Original Article

Complete Genome Sequencing, Annotation, and Mutational Profiling of the Novel Clade I Human Mpox Virus, Kamituga Strain

Leandre Murhula Masirika[1,2 #], Anuj Kumar[3,4,5 #], Mansi Dutt[3,4,5 #], Ali Toloue Ostadgavahi[3,4], Benjamin Hewins[3,4], Maliyamungu Bubala Nadine[6], Bilembo Kitwanda Steeven[6], Franklin Kumbana Mweshi[6], Léandre Mutimbwa Mambo[7], Justin Bengehya Mbiribindi[8], Freddy Belesi Siangoli[8], Alyson A Kelvin[9], Jean Claude Udahemuka[10], Patricia Kelvin[3,5], Luis Flores[1,11,12,13], David J Kelvin[3,4,5], Gustavo Sganzerla Martinez[3,4,5]

[1] Centre de Recherche en Sciences Naturelles de Lwiro, South Kivu, DS Bukavu, Democratic Republic of the Congo
[2] Health and Biotechnology, (SaBio), Institute for Game and Wildlife Research (IREC), Castilla-LaMancha University & CSIC, Ciudad Real, Spain
[3] Department of Microbiology and Immunology, Canadian Center for Vaccinology (CCfV), Faculty of Medicine, Dalhousie University, Halifax, Canada
[4] Laboratory of Immunity, Shantou University Medical College, Jinping, Shantou, Guangdong, China
[5] BioForge Canada Limited. Halifax, NS, Canada.
[6] Hospital General de Reference de Kamituga, South Kivu, Democratic Republic of the Congo
[7] Zone de Santé de Kamituga, South Kivu, Democratic Republic of the Congo
[8] Division Provinciale de la Santé, South Kivu, Democratic Republic of the Congo
[9] Vaccine and Infectious Disease Organization (VIDO), University of Saskatchewan, Saskatoon, Canada
[10] Department of Veterinary Medicine, University of Rwanda, Nyagatare, Rwanda
[11] Volunteers for the Conservation of Fauna and Flora (VCFF)
[12] One Health Conservation Initiative, Katana, Kabare, South Kivu, Democratic Republic of the Congo
[13] Centre de Rehabilitation des Primates de Lwiro, Lwiro, South Kivu, DS Bukavu, Democratic Republic of the Congo

# Authors contributed equally to this work.

Abstract

Introduction: Human Mpox (formerly monkeypox) infection is an emerging zoonotic disease caused by the Mpox virus (MPXV). We describe the complete genome annotation, phylogeny, and mutational profile of a novel, sustained Clade I Mpox outbreak in the city of Kamituga in Eastern Democratic Republic of the Congo (DRC).

Methodology: A cross-sectional, observational, cohort study was performed among patients of all ages admitted to the Kamituga Hospital with Mpox infection symptoms between late September 2023 and late January 2024. DNA was isolated from Mpox swabbed lesions and sequenced followed by phylogenetic analysis, genome annotation, and mutational profiling.

Results: We describe an ongoing Clade I Mpox outbreak in the city of Kamituga, South Kivu Province, Democratic Republic of Congo. Whole-genome sequencing of the viral RNA samples revealed, on average, 201.5 snps, 28 insertions, 81 deletions, 2 indels, 312.5 total variants, 158.3 amino acid changes, 81.66 intergenic variants, 72.16 synonymous mutations, 106 missense variants, 41.16 frameshift variants, and 3.33 inframe deletions across six samples. By assigning mutations at the proteome level for Kamituga MPXV sequences, we observed that seven proteins, namely, C9L (OPG047), I4L (OPG080), L6R (OPG105), A17L (OPG143), A25R (OPG151), A28L (OPG153), and B21R (OPG210) have emerged as hot spot mutations based on the consensuses inframe deletions, frameshift variants, synonymous variants, and amino acids substitutions. Based on the outcome of the annotation, we found a deletion of the *D14L (OPG032)* gene in all six samples. Following phylogenetic analysis and whole genome assembly, we determined that this cluster of Mpox infections is genetically distinct from previously reported Clade I outbreaks, and thus propose that the Kamituga Mpox outbreak represents a novel subgroup (subgroup VI) of Clade I MPXV.

Conclusions: Here we report the complete viral genome for the ongoing Clade I Mpox Kamituga outbreak for the first time. This outbreak presents a distinct mutational profile from previously sequenced Clade I MPXV oubtreaks, suggesting that this cluster of infections is a novel subgroup (we term this subgroup VI). These findings underscore the need for ongoing vigilance and continued sequencing of novel Mpox threats in endemic regions.

Key words: DRC, Mpox, Kamituga outbreak, hMPXV, Clade I, Sequencing, South Kivu

Introduction

Mpox (formerly monkeypox) is an emerging zoonotic disease caused by the Mpox virus (MPXV) first isolated in 1958 from smallpox-like lesions in cynomolgus monkeys [1-4]. The first human infection caused by Mpox was diagnosed in a boy from the Democratic Republic of the Congo (DRC) in 1970 [5]. MPXV is endemic in various regions of Central and West Africa and outbreaks are sporadic in frequency [3, 6]. The transmission of Mpox predominantly occurs

between animals (small mammals, rodents, and monkeys), between animals and humans, and via human-to-human contact with infectious individuals [7]. The route of Mpox transmission also depends on the specific Clade responsible for infection. Symptoms associated with Mpox infection also vary depending on the Clade, but Mpox generally presents as a febrile, self-limiting disease with headache, myalgia, back pain, lethargy, lymphadenopathy, and rash, which may progress to larger ulcers in areas of the face, genitals, and eyes during instances of severe infection [4].

MPXV contains a linear, double-stranded DNA genome and belongs to the genus orthopoxvirus, poxviridae family, and subfamily chordopoxvirinae. The MPXV genome is roughly 200 kb in length (encoding ~200 genes) and encodes for an estimated 190 proteins [8]. MPXV is classified into two distinct Clades: Clade I (formerly the Congo Basin clade-Central Africa), and Clade II (West African clade). Clade II may be further divided into two subclades, termed IIa and IIb[9]. Subclade IIb was responsible for the global Mpox outbreak in 2022, which was primarily described in the men who have sex with men (MSM) population and for individuals who engage in sexual activity with infectious individuals [10]. Clade I MPXV is associated with more severe symptoms and a high case fatality rate (CFR) of around 10.6%, while Clade II is associated with a lower CFR of roughly 3.6% [11]. Clade I and IIa MPXV exhibits very little human-to-human spread and is mainly transmitted zoonotically by rodents and small mammals, whereas Clade IIb is efficiently transmitted between humans via both sexual and non-sexual contact with an infected individual [7].

In this study, we describe an ongoing Clade I Mpox outbreak in the city of Kamituga, South Kivu Province, Democratic Republic of the Congo (DRC). The Kamituga outbreak is unique for the following reasons: i) MPXV is primarily being spread via heterosexual contact with infectious individuals; ii) community spread has been observed among a broad age range, including children; and, iii) sequence data from the Kamituga outbreak is genetically distinct from previously sequenced Clade I MPXV. To our knowledge, this is the largest self-sustaining Clade I Mpox outbreak on record. It is therefore critical to better understand the transmission patterns and genomic evolution of the Kamituga Mpox outbreak. We performed next-generation sequencing (NGS), revealing the complete viral genome of samples collected from infected individuals during the MPXV/Kamituga outbreak. Following a phylogenetic analysis, mutational profiling, and genome annotation, we determined that the Kamituga Mpox cluster is a novel, Clade I subgroup (subgroup VI) of MPXV.

Methodology

Study design and population

This prospective, observational cohort study was conducted at the Kamituga hospital, in the city of Kamituga, South Kivu Province, DRC. Individuals who presented with Mpox infection symptoms were admitted to the Kamituga hospital from September 2023 to January 2024. Importantly, the Kamituga outbreak is ongoing and has not yet been resolved. All prospective study participants were introduced to the study and provided the opportunity to participate following an informed consent procedure. Admission to the Kamituga hospital was based on the clinical diagnosis of human MPXV infection, which was achieved using reverse transcriptase polymerase chain reaction (RT-PCR) at the Lwiro Laboratory (Centre de Recherche en Sciences Naturelles de Lwiro, CRSN-Lwiro). Diagnosis was also based on symptoms present at the hospital visit and the epidemiological link to confirmed or probable Mpox cases. Rash was also observed and skin lesions were defined as a single, circumscribed area with presentations consistent with early vesicles, small pustules, umbilicated pustule, ulcerated lesion, or fluid-filled vesicles, with or without changes in skin texture, colour, or inflammatory changes. A total of fifty-one individuals were enrolled in this study and 37 (73%) of these participants were PCR diagnosed for Mpox infection. Swab samples were obtained from the skin lesions of enrolled participants and DNA extraction was performed. Extracted DNA was quantified and reverse transcription quantitative PCR (RT-qPCR) was performed for obtaining cycle threshold (CT) values and library preparation for NGS.

