacy-Hymann, E., Thomson, R.G., 2003. Genotyping field strains of African swine fever virus by partial p72 gene characterisation. Arch. Virol. 148, 693–706. https://doi.org/10.1007/s00705-002-0946-8.

6. Sánchez-Vizcaíno, J., Mur, L., Gomez-Villamandos, J., Carrassco, L. (2015). An update on the epidemiology and pathology of African swine fever. Journal of Comparative Pathology, 152 (1):9–21.

7. Achenbach JE, Gallardo C, Nieto-Pelegrín E, Rivera-Arroyo B, Degefa-Negi T, Arias M, Jenberie S, Mulisa DD, Gizaw D, Gelaye E, Chibssa TR, Belaye A, Loitsch A, Forsa M, Yami M, Diallo A, Soler A, Lamien CE, Sánchez-Vizcaíno JM: Identification of a new genotype of African swine fever virus in domestic pigs from Ethiopia. Transbound Emerg Dis. 2017; 64:1393–1404.

8. Quembo, C.J., Jori, F., Vosloo, W., Heath, L., 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. Transbound. Emerg. Dis. 65, 420–431. https://doi.org/10.1111/tbed.12700.

9. Bisimwa, N.P., Machuka, M.E, Amimo, O.J, Dedan, G., Banswe, G, Masembe, C., Ongus, R.J., Steinaa, L., Pelle, R., 2019. Evidence for the Presence of African Swine fever virus in apparently healthy pigs in South Kivu province of the Democratic Republic of Congo. Vet. Microbiol. 240, 108521.

10. Mulumba–Mfumu, L.K., Achenbach, E.J., Mauldin, M.R., Dixon, L.K., Tshilenge, C.G., Thiry, E., Moreno, N., Blanco, E., Saegerman, C., Lamien, C.E., Diallo, A., 2017. Genetic Assessment of African Swine Fever Isolates Involved in Outbreaks in the Democratic Republic of Congo between 2005 and 2012 Reveals Co-Circulation of p72 Genotypes I, IX and XIV, Including 19 Variants. Viruses. 9(2), 31. doi:10.3390/v9020031.

11. Bisimwa, N.P., Ongus, R.J., Tiambo, K.C., Machuka, M.E, Bisimwa, B.E, Steinaa, L., Pelle, R., 2020. First detection of African swine fever (ASF) virus genotype X and serogroup 7 in the Democratic Republic of Congo through survey of symptomatic pigs. BMC Virol. J. 17, 135. https://doi.org/10.1186/s12985-020-01398-8.

12. de Villiers, E.P., Gallardo, C., Arias, M., da Silva, M., Upton, C., Martin, R., Bishop, R.P., 2010. Phylogenomic Analysis of 11 Complete African Swine Fever Virus Genome Sequences. Virol. 400, 128–136.

13. Aguero, M., Fernandez, J., Romero, L., Sanchez-Mascaraque, C., Arias, M., Sanchez-Vizcaino J.M., 2003. Highly sensitive PCR assay for routine diagnosis of African swine fever virus in clinical samples. J. Clin. Microbiol. 41 (9), 4431–4434.

14. Tcherepanov, V., Ehlers, A., Upton, C., 2006. Genome Annotation Transfer Utility (GATU): rapid annotation of viral genomes using a closely related reference genome. BMC Genomics 7, 150.
15. Thompson, J., Higgins, D., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

16. Kumar, S., Stecher, G., Tamura, K., 2016. Molecular evolutionary genetic analysis version 7.0 for bigger datasets. Mol. Biol. and evol. 33, 1870-1874.

17. Pan, I.C., 1992. African swine fever virus: Generation of subpopulations with altered immunogenicity and virulence following passage in cell cultures. The J. Vet. Med. Sci. 54, 43–52. https://doi.org/10.1292/jvms.54.43.

18. Chapman, D.A.G., Darby, A.C., Da Silva, M., Upton, C., Radford, A.D., Dixon, L.K., 2011. Genomic analysis of highly virulent isolate of African swine fever virus. Emerg. Infect. Dis. 17(4), 599–605. DOI: 10.3201/eid1704.101283.

19. Zsak, L., Borca, M.V., Risatti, G.R., Zsak, A., French, R.A., Lu, Z., Rock, D.L., 2005. Preclinical diagnosis of African swine fever in contact-exposed swine by a real-time PCR assay. Journal of Clinical Microbiology, 43, 112–119. https://doi.org/10.1128/JCM.43.1.112-119.2005.

20. Bishop, R.P., Fleischauer, C., de Villiers, E.P., Okoth, E.A., Arias, M., Gallardo, C., Upton, C., 2015. Comparative analysis of the complete genome sequences of Kenyan African swine fever virus isolates within p72 genotypes IX and X. Virus Gene. 50(2), 303–9.

21. Chapman, D.A.G., Tcherepanov, V., Upton, C., Dixon, L.K. (2008). Comparison of the genome sequences of nonpathogenic and pathogenic African swine fever virus isolates. Journal of General Virology, 89, 397–408.

