Science
AAAS

Supplementary Materials for

Multiple lineages of monkeypox virus detected in the United States, 2021- 2022

Crystal M. Gigante, Bette Korber, Matthew H. Seabolt1, Kimberly Wilkins, Whitni Davidson, Agam K. Rao, Hui Zhao, Todd G. Smith, Christine M. Hughes, Faisal Minhaj, Michelle A. Waltenburg, James Theiler, Sandra Smole, Glen R. Gallagher, David Blythe, Robert Myers, Joann Schulte, Joey Stringer, Philip Lee, Rafael M. Mendoza, LaToya A. Griffin-Thomas, Jenny Crain, Jade Murray, Annette Atkinson, Anthony H. Gonzalez, June Nash, Dhwani Batra, Inger Damon, Jennifer McQuiston, Christina L. Hutson, Andrea M. McCollum, and Yu Li

Correspondence to: yuli@cdc.gov

This PDF file includes:

Materials and Methods

PCR testing

DNA was extracted from swabs collected from patient lesions using EZ-1 DNA tissue kit (Qiagen) followed by heat inactivation at 56°C for ≥ 1 hour. Monkeypox infection was confirmed by real-time PCR at CDC. CladeII-specific Monkeypox virus real-time PCR assay was performed as described in Li et al. (2010) (*15*). Orthopoxvirus OPX3 real-time PCR assay was performed as described (*14*), with the following changes: Each reaction (20 µL) contained 5 µL of template DNA, 0.5 µL of each primer (20 µM), 0.5 µL probe (10 µM) added to the 2X TaqMan Fast Advanced master mix (Applied Biosystems). Thermal cycling conditions for the ABI 7500 Fast Dx Real-Time PCR System (Applied Biosystems) were one cycle at 95°C for 20 seconds and 40 cycles at 95°C for 3 seconds and 60°C for 30 seconds. Primer and probe sequences for the Orthopoxvirus OPX3 real-time PCR assay (called OPX in (*14*)) are as follows:

Name	Sequence
OPX3 Forward	5'-TCA AAT ATT GAT CGT CCA ACG A-3'
OPX3 Reverse	5'-TGG ATG AAT TTC TCA ATA TTA GTT GG-3'
OPX3 Probe	5' FAM-TAA CAT CCG TCT GGA GAT ATC CCG TTA GA -BHQ1-3'

Sequencing

ONT: Library preparation was performed on extracted DNA using Ligation Sequencing