Sample Collection and DNA Extraction

Samples were collected from lesion swabs from suspected Mpox virus-infected patients following informed consent. DNA extraction was performed using the Qiagen DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany). Following extraction, DNA concentration was quantified using the Qubit 1X dsDNA HS Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA) on a Qubit 4 Fluorometer, following the manufacturer's protocol with a 1 µL sample volume. This high-sensitivity assay was crucial for the accurate quantification of DNA, even at low concentrations, setting the stage for successful library preparation.

Library Preparation, Barcoding, and Sequencing

All sequencing was performed in Lwiro, DRC. Library preparation was conducted using two kits: the Rapid Sequencing DNA - PCR Barcoding Kit 24 V14 (SQK-RPB114.24) for R10 flow cells and the Rapid PCR Barcoding Kit 24 (Version 14, SQK-RPB004) for R9 flow cells (Oxford Nanopore Technologies). The procedure involved initial DNA tagmentation; barcode addition using rapid barcode primers; PCR amplification; and library cleanup with AMPure XP beads (Beckman Coulter), followed by a final quantification step using the Qubit fluorometer. Proper flow cell priming and library loading were conducted per ONT protocols. Sequencing runs were extended to 24 hours with real-time basecalling enabled through MinKNOW (version 23.11.5, Oxford Nanopore Technologies) for live monitoring and data quality assessment.

Sequence assembly, genome analysis, and phylogeny

After the basecalling, generated reads were retrieved in FASTQ format with the help of MinKNOW software (version 23.11.5, Oxford Nanopore Technologies). The fastp tool [12] was used for quality control to analyze the obtained reads to trim adapter sequences, and for other quality filtering. After the quality check of generated reads, an alignment with the reference genome (NC_003310.1) was performed using the 'Mininmap2' tool (version 2.26-r1175) and a SAM file was created. The produced SAM file was converted into BAM file format with the aid of samtools (v1.19.2). The VCF file was generated to perform the variant calling using the mpileup function embedded in Bcftools (version1.19). The output of the bcftools was compressed and indexed using tabix (v1.19.1) in the process of obtaining a consensus sequence. The MAFFT program (version 7) [13] was employed to perform the multiple sequence alignment (MSA) between the obtained consensus sequence and publicly available MPXV genomes for Clade I and Clade II on NCBI [14] and GISAID databases. After getting the MSA, two phylogenetic trees were calculated using the neighbor-joining method utilizing the Jukes-Cantor model. Both constructed trees were calculated for a percentage of 1000 bootstrap values iterations. Molecular Evolutionary Genetics Analysis (MEGA 6.0) software [15] was used to render the constructed trees.

Mutational profiling and annotation

The "Monkeypox virus analysis Platform" of the Monkeypox Virus Resource (MpoxVR) [16] was used to identify the mutation pattern(s) and to annotate the genome sequences representing the ongoing Kamituga outbreak. The NC_003310 genome was used as a reference to predict the different types of mutations (SNPs, intergenic variants, synonymous variants, missense variants, etc.) in the Kamituga outbreak sequences.

Ethical statement

Ethical clearance for conducting this study was approved by the Ethical Review Committee of the Catholic University of Bukavu (Number UCB/CIES/NC/022/2023). All study participants were introduced to the observational study and given the option to participate by providing informed consent, or in the case that the participant was a minor, parental permission or consent was obtained.

Results

Sequencing

A nanopore (Oxford Nanopore Technologies, Oxford, UK) sequencing strategy was used, and a total of 964.63K, 2.07M, 3M, 1.29M, 1.96M, and 1.58M raw reads were generated for samples 1, 2, 3, 4, 5, and 6, respectively. The first three sequencing runs have previously been reported in Murhula et al., 2024 (submitted). After basecalling the raw reads with a minimum Q score of 8, a total of 19102, 65741, 57225, 7372, 93103, and 44973 reads were aligned to the Mpox Clade I reference genome (i.e., NC_003310), respectively. The aligned reads generated a coverage of 203.03, 1021.43, 1012.40, 110.6, 1107.96, and 669.209 folds over the entire Mpox Clade I genome.

Mutation map of Kamituga MPXV

To identify mutations of the MPXV genome during the Kamituga outbreak, mapping to the reference genome (NC_003310.1) was performed in a genome-wide manner. Our mutational analysis revealed several SNPs, insertions, deletions, indels, amino acid (aa) changes, intergenic variants, synonymous variants, missense variants, frameshift variants, and inframe deletions across all six samples. Earlier samples collected during the outbreak (samples 1-3) showed a greater number of genome changes compared to samples (samples 4-6) collected later in the outbreak (Figure 1).

Our mutational analysis revealed that the Kamituga subgroup contains a set of highly mutative proteins when compared with the reference genome (NC_003310) (Figure 2).

Among these mutations, seven proteins (including C9L [OPG047], I4L [OPG080], L6R [OPG105], A17L [OPG143], A25R [OPG151], A28L [OPG153], and B21R [OPG210]) emerged as mutational hotspots with a number of consensus inframe deletions, frameshift variants, synonymous variants, and amino acids substitutions (Supplementary Table 1).

C9L (OPGO47), a Kelch-like protein, was found to have an inframe deletion (DGMG483-486E) and one frameshift variant (D483X) with three consensus amino acid substitutions (L151H, V127E, and Y114N). When compared with subgroup V (JX878417), this protein was found to have a unique inframe deletion. I4L (OPG080), also known as the ribonucleotide reductase large subunit, was found to have two consensus synonymous variants (T372 and T31) with three amino acid substitutions (A685P, R358G, and G260S). The *L6R (OPG105)* gene codes for an RNA polymerase subunit (RPO147); our analysis revealed that this protein contained one consensus frameshift variant (P285X), one synonymous variant (A396), and three amino acid substitutions (V431A, R475C, and D506Y). A17L (OPG143) contained two consensus synonymous variants (E110 and R127) and three amino acid

substitutions (A289T, and V304A, V322A). *A25R (OPG151)* gene encodes the RNA polymerase and the 132 kDa subunit (RPO132) demonstrated a total of seven consensus changes. Among them, S220, L360, P903, and T910 have been observed as synonymous variants, while R273Q, T601M, and D633E are annotated as amino acid substitutions. A seven-residue frameshift deletion (DDDDDII367-373-), frameshift variants (D384X and D386X and DD386-387X), and two amino acid substitutions were observed in the *A28L(OPG153)*, which is known to code the P4c precursor. Further, among these seven proteins, B21R (OPG210), which codes a surface glycoprotein, was found to be associated with the highest number of mutations including two consensus frameshift variants (K431X and P726X), synonymous variants (A1207 and C1715), and nine amino acid substitutions (A475V, S533L, T713P, T723P, E1056D, A1339T, T1372A, V1725S, and D1862E). In the context of B21R, the number of mutations was found to be consistent across all six samples. When compared with the sub-group V genome, we found that the B21R (OPG210) protein of the Kamituga strain contains a total of four unique

Figure 1 (A). Graphical representation of calculated mutation patterns in the Kamituga sequences. The SNPs, insertions, deletions, indels, total variants, aa changes, intergenic variants, synonymous variants, missense variants, frameshift variants, and inframe deletions are indicated by dark blue, orange, red, gray, purple, green, sky blue, brown, black, yellow and pink colors, respectively; **(B).** Venn diagram depicting the common and unique mutative proteins among all six samples; and **(C).** Calculated counts of categorized (amino acid substitution, frameshift variant, inframe deletion, and synonymous) mutations across the seven genes (X-axis). The mutations are consensual between six MPXV Clade I subgroup VI samples.
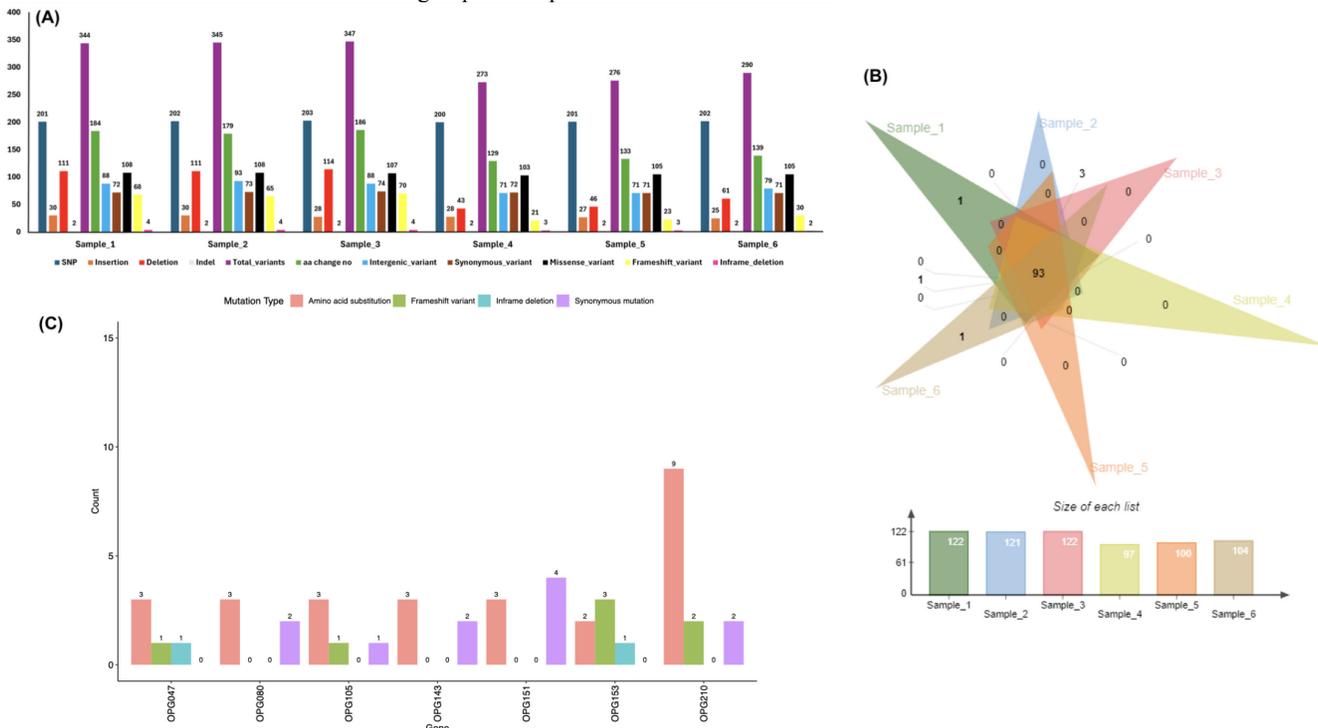
changes, two frameshift variants (K431X and A475V), and two amino acid substitutions (P726X and E1056D).