22. Pikalo, J., Schoder, M., Sehl, J., Breithaupt, A., Tignon, M., Cay, A.B., Gager, A.M., Fischer, M., Beer, M., Blome, S., 2020. The African Swine Fever Virus Isolate Belgium 2018/1 Shows High Virulence in European Wild Boar. Transbound. Emerg. Dis. 67(4), 1654-1659.

23. Jia, L., Jiang, M., Wu, K., Hu, J., Wang, Y., Quan, W., Hao, M., Liu, H., Wei, H., Fan, W., Liu, W., Hu, R., Wang, D., Li, J., Chen, J., Liu, D., 2020. Nanopore sequencing of African swine fever virus. Sci. China Life Sci. 63, 160–164. https://doi.org/10.1007/s11427-019-9828-1.

24. Torresi, C., Fiori, M., Bertolotti, L., Floris, M., Colitti, B., Giammarioli, M., Giudici, S.D., Oggiano, A., Malmberg, M., De Mia, G.M., Belák, S., Granberg, F., 2020. The Evolution of African Swine Fever Virus in Sardinia (1978-2014) as Revealed by Whole-Genome Sequencing and Comparative Analysis. Transbound. Emerg. Dis. 00, 1–10. DOI: 10.1111/tbed.13540.

25. Zsak, L., Caler, E., Lu, Z., Kutish, G.F., Neilan, J.G., Rock, D.L., 1998. A nonessential African swine fever virus gene UK is a significant virulence determinant in domestic swine. J. Virol. 72(2), 1028-1035.

26. Rodriguez, J.M., Moreno, L.T., Alejo, A., Lacasta, A., Rodriguez, F., Salas, M.L., 2015. Genome Sequence of African Swine Fever Virus BA71, the Virulent Parental Strain of the Nonpathogenic and Tissue-Culture Adapted BA71V. PLoS ONE 10(11):e0142889. doi:10.1371/journal.pone.014288.

27. Dixon, L.K., Chapman, D.A.G., Netherton, C.L., Upton, C., 2013. African swine fever virus replication and genomics. Virus Res. 173, 3–14. https://doi.org/10.1016/j.virusres.2012.10.020.
28. Zsak, L., Lu, Z., Burrage, T.G., Neilan, J.G., Kutish, G.F., Moore, D.M., Rock, D.L., 2001. African Swine Fever Virus Multigene Family 360 and 530 Genes Are Novel Macrophage Host Range Determinants. J. Virol. 7, 3066–3076.

29. Burrage, T.G., Lu, Z., Neilan, J.G., Rock, D.L., Zsak, L., 2004. African Swine Fever Virus Multigene Family 360 Genes Affect Virus Replication and Generalization of Infection in Ornithodoros porcinus Ticks. J. Virol. 78, 2445–2453.

30. O’Donnell, V., Holinka, L.G., Gladue, D.P., Sanford, B., Krug, P.W., Lu, X., Arzt, J., Reese, B., Carrillo, C., Risatti, G.R., Borca, M.V., 2015. African swine fever virus Georgia isolate harboring deletions of MGF360 and MGF505 genes is attenuated in swine and confers protection against challenge with virulent parental virus. J. Virol. 89, 6048–6056. doi:10.1128/JVI.00554-15.

31. Lubisi, B.A., Bastos, A.D., Dwarka., R.M., Vosloo, W., 2005. Molecular epidemiology of African swine fever in East Africa. Arch. Virol. 150(12), 2439-2452.

32. Gallardo, C., Okoth, E., Pelayo, V., Anchuelo, R., Martin, E., Simon, A., Llorente, A., Nieto, R., Soler, A., Martin, R., Arias, M., Bishop, R.P., 2011. African swine fever viruses with two different genotypes, both of which occur in domestic pigs, are associated with ticks and adult warthogs, respectively, at a single geographical site. J. Gen. Virol. 92, 432–444.

33. Hurtado, C., Granja, A.G., Bustos, M.J., Nogal, M.L., Gonzalez, de Buitrago, G., de Yebenes, V.G., Salas, M.L., Revilla, Y., Carrascosa, A.L., 2004. The C-type lectin homologue gene (EP153R) of African swine fever virus inhibits apoptosis both in virus infection and in heterologous expression. Virol. 326, 160–170.

34. Sanna, G., Dei Giudici, S., Bacci, D., Angioi, P. P., Giammarioli, M., De Mia, G. M., Oggiaro, A. (2017). Improved strategy for molecular characterization of African swine fever viruses from Sardinia based on analysis of p30, CD2V and I73R/I329L variable regions. Transboundary and Emerging Diseases, 64, 1280–1286.

35. Malogolovkin, A., Burmakina, G., Tulman, E.R., Delhon, G., Diel, D.G., Salnikov, N., Kutish, G.F., Kolbasov, D., Rock, D.L., 2015. African swine fever virus CD2v and C-type lectin gene loci mediate serological specificity. J. Gen. Virol. 96, 866–73.