Based on our annotation analysis, we found that fourteen genes (*D2L, D4L, D17L, D18L, C3L, A26L, A27L, A48R, B15L, B1R, B18R, K1R, R1R* and *N2R)* were present in our samples but not in the reference genome. When we compared the genomic coordinates of these genes with reference coordinates, we found that they are identical to various miscellaneous features described in the RefSeq file of the reference genome with orthopox group gene names (i.e *OPG016*). As shown in Figure 2, we are describing these miscellaneous features as "new/rare" features, according to the standard definition by ebi [17]. Out of these fourteen candidates, five (*D2L, A27L, B1R, K1R,* and *R1R*) genes also have consensus mutations in all six samples with a few synonymous variants and amino acid substitutions (Figure 2). Of note, we found that the gene *C12L* codes for a hypothetical protein that was missing in our mapped annotation files, but when we performed the sequence and read the alignments, we found it matching with our reads.

Phylogenetic analysis

To rapidly investigate the phylogeny of the MPXV consensus sequences representing the Kamituga outbreak, we constructed two phylogenetic trees (one for all clades and one for Clade I sub-groups) by integrating six MPXV sequences from the Kamituga outbreak into the publicly available MPXV genome sequences from NCBI and GISAID. Six MPXV consensus sequences from the Kamituga outbreak exhibited 99.47 to 99.79**%** identity with the reference genome (NC_003310) based on the local pairwise sequence alignment. As depicted in Figure 3, all six Kamituga sequences belonged to Cluster I, forming an individual cluster (green). This cluster differed from those of MPXV strains recently reported in the DRC (23-MPX-0099 bwa 3X, 23MPX383V bwa 3X, and 23-MPX-0098 bwa 3X) and had the closest evolutionary relationship with the subgroup V sequence (JX878417) compared to other sub-groups (Figure 3). The separation of the Kamituga sequences from other Clade I members may be due to the substantial number of frameshift mutations, insertions, and deletions. Taken

Figure 2. Mutation map of the Kamituga MPXV strain: A schematic representation of the MPXV genome structure based on the reference genome (NC_003310). Gene names and positions are given per orthopoxvirus gene nomenclature (ex. OPG001) and updated MPXV RefSeq (J1L) annotation. Gene orientation is reflected in the arrowheads. Missense, synonymous, and protein-altering variant names are labeled with red, blue, and pink colors, respectively. Inframe deletions are indicated with a red circle, while frameshift variants are denoted by a blue rectangle. A black line with a circle on the top indicates the new/rare features per the EBI definition. Genes are colored based on their biological functions. Genes involved in immunomodulation, transcription, viral replication/repair, cell entry/attachment, morphogenesis, virulence, cell-to-cell fusion, hypothetical, and unknown are shown in green, orange, purple, yellow, rose, dark blue, light blue, black, and light gold, respectively. The deleted Clade I specific OPG032 gene is shown with a red arrowhead.

together, our phylogenetic analysis suggests that the ongoing Kamituga outbreak is spreading by a novel strain of MPXV, which can be taxonomically proposed as a novel sub-group VI of Clade I.

Discussion

This article is the first to describe the underlying genomic variability of the novel Kamituga Clade I Mpox outbreak in the city of Kamituga, South Kivu Province, DRC. We highlight the distinct nature of this outbreak based on transmission patterns not previously observed for Clade I MPXV. We also describe the complete genome of this virus and draw special attention to the phylogeny and unique mutational profile, which we propose are responsible for the sustained nature of this outbreak. In this study, we isolated viral DNA and reported six complete viral genomes, phylogeny, and mutational profiling associated with the Clade I Kamituga Mpox outbreak. Our sequences were compared to reference Clade I Mpox genomes (NC_003310.1) to determine the genomic variability of our samples.

The *C9L* gene in Mpox encodes a Kelch-like protein that serves an important role as an antagonist of the innate immune response. C9L is the only Kelch-like protein in the MPXV genome, while other orthopoxviruses contain multiple Kelch-like proteins [18]. The *C9L* gene is highly conserved across orthopoxviruses. In VACV, *C9L* is expressed early during infection and functions as an antagonist of the Type I interferon response where *C9L* encodes an ankyrin repeat/F-box protein shown to interfere with antiviral activity in human cells infected with VACV [19, 20]. The Mpox *C9L* gene contains an RNA G-quadraplex (RG4), which is susceptible to structural changes (instability) following genome evolution [18]. In Mpox, much like VACV, the C9L protein is thought to inhibit the innate immune response of the host, although further exploration is needed to delineate the exact role of the protein. It has been proposed that the Mpox genomes sequenced during the 2022 global outbreak contained the least stable RG4 configuration (i.e., C9L-RG4-5), and variability in the C9L may drive transmission of the virus [18, 21].

I4L is a gene that encodes a ribonucleotide reductase (RR) large subunit 1 protein, which is likely involved in the synthesis of nucleotides. This gene plays a role in virus replication in non-dividing cells [22]. RRs are highly conserved enzymes that catalyze the rate-limiting step during deoxynucleoside

Figure 3. Calculated phylogenetic trees of six Kamituga sequences with publicly available sequences for Clade I and II: **(A).** Phylogenetic tree of Clade I sub-group sequences with newly sequenced Kamituga genomes, proposed sub-group VI is labelled with green color; and **(B).** Rectangular representation of phylogenetic tree of Clade I, II, and Kamituga genomes (shown in green color).

triphosphate (dNTP) synthesis; in orthopoxviruses (including VACV and Mpox), RRs encode both R1 and R2 subunits [23]. Moreover, the inhibition of RR has been proposed as an anti-viral target [24].

The role of the *L6R* gene in Mpox is minimally described; however, this gene plays a likely role in the transcription of the viral genome. It is not unreasonable to assume that the *L6R* gene plays a role in transcribing and replicating the viral genome. These nucleic acid polymerases catalyze RNA synthesis from RNA templates, thus facilitating viral genome replication and transcription [25]. The *A25R* gene encodes the RPO132 subunit, which is predicted to be a component of the viral RNA polymerase complex fundamental to the virus's ability to replicate its DNA into RNA [26]. Although the specific contributions of *A25R* to the pathogenicity of MPXV are not yet fully understood, its high conservation across various Mpox strains underlines its critical role in the viral life cycle. The location of the *A25R* gene in the central core region of the virus genome which has over 90% sequence homology with other orthopoxviruses further highlights its importance and potential as a drug target [4]. Further studies are needed to elucidate the exact contribution of *A25R* to virulence and host interactions.

A17L is a viral envelope protein essential for host-cell entry and fusion. In the closely related VACV, A17L plays an important role in virion assembly and helps determine the intracellular sites of protein interactions, cleavage, phosphorylation, and disulfide bond formation [27]. A17L is phosphorylated by the F10 kinase and represents a major transmembrane component of VACV expressed late in infection [28]. Although this gene has not explicitly been described in Mpox, it likely augments viral entry and infection establishment, especially given its high level of conservation across samples in our study.

The role of the *C7L* gene in MPXV is analogous to the function of the *F1L* gene in VACV [29]. In VACV, F1L helps the virus in evade the host's innate immune defenses and acts as an inhibitor of apoptosis by binding to the pro-apoptotic Bcl-2 family of proteins, Bak and Bax. These proteins are key regulators of mitochondrial membrane integrity and cytochrome c release. *F1L* can also inhibit caspase-9 and NLRP1 inflammasome, thereby enhancing viral replication and survival within the host [30, 31]. While specific studies directly addressing the role of *C7L* in MPXV are not prevalent, it is reasonable to infer that *C7L* could play a similar role in MPXV, given the genetic and functional similarities among members of the orthopoxvirus family. The gene likely contributes to the virus's ability

to suppress immune responses, thereby enhancing the survival and transmission of the virus within its host [29, 32].

The A28L protein is involved in the host's cellular entry and exit processes [33]. It has shown high conservation rates, suggesting its crucial role in the virus's lifecycle and potential as a target for vaccine development. In a recent study, a significant number of epitopes from the A28L protein were identified and proposed for a multiepitope vaccine candidate, indicating its immunogenic potential. The in silico analysis of these epitopes suggests that they could elicit both humoral and cellular immune responses, thus representing a promising addition to the design of an immunogenic construct against MPXV [34].

The ongoing outbreak in Kamituga has been self-sustaining since late September 2023, where more than 200 men, women, and children have been infected with the virus. Interestingly, this Clade I outbreak demonstrates unique characteristics rarely observed during Clade I outbreaks, such as human-to-human transmission by heterosexual contact, and non-sexual contact (community spread). Sporadic Mpox outbreaks have been occurring in DRC since the 1970s, where the virus is endemic. Prior to the Kamituga Clade I outbreak presented in this study, there have been few documented cases of human-to-human spread of Clade I Mpox. One such outbreak occurred in the village of Bokungu (Eastern DRC) from July to December 2013 and involved 104 suspected MPXV infections and 10 deaths with the majority of cases occurring in males [35]. The main mode of transmission during this outbreak was human-to-human transmission, where community-level spread was favoured. A separate study (2023) recently documented a cluster of Clade I MPXV in the Kwango province of Western DRC, which was characterized by human-to-human transmission via heterosexual contact [10]. Although the timeline of this cluster overlaps with our current study, our phylogenetic analysis suggests that the causative virus for the Clade I Kamituga outbreak is genetically distinct from the previously described Clade I outbreaks.

The Clade I Kamituga outbreak is primarily being spread by heterosexual contact and also exhibits community-level transmission, where children have also been affected. Importantly, we underscore that classical routes of Clade I MPXV transmission should be reevaluated in endemic regions and adjusted based on these findings. To our knowledge, this paper is the first of its kind to annotate the novel Clade I Kamituga MPXV genome, providing biological context to key

genetic and protein level mutations. Our findings serve as a foundation for additional research efforts in the DRC to further elucidate the immunlogical underpinnings of infection, as the specific mechanisms governing lesion dissemination and transmission remain largely unknown. We acknowledge that only a fraction of our cohort was PCR confirmed for MPXV infection, which demonstrates the urgent need for increased supplies, resources, and infrastructure in Mpox endemic regions of East Africa.

Conclusions

We describe the phylogeny, mutational profiling, and genome annotation associated with the novel Clade I Mpox outbreak in Kamituga, Eastern DRC. Our phylogenetic analysis and genome annotation findings suggest that this cluster of infections is being driven by a novel lineage (here we term this subgroup VI) of Clade I Mpox with a unique, pathogen-favouring mutational profile. Noteworthy is the fact that this Clade I outbreak is being transmitted by sexual contact, a feature only previously described in Clade II Mpox. These findings underscore the importance of continued genomic surveillance of MPXV in endemic regions and suggest that emerging subgroups of MPXV are broadening the demographic populations at risk of contracting the disease.

Data availability statement

All six genome sequences were deposited to GISAID with the following accession IDs: EPI_ISL_19004044, EPI_ISL_19004045, EPI_ISL_19004046, EPI_ISL_19079342, EPI_ISL_19079343, and EPI_ISL_19079344.

References

1. Otu A, Ebenso B, Walley J, Barceló JM, Ochu CL (2022) Global human monkeypox outbreak: atypical presentation demanding urgent public health action. Lancet Microbe 3: e554-e555. doi: 10.1016/S2666-5247(22)00153-7.
2. McCollum AM, Damon IK (2014) Human monkeypox. Clin Infect Dis Off Publ Infect Dis Soc Am 58: 260-267. doi: 10.1093/cid/cit703.
3. Gessain Antoine, Nakoune Emmanuel, Yazdanpanah Yazdan (2022) Monkeypox. N Engl J Med 387: 1783-1793. doi: 10.1056/NEJMra2208860.
4. Lu J, Xing H, Wang C, Tang M, Wu C, Ye F, Yin L, Yang Y, Tan W, Shen L (2023) Mpox (formerly monkeypox): pathogenesis, prevention, and treatment. Signal Transduct Target Ther 8: 458. doi: 10.1038/s41392-023-01675-2.
5. World Health Organization (WHO) (2023) Mpox (monkeypox). Available: https://www.who.int/news-room/fact-sheets/detail/monkeypox. Accessed: 24 January 2024.
6. Durski KN, McCollum AM, Nakazawa Y, Petersen BW, Reynolds MG, Briand S, Djingarey MH, Olson V, Damon IK, Khalakdina A (2018) Emergence of Monkeypox - West and Central Africa, 1970-2017. MMWR Morb Mortal Wkly Rep 67: 306-310. doi: 10.15585/mmwr.mm6710a5.
7. Moore MJ, Rathish B, Zahra F (2024) Mpox (Monkeypox). StatPearls. StatPearls Publishing, Treasure Island (FL).
8. Giorgi FM, Pozzobon D, Di Meglio A, Mercatelli D (2024) Genomic and transcriptomic analysis of the recent Mpox outbreak. Vaccine 42: 1841-1849. doi: 10.1016/j.vaccine.2023.12.086.
9. Americo JL, Earl PL, Moss B (2023) Virulence differences of mpox (monkeypox) virus clades I, IIa, and IIb.1 in a small animal model. Proc Natl Acad Sci 120: e2220415120. doi: 10.1073/pnas.2220415120.
10. Kibungu EM, Vakaniaki EH, Kinganda-Lusamaki E, Kalonji-Mukendi T, Pukuta E, Hoff NA, Bogoch II, Cevik M, Gonsalves GS, Hensley LE, Low N, Shaw SY, Schillberg E, Hunter M, Lunyanga L, Linsuke S, Madinga J, Peeters M, Cigolo J-CM, Ahuka-Mundeke S, Muyembe J-J, Rimoin AW, Kindrachuk J, Mbala-Kingebeni P, Lushima RS (2024) Clade I-Associated Mpox Cases Associated with Sexual Contact, the Democratic Republic of the Congo. Emerg Infect Dis 30: 172-176. doi: 10.3201/eid3001.231164.
11. Shantier SW, Mustafa MI, Abdelmoneim AH, Fadl HA, Elbager SG, Makhawi AM (2022) Novel multi epitope-based vaccine against monkeypox virus: vaccinomic approach. Sci Rep 12: 15983. doi: 10.1038/s41598-022-20397-z.
12. Chen S, Zhou Y, Chen Y, Gu J (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34: i884-i890. doi: 10.1093/bioinformatics/bty560.
13. Katoh K, Misawa K, Kuma K, Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 30: 3059-3066. doi: 10.1093/nar/gkf436.
14. National Library of Medicine, National Center for Biotechnology Information (NCBI) (2024) Welcome to NCBI. Available: https://www.ncbi.nlm.nih.gov/. Accessed: 24 April 2024.
15. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 30: 2725-2729. doi: 10.1093/molbev/mst197.

16. Ma Y, Chen M, Bao Y, Song S (2022) MPoxVR: A comprehensive genomic resource for monkeypox virus variant surveillance. The Innovation 3: 100296. doi: 10.1016/j.xinn.2022.100296.

17. EMBL's European Bioinformatics Institute (EMBL-EBI) (2024) EMBL-EBI. Unleashing the potential of big data in biology. Available: https://www.ebi.ac.uk/. Accessed: 24 April 2024.

18. Dai Y, Teng X, Hu D, Zhang Q, Li J (2023) Evolution of Monkeypox C9L RNA G-Quadruplex Elevates the Translation of Its Only Kelch-like Protein by Accelerating G-Quadruplex Unfolding Kinetics. CCS Chem 6: 953-963. doi: 10.31635/ccschem.023.202302807.

19. Liu R, Moss B (2018) Vaccinia Virus C9 Ankyrin Repeat/F-Box Protein Is a Newly Identified Antagonist of the Type I Interferon-Induced Antiviral State. J Virol 92: e00053-18. doi: 10.1128/JVI.00053-18.

20. Ingham RJ, Loubich Facundo F, Dong J (2022) Poxviral ANKR/F-box Proteins: Substrate Adapters for Ubiquitylation and More. Pathogens 11: 875. doi: 10.3390/pathogens11080875.

21. Zebardast A, Latifi T, Shafiei-Jandaghi N-Z, Gholami Barzoki M, Shatizadeh Malekshahi S (2023) Plausible reasons for the resurgence of Mpox (formerly Monkeypox): an overview. Trop Dis Travel Med Vaccines 9: 23. doi: 10.1186/s40794-023-00209-6.

22. Chaurasiya S, Chen NG, Lu J, Martin N, Shen Y, Kim S-I, Warner SG, Woo Y, Fong Y (2020) A chimeric poxvirus with J2R (thymidine kinase) deletion shows safety and anti-tumor activity in lung cancer models. Cancer Gene Ther 27: 125-135. doi: 10.1038/s41417-019-0114-x.

23. Gammon DB, Gowrishankar B, Duraffour S, Andrei G, Upton C, Evans DH (2010) Vaccinia Virus-Encoded Ribonucleotide Reductase Subunits Are Differentially Required for Replication and Pathogenesis. PLoS Pathog 6: e1000984. doi: 10.1371/journal.ppat.1000984.

24. Reardon JE, Spector T (1991) Acyclovir: Mechanism of Antiviral Action and Potentiation by Ribonucleotide Reductase Inhibitors. In: August JT, Anders MW, Murad F (eds) Advances in Pharmacology. Academic Press, pp 1-27. doi: 10.1016/S1054-3589(08)60031-9.

25. Jia H, Gong P (2019) A Structure-Function Diversity Survey of the RNA-Dependent RNA Polymerases From the Positive-Strand RNA Viruses. Front Microbiol 10: 1945. doi: 10.3389/fmicb.2019.01945.

26. Chen N, Li G, Liszewski MK, Atkinson JP, Jahrling PB, Feng Z, Schriewer J, Buck C, Wang C, Lefkowitz EJ, Esposito JJ, Harms T, Damon IK, Roper RL, Upton C, Buller RML (2005) Virulence differences between monkeypox virus isolates from West Africa and the Congo basin. Virology 340: 46-63. doi: 10.1016/j.virol.2005.05.030.

27. Betakova T, Moss B (2000) Disulfide Bonds and Membrane Topology of the Vaccinia Virus A17L Envelope Protein. J Virol 74: 2438-2442. doi: 10.1128/JVI.74.5.2438-2442.2000.

28. Moss B (2015) Poxvirus membrane biogenesis. Virology 479-480: 619-626. doi: 10.1016/j.virol.2015.02.003.

29. Lum F-M, Torres-Ruesta A, Tay MZ, Lin RTP, Lye DC, Rénia L, Ng LFP (2022) Monkeypox: disease epidemiology, host immunity and clinical interventions. Nat Rev Immunol 22: 597-613. doi: 10.1038/s41577-022-00775-4.

30. Pelin A, Foloppe J, Petryk J, Singaravelu R, Hussein M, Gossart F, Jennings VA, Stubbert LJ, Foster M, Storbeck C, Postigo A, Scut E, Laight B, Way M, Erbs P, Le Boeuf F, Bell JC (2019) Deletion of Apoptosis Inhibitor F1L in Vaccinia Virus Increases Safety and Oncolysis for Cancer Therapy. Mol Ther Oncolytics 14: 246-252. doi: 10.1016/j.omto.2019.06.004.

31. Gerlic M, Faustin B, Postigo A, Yu EC-W, Proell M, Gombosuren N, Krajewska M, Flynn R, Croft M, Way M, Satterthwait A, Liddington RC, Salek-Ardakani S, Matsuzawa S, Reed JC (2013) Vaccinia virus F1L protein promotes virulence by inhibiting inflammasome activation. Proc Natl Acad Sci U S A 110: 7808-7813. doi: 10.1073/pnas.1215995110.

32. Alakunle E, Kolawole D, Diaz-Cánova D, Alele F, Adegboye O, Moens U, Okeke MI (2024) A comprehensive review of monkeypox virus and mpox characteristics. Front Cell Infect Microbiol 14: 1360586. doi: 10.3389/fcimb.2024.1360586.

33. Hudu SA, Alshrari AS, Al Qtaitat A, Imran M (2023) VP37 protein inhibitors for Mpox treatment: highlights on recent advances, patent literature, and future directions. Biomedicines 11: 1106. doi: 10.3390/biomedicines11041106.

34. de Araújo LP, de Melo Santos NC, Corsetti PP, de Almeida LA (2024) Immunoinformatic approach for rational identification of immunogenic peptides against host entry and/or exit Mpox proteins and potential multiepitope vaccine construction. J Infect Dis 229: S285-S292. doi: 10.1093/infdis/jiad443.

35. Nolen LD, Osadebe L, Katomba J, Likofata J, Mukadi D, Monroe B, Doty J, Hughes CM, Kabamba J, Malekani J, Bomponda PL, Lokota JI, Balilo MP, Likafi T, Lushima RS, Ilunga BK, Nkawa F, Pukuta E, Karhemere S, Tamfum J-JM, Nguete B, Wemakoy EO, McCollum AM, Reynolds MG (2016) Extended human-to-human transmission during a Monkeypox outbreak in the Democratic Republic of the Congo. Emerg Infect Dis 22: 1014-1021. doi: 10.3201/eid2206.150579.

Corresponding author

David J. Kelvin, PhD
Canada Research Chair in Translational Vaccinology and Inflammation, The Department of Microbiology and Immunology, Dalhousie University, Halifax, Nova Scotia, B3H 4H7, Canada
Tel: +1 647-529-3556
Email: David.Kelvin@dal.ca

Conflict of interests: The authors DJK, AK, MD, PK, and GSM are shareholders of the BioForge Canada Limited company. The company specializes in developing computational solutions for biological problems. The authors declare that the research reported in this work was conducted independently, and the reported results were not influenced by any financial or personal relationships with the company.

Annex – Supplementary Items

Supplementary Table 1. List of the mutations calculated in the Kamituga MPXV sequences based on the mapping to reference genome (NC_003310).

GENE	ORTHOPOXVIRUS GENE	PROTEIN FUNCTION	SAMPLE_1	SAMPLE_2	SAMPLE_3	SAMPLE_4	SAMPLE_5	SAMPLE_6	CONSENSUS MUTATIONS
J1L	OPG001	CC-chemokine binding Chemokine binding protein (Cop-C23L) J1L most abundant secreted protein	V194	V194	V194	V194	V194	V194	V194
J2L	OPG002	TNF receptor (CrmB) (Cop-C22L)	L22I	L22I	L22I	L22I	L22I	L22I	L22I
J3L	OPG003	Ankyrin (Cop-C19L)	T546; T487I	T546; T487I; T487I	T546; T487I; K103X	T546;T487I	T546; T487I	T546; T487I	T546;T487I
D1L	OPG015	Ankyrin (CPXV-017)	T273A; A249T; I48T	T273A; A249T; I48T	T273A; A249T; I48T	T273A; A249T; I48T	T273A; A249T; I48T	T273A; A249T; I48T	T273A; A249T; I48T
#D2L			C33Y	C33Y	C33Y	C33Y	C33Y	C33Y	C33Y
D3R	OPG019	Secreted EGF-like protein (Cop-C11R) secreted growth factor	G112C	G112C	G112C	G112C	G112C	G112C	G112C
D7L	OPG023	Ankyrin/Host Range	G384X; S107	G384X; S107	G384X; S107	S107	G384X; S107	S107	S107
D9L	OPG025	Ankyrin D9L Type I IFN resistance (Cop-C9L), ankyrin-like protein	A276; K47X; I20S	A276; K47X; I20S	A276; K47X; I20S	A276; I20S	A276; K47X; I20S	A276; I20S	A276; I20S
D10L	OPG027	Type 1 IFN inhibitor (Cop-C7L), host range	L79F	L79F	L79F	L79F	L79F	L79F	L79F
D13L	OPG031	IL-1 receptor antagonist (Cop-C10L)	NA	NA	NA	P297X	P297X	P297X	NA
#D17L		D17L	NA	E95X; D56GN	E95X; D56GN	E95X; M94IX; D56GN	E95X; M94IX; D56GN	E95X; D56GN	NA
D19L	OPG034	Putative TLR signalling inhibitor (Cop-C1L)	P184H; T87	P184H; T87	P184H; T87	P184H; T87	P184H; T87	P184H; T87	P184H; T87
P2L	OPG036	Alpha amanatin target protein (Cop-N2L)	V17I	V17I	V17I	V17I	V17I	V17I	V17I
O1L	OPG037	ANK-containing protein O1L ankyrin-like apoptosis inihibitor (Cop-M1L)	S395F; I355M	I355M	I355M	NA	NA	I355M	NA
C1L	OPG039	Ankyrin/NFkB inhibitor (Cop-K1L) C1L ankyrin-like host range	V173	V173	V173	V173	V173	V173	V173
C2L	OPG040	SPI-3 Serpin 1,2,3 (Cop-K2L) inhibition of the ability of infected cells to fuse serine protease inhibitor-like	I362; T178; M160I	I362; T178; M160I	I362; T178; M160I	I362; T178; M160I	I362; T178; M160I	I362; T178; M160I	I362; T178; M160I
C4L	OPG042	Phospholipase-D-like protein (Cop-K4L)	R228G; T159I	R228G; T159I	R228G; T159I	R228G; T159I	R228G; T159I	R228G; T159I	R228G; T159I
C5L	OPG043	Monoglyceride lipase (Cop-K5L) lysophospholipase-like	D210N	D210N; N68S	D210N; N68S	D210N; N68S	D210N; N68S	D210N; N68S	D210N; N68S
C6R	OPG044	Host immune response repressor (Cop-K7R) VAC B15R-like	K147X	K147X	K147X	NA	NA	NA	NA
C7L	OPG045	Caspase-9 (apoptosis) inhibitor (mitochondrial- associated) (Cop-F1L)	G78; G78R; YVH13-15Y	G78; G78R; YVH13-15Y	G78; G78R; YVH13-15Y	G78; G78R; YVH13-15Y	G78; G78R; YVH13-15Y	G78;G78R; YVH13-15Y	G78; G78R; YVH13-15Y (inframe deletion)
C9L	OPG047	Kelch-like protein (Cop-F3L)	DGMG483-486E; D483X; L151H; V127E; Y114N	DGMG483-486E; D483X; L151H; V127E; Y114N	DGMG483-486E; D483X; L151H; V127E; Y114N	DGMG483-486E; D483X; L151H; V127E; Y114N	MG485-486X; DG483-484X ; D483X; L151H; V127E; Y114N	DGMG483-486E; D483X; L151H; V127E; Y114N	DGMG483-486E (inframe deletion); D483X; L151H; V127E; Y114N
C10L	OPG048	Ribonucleotide reductase small subunit (Cop-F4L)	P215; D47	P215; D47	P215; D47	P215; D47	P215; D47	P215; D47	P215; D47
C11L	OPG049	36kDa major membrane protein (Cop-F5L)	T303M; A301T; R24H	T303M; A301T; R24H	T303M; A301T; R24H	T303M; A301T; R24H	T303M; A301T; R24H	T303M; A301T; R24H	T303M; A301T; R24H
C16L	OPG054	Essential Ser/Thr kinase morph (Cop-F10L)	I399; G351; V256A; P215; E142X; S74	I399; G351; V256A; P215;E142X; S74	I399; G351; V256A; P215; E142X; S74	I399; G351; V256A; P215; S74	I399; G351; V256A; P215; S74	I399; G351; V256A; P215	I399; G351; V256A; P215
C17L	OPG055	RhoA signalling inhibitor, virus release protein (Cop-F11L)		G154X	NA	NA	NA	Y54FN	NA
C18L	OPG056	EEV maturation protein (Cop-F12L) actin tailformation	S360T; Y203C; I126N	S360T; Y203C; I126N	S360T; Y203C; I126N	S360T; Y203C	S360T; Y203C	S360T; Y203C	S360T; Y203C
C19L	OPG057	Palmitylated EEV membrane glycoprotein (Cop-F13L)C19L phospholipase D-like, major envelope antigen of EEV wrapping of IMV to form IEV	S250N; K110X; V5A	S250N; K110X; V5A	S250N; K110X; V5A	S250N; V5A	S250N; V5A	S250N; V5A	S250N; V5A
C22L	OPG061	Non-functional Serine Recombinase (Cop-F16L)	A157; E47X	A157; E47X	A157; E47X	A157	A157	A157	A157
C23R	OPG062	DNA-binding phosphoprotein (VP11) mTOR antagonist (Cop-F17R)	A8; N94	A8; N94	A8; N94	A8; N94; P63X	A8; N94	A8; P63X; N94	A8; N94
F1L	OPG063	Poly (A) polymerase catalytic subunit (VP55) (Cop-E1L) poly-A polymerase	F63X; R62X	F63X; R62X	F63X; R62X			F65X	
F2L	OPG064	IEV morphogenesis (Cop-E2L)	H89	H89	H89	H89	H89	H89	H89
F4L	OPG066	RNA polymerase 30 kDa subunit RNA polymerase subunit (RPO30) (Cop-E4L) simultaneously intermediate stage promoter-specific transcription factor, VITF-1	L222I; K140N	L222I; K140N	L222I; K140N	L222I; K140N	L222I; K140N	L222I; K140N	L222I; K140N
F6R	OPG069	Myristylated protein (Cop-E7R)	L103	L103	L103	L103	L103	L103	L103
F7R	OPG070	ER-localized membrane protein, virion core protein (Cop-E8R)	Y12H	Y12H	Y12H	Y12H	Y12H	Y12H	Y12H
F8L	OPG071	DNA polymerase (Cop-E9L) DNA polymerase, catalytic subunit	D785N; Y668	D785N; Y668	D785N; Y668	D785N; Y668	D785N; Y668	D785N; Y668	D785N; Y668
F9R	OPG072	Sulfhydryl oxidase (FAD-linked) (Cop-E10R) protein disulfide bond-forming enzyme	I34L	I34L	I34L	I34L	I34L	I34L	I34L
F10L	OPG073	Virion core protein (Cop-E11L)	I6X		I6X			D23N	
O1L	OPG074	Membrane protein (Cop-O1L)	K362X; T243	K362X; T243	K362X; T243	T243	T243	T243; S99X	T243
Q2L	OPG075	Glutaredoxin 1 (Cop-O2L)	F74X	F74X	F74X	NA	NA	NA	NA
I1L	OPG077	DNA-binding core protein (Cop-I1L), virosomal protein essential for virus multiplication	V302	V302	V302	V302	V302	V302	V302
I2L	OPG078	IMV membrane protein (Cop-I2L)	I54X	I54X	I54X	NA	NA	NA	NA

GENE	ORTHOPOXVIRUS GENE	PROTEIN FUNCTION	SAMPLE_1	SAMPLE_2	SAMPLE_3	SAMPLE_4	SAMPLE_5	SAMPLE_6	CONSENSUS MUTATIONS
I4L	OPG080	Ribonucleotide reductase large subunit (Cop-I4L), R1	A685P; T372; R358G; G260S; T31	A685P; T372; R358G; G260S; T31	A685P; T372; R358G; G260S; T31	A685P; T372; R358G; G260S; T31	A685P; T372; R358G; G260S; T31	A685P; T372; R358G; G260S; T31	A685P; T372; R358G; G260S; T31
I5L	OPG081	IMV protein VP13 (Cop-I5L) IMV surface membrane protein	T39	T39	T39	T39	T39	T39	T39
I6L	OPG082	Telomere-binding protein (Cop-I6L)	K360X; I276; K171T	I276; K171T	I276; K171T	I276; K171T	I276; K171T	I276; K171T	I276; K171T
I8R	OPG084	DNA and RNA helicase I8R RNA helicase, DExH-NPH-II domain (Cop-I8R) nucleoside triphosphate phosphohydrolase II	F11X; K72X; G140; T383	F11X; K72X; G140; 383	F11X; K72X; G140; T383	F11X; G140; T383	F11X; G140; T383	G140; T383	G140; T383
G1L	OPG085	Metalloprotease (Cop-G1L)	K584X; L467; P395L; A393	K584X; L467; P395L; A393	K584X; L467; P395L; A393	L467; P395L; A393	L467; P395L	L467; P395L	L467; P395L
G3R	OPG087	VLTF (late transcription elongation factor) (Cop-G2R)				K178X	K178X	K178X	
G4L	OPG088	Glutaredoxin-like protein (Cop-G4L)	F40X	F40X	F40X	NA	NA	NA	NA
G5R	OPG089	FEN1-like nuclease (Cop-G5R)	I213K	I213K	I213K	I213K	I213K	I213K	I213K
G7R	OPG091	NLPc/P60 superfamily protein (Cop-G6R)	S18F	S18F	S18F	S18F	S18F	S18F	S18F
G8L	OPG092	Virion phosphoprotein, early morphogenesis (Cop-G7L)	Q279	Q279	Q279	Q279	Q279	Q279	Q279
G9R	OPG093	VLTF-1 (late transcription factor 1) (Cop-G8R) late gene transcription factor	NA	NA	NA	NA	NA	K221X	
G10R	OPG094	Entry/fusion complex component, myristylprotein (Cop-G9R)	G111	G111	G111	G111	G111	G111	G111
M1R	OPG095	IMV membrane protein (Cop-L1R) M1R myristylated IMV surface membrane protein	G183	G183	G183	G183	G183	G183	G183
M2R	OPG096	Viral membrane assembly proteins (VMAP) (Cop-L2R)	L75P	L75P	L75P	L75P	L75P	L75P	L75P
M3L	OPG097	Internal virion protein (Cop-L3L)	T298; V285; K264X; N198D	T298; V285; K264X; N198D	T298; V285; K264X; N198D	T298; V285; K264X; N198D	T298; V285; K264X; N198D	T298; V285; K264X; N198D	T298; V285; K264X; N198D
M4R	OPG098	ssDNA binding stimulation of I8R helicase activity virion core protein	P248	P248	P248	P248	P248	P248	P248
L1R	OPG100	Virion morph (Cop-J1R)	R134H	R134H	R134H	R134H	R134H	R134H	R134H
L2R	OPG101	Thymidine kinase (Cop-J2R)	L14F; L109	L14F; L109	L14F; L109	L14F; L109	L14F; L109	L14F; L109	L14F; L109
L3R	OPG102	Poly (A) polymerase small subunit (VP39) (Cop-J 3R) poly-A polymerase, stimulatory subunit, simultaneously cap-specific mRNA (nucleoside-O2'-)-methyltransferase	K33X; K235X	K33X; K235X	K33X; K235X	NA	NA	NA	NA
L5L	OPG104	IMV membrane protein (Cop-J5L) L5L essential for virus multiplication	K58X	K58X	K58X	NA	NA	NA	NA
L6R	OPG105	RNA polymerase subunit (RPO147) (Cop-J6R) RNA polymerase, 147 kDa	_281-282X; P285X; A396; V431A; R475C; D506Y	_281-282X; P285X; A396; V431A; R475C; D506Y; F944S	_281-282X; P285X; A396; V431A; R475C; D506Y	P285X; A396;V431A; R475C; D506Y;R475C; F944S	P285X; A396; V431A; R475C; D506Y; F944S	P285X; A396; V431A; P453X; R475C; D506Y; F944S	P285X (frameshift variant); A396;V431A; R475C; D506Y
H3L	OPG108	IMV heparin binding surface protein (Cop-H3L) IMV surface membrane protein	H239	H239	H239	H239; E45X	H239; E45X	H239; E45X	H239
H4L	OPG109	virion core RNA polymerase-associated protein(Cop-H4L), RAP94	V727; E451X; V377A; K204X; L114X; N46	V727; H239Z; V377A; K204X; L114X; N46	V727; H239Z; V377A; K204X; L114X; N46	V727; V377A; L114X; N46	V727; V377A; L114X; N46	V727; V377A; LF114-115LX; N46	V727; V377A; L114X (frameshift variant); N46
E2L	OPG114	Virion core protein (Cop-D2L)	K8X	K8X	K8X	NA	NA	NA	NA
E3R	OPG115	Virion core protein (Cop-D3R)	F45X	NA	F45X	NA	K228X	K228X	NA
E5R	OPG117	NTPase, DNA primase (Cop-D5R) nucleic acid-independent nucleoside triphosphatase	NA	F502X; K678X;	F502X	NA	NA	NA	NA
E6R	OPG118	Morphogenesis, early transcription factor, VETF, small subunit	A137	A137	A137	A137	A137	A137	A137
E10R	OPG122	Mut-like down regulation of gene expression mRNA decapping enzyme (Cop-D10R)	F160X	F160X	F160X	NA	NA	NA	NA
E11L	OPG123	ATPase, NPH1 (Cop-D11L) DNA-dependent ATPase E11L early gene transcription termination factor nucleoside triphosphate phosphohydrolase I, NPH I	K568X	K568X	K568X	NA	NA	NA	NA
E12L	OPG124	mRNA (guanine-N7-)-methyl-transferase mRNA capping enzyme, small subunit transcription initiation factor	K203X	K203X	K203X	NA	NA	NA	NA
E13L	OPG125	Trimeric virion coat protein (rifampicin res) (Cop-D13L) protein needed for the formation of immature IMV surface membrane	T493M	T493M	T493M	T493M	T493M	T493M	T493M
A1L	OPG126	VLTF-2 (late transcription factor 2) (Cop-A1L)	A26S	A26S	A26S	A26S	A26S	A26S	A26S
A2L	OPG127	VLTF-3 (late transcription factor 3) (Cop-A2L)	R175K; Q37X; C6	R175K; Q37X; C6	R175K; Q37X; C6	R175K; C6	R175K; C6	R175K; C6	R175K; C6
A3L	OPG128	S-S bond formation pathway protein (Cop-A2.5L)	S30	S30	S30	NA	NA	NA	NA
A4L	OPG129	P4b precursor (Cop-A3L) major virion core protein	P490; V9X	P490	P490; V9X	P490; G61X	P490; G61X	P490; G61X	P490
A8L	OPG133	early transcription factor, VETF, large subunit needed for morphogenesis of the virion core (Cop- A7L)	K583X	K583X	K583X	NA	NA	NA	NA
A9R	OPG134	VITF-3 34kda subunit (Cop-A8R) intermediate transcription factor	E15; G17X; N89X	E15; G17X; N89X	E15; G17X; N89X	E15	E15	SS	E15
A11L	OPG136	P4a precursor (Cop-A10L) major virion core protein	K875X; M740I; N313; L303	K875X; M740I; N313; L303	K875X; M740I; N313; L303	M740I; N313; L303	M740I; N313; L303	M740I; N313; L303	M740I; N313; L303

GENE	ORTHOPOXVIRUS GENE	PROTEIN FUNCTION	SAMPLE_1	SAMPLE_2	SAMPLE_3	SAMPLE_4	SAMPLE_5	SAMPLE_6	CONSENSUS MUTATIONS
A13L	OPG138	Virion core and cleavage processing protein (Cop-A12L)	K4X	K4X	K4X	NA	NA	NA	NA
A15L	OPG140	Essential IMV membrane protein (Cop-A14L) IMV inner membrane protein	P39H; L15	P39H; L15	P39H; L15	P39H; L15	P39H; L15	P39H; L15	P39H; L15
A17L	OPG143	Myristylated protein, essential for entry/fusion (Cop-A16L)	V322A; V304A; A289T; R127; E110	V322A; V304A; A289T; R127; E110	V322A; V304A; A289T; R127; E110	V322A; V304A; A289T; R127; E110	V322A; V304A; A289T; R127; E110	V322A; V304A; A289T; R127; E110	V322A; V304A; A289T; R127; E110
A19R	OPG145	DNA helicase, transcript release factor (Cop-A18R) post-replicative negative transcription elongation factor	K16X; P76Q; K260X; V473	K16X; P76Q; K260X; V473	K16X; P76Q; K260X; V473	P76Q; V473	P76Q; K260X; V473	P76Q; K260X; V473	P76Q; V473
A20L	OPG146	Zinc finger-like protein (Cop-A19L)	G25D	G25D	G25D	G25D	G25D	G25D	G25D
A21L	OPG147	IMV membrane protein, entry/fusion complex component (Cop-A21L)	G52S	G52S	G52S	G52S	G52S	G52S	G52S
A22R	OPG148	DNA polymerase processivity factor (Cop-A20R)	F69; Y82; N326; K422X	F69; Y82; N326; K422X	F69; Y82; N326; K422X	F69; Y82; N326	F69; Y82; K244X; N326	F69; Y82; K244X; N326	F69; Y82; N326
A23R	OPG149	Holliday junction resolvase (Cop-A22R)	K186X	K186X	K186X	K125X	K125X	N35X; K125X	
A25R	OPG151	RNA polymerase, 132 kDa subunit (RPO132) (Cop-A24R)	N4X; S220; R273Q; L360; T601M; D633E; P903; T910; K1034X	N4X; S220; R273Q; L360; T601M; D633E; P903; T910; K1034X	N4X; S220; R273Q; L360; T601M; D633E; P903; T910; K1034X	S220; R273Q; D324X; L360; T601M; D633E; P903; T910	S220; R273Q; L360; T601M; D633E; P903; T910	S220; R273Q; L360; T601M; D633E; P903; T910	S220; R273Q; L360; T601M; D633E; P903; T910
#A27L		A-type inclusion protein (Cop-A25L)	S589; T494I; V394A	S589; T494I; V394A	S589; T494I; V394A	S589; T494I; V394A	S589; T494I; V394A	S589; T494I; V394A	S589; T494I; V394A
A28L	OPG153	P4c precursor (Cop-A26L) major component of IMV surface tubules p4c	DD386-387X; D386X; D385H; D384X; DDDDDII367-373-; S344N	DD386-387X; D386X; D385H; D384X; DDDDDII367-373-; S344N	DD386-387X; D386X; D385H; D384X; DDDDDII367-373-; S344N	DDDDDDD381-387X; D381X; D380H; D379X; II372-373X; S344N	DDDDDDD381-387X; D387X; D386H; D385X; DDDDDII367-373-; S344N	DDDDDDD381-387X; D381X; D380H; D379X; II372-373X; S344N	DD386-387X D386X; D384X D385H DDDDDII367-273-S344N
A29L	OPG154	IMV surface membrane 14 kDa fusion protein IMV surface protein, fusion protein (Cop-A27L) binding to cell surface heparan	E55X	E55X	E55X	NA	NA	NA	NA
A30L	OPG155	IMV MP/Virus entry (Cop-A28L)	NA	I6X	I6X	NA	NA	NA	NA
A31L	OPG156	RNA polymerase, 35 kDa subunit (RPO35) (Cop-A29L)	M261I; M154I	M261I; M154I	M261I; M154I	M261I; M154I	M261I; M154I	M261I; M154I	M261I; M154I
A32L	OPG157	IMV protein (Cop-A30L)	H61Q	H61Q	H61Q	NA	NA	NA	NA
A34L	OPG160	ATPase/DNA packaging protein (Cop-A32L) DNA packaging into virion NTP-binding motif A	R61K	R61K	R61K	R61K	R61K	R61K	R61K
A35R	OPG161	EEV envelope glycoprotein, needed for formation of actin-containing microvilli and cell-to-cell spread of virion EEV membrane phosphoglycoprotein, C-type lectin-like domain (Cop-A33R) interacts with VAC	Q59L	Q59L	Q59L	NA	NA	NA	NA
A36R	OPG162	C-type lectin-like IEV/EEV glycoprotein (Cop-A34R) EEV envelope glycoprotein, lectin-like required for infectivity of EEV, formation of actin-containing microvilli, and cell-to-cell spread of virion	G84R; E139	G84R; E139	G84R; E139	G84R; E139	G84R; E139	G84R; E139	G84R; E139
A37R	OPG163	MHC class II antigen presentation inhibitor (Cop-A35R)	N89X	E139	E139	NA	NA	NA	NA
A38R	OPG164	IEV but not CEV envelope protein IEV transmembrane phosphoprotein (Cop-A36R) interacts with VAC A33R and A34R plays critical role for actin tail formation	Y83N; K117E	Y83N; K117E	Y83N; K117E	Y83N	Y83N; K117E	Y83N	Y83N
A39R	OPG165	Hypothetical protein (Cop-A37R)	G63D; D97; T127I; K238X	G63D; D97; T127I; K238X	G63D; D97; T127I; K238X	G63D; D97; T127I	G63D; D97; T127I	G63D; D97; T127I; E218X	G63D; D97; T127I
A41L	OPG170	Chemokine binding protein (Cop-A41L) secreted protein reducing infiltration of inflammatory cells into the infected area	Y116H; V96I	Y116H; V96I	Y116H; V96I	Y116H; V96I	Y116H; V96I	Y116H; V96I	Y116H; V96I
A43R	OPG172	Type I membrane glycoprotein (Cop-A43R)	M4I****W	M4I*MMMMX; K191X	M4I****W; K191X;	M4I****W	M4I****W	M4I****W	M4I****W
A44R	OPG173	Hypothetical protein (Cop-A43.5R)	N21S; N59X	N21S	N21S	N21S	N21S	N21S	N21S
A45L	OPG174	3 beta-hydroxysteroid dehydrogenase/delta 5->4 isomerase (Cop-A44L) 3-b-Hydroxy-delta5-steroid dehydrogenase	R315C	R315C	R315C	R315C	R315C	R315C	R315C
A47R	OPG176	IL-1/TLR signaling inhibitor (Cop-A46R)	T115	T115	T115	T115	T115	T115	T115
A49R	OPG178	Thymidylate kinase (Cop-A48R)	NA	K57X	K57X	NA	NA	NA	NA
#A48R		Thymidylate kinase (Cop-A48R)	D2VKX; I7RX			NA	NA	Y54FN	NA
A50R	OPG180	ATP-dependent DNA ligase (Cop-A50R	N245X; K479X	N245X; K479X	N245X; K479X				
A51R	OPG181	Hypothetical protein (Cop-A51R)	G45; D90; 107A; F151X	G45; D90; 107A; F151X	G45; D90; 107A; F151X	G45; D90; T107A	G45; D90; T107A	G45; D90; T107A	G45; D90; T107A
#B1R		kelch-like	N66X	N66X	N66X	N66X	N66X	N66X	N66X (frameshift varint)
B2R	OPG185	EEV envelope and cell membrane glycoprotein hemagglutinin Hemagglutinin (Cop-A56R) inhibition of the ability of infected cells to fuse interacts with VAC	V16A; T239I	V16A; T239I	V16A; T239I	V16A	V16A; T239I; T239I	V16A; T239I	V16A
B4R	OPG188	Schlafen (Cop-B2R) schlafen-like	R19G	R19G	R19G	R19G; K437X	R19G; K437X	R19G; K437X	R19G
B5R	OPG189	Ankyrin (Cop-B4R) B5R ankyrin-like	I87X	I87X	I87X; K209X	G101X	G101X	G101X	NA
B6R	OPG190	EEV type-1 membrane glycoprotein, protective antigen (Cop-B5R) complement control protein-like palmitated 42 kDa glycoprotein located both on the membranes of infected cells and on EEV envelope	C61; N114X	C61; N114X; P251	C61; N114X; P251	C61; P251	C61; P251	C61	C61

GENE	ORTHOPOXV IRUS GENE	PROTEIN FUNCTION	SAMPLE_1	SAMPLE_2	SAMPLE_3	SAMPLE_4	SAMPLE_5	SAMPLE_6	CONSENSUS MUTATIONS
B7R	OPG191	Ankyrin-like protein (Cop-B6R)	T120A	T120A	T120A	T120A	T120A	T120A	T120A
B8R	OPG192	Virulence, ER resident (Cop-B7R)	Y112H	Y112H	Y112H	Y112H	Y112H	Y112H	Y112H
B11R	OPG198	Ser/Thr Kinase (Cop-B12R)	E14K; K231X; F241X	E14K; K231X; F241X	E14K; K231X; F241X	E14K	E14K	E14K	E14K
B12R	OPG199	SPI-2 Serpin 1,2,3 (Cop-K2L) apoptosis inhibition inhibition of the IL-1b converting enzyme serine protease inhibitor-like, SPI-2	A74; S192P	A74; S192P	A74; S192P	A74; S192P	A74; S192P	A74; S192P	A74; S192P
B13R	OPG200	Hypothetical protein (Cop-C16L)	G52D	G52D	G52D; K96X	G52D	G52D	G52D	G52D
B14R	OPG201	IL-1 beta receptor (Cop-B16R) inhibition of virus infection induced fever secreted IL-1b binding protein	E197DIYI*	_200-201YIYIX	E197DIYI*	E197*	E197*	NA	NA
B17R	OPG205	Ankyrin (Cop-B20R) B17R ankyrin-like	SQSQSQS736-742S	SQSQSQS736-742S	SQSQSQS736-742S	SQSQSQS736-742S	SQSQSQS736-742S	K368X; QSQSQ733-737X	SQSQSQS736-742S (inframe deletion)
B20R	OPG209	Hypothetical protein (Cop-C14L)	G179X	G179X	G179X	NA	NA	NA	NA
B21R	OPG210	Surface glycoprotein cadherin-like domain putative membrane-associated glycoprotein	K431X; A475V; S533L; T713P; T723P; P726X; E1056D; A1207; A1339T; T1372A; C1715; V1725S; D1862E	K431X; A475V; S533L; T713P; T723P; P726X; E1056D; A1207; A1339T; T1372A; C1715; V1725S; D1862E	K431X; A475V; S533L; T713P; T723P; P726X; E1056D; A1207; A1339T; T1372A; C1715; V1725S; D1862E	K431X; A475V; S533L; T713P; T723P; P726X; E1056D; A1207; A1339T; T1372A; C1715; V1725S; D1862E	K431X; A475V; S533L; T713P; T723P; P726X; E1056D; A1207; A1339T; T1372A; C1715; V1725S; D1862E	K431X; A475V; S533L; T713P; T723P; P726X; E1056D; A1207; A1339T; T1372A; C1715; V1725S; D1862E	K431X; A475V; S533L; T713P; T723P; P726X; E1056D; A1207; A1339T; T1372A; C1715; V1725S; D1862E
#K1R		K1R	T12A	T12A	T12A	T12A	T12A	T12A	T12A
#R1R		R1R	I49M	I49M	I49M	I49M	I49M	I49M	I49M
N1R	0PG005	NA	NA	NA	NA	NA	NA	NA	NA
N4R	OPG015	Ankyrin (CPXV-017) N4R ankyrin-like	I48T; A249T; T273A	I48T; A249T; T273A	I48T; A249T; T273A	I48T; A249T; T273A	I48T; A249T; T273A	I48T; A249T; T273A	I48T; A249T; T273A
J1R	OPG003	Ankyrin (Cop-C19L) J1R ankyrin-like	T487I;T546	T487I; T546	K104X; T487I; T546	T487I; T546	T487I; T546	T487I; T546	T487I; T546
J2R	OPG002	TNF receptor (CrmB) (Cop-C22L) secreted TNF binding protein	L22I	L22I	L22I	L22I	L22I	L22I	L22I
J3R	OPG001	CC-chemokine binding Chemokine binding protein (Cop-C23L)	V194	V194	V194	V194	V194	V194	V194

#New/rare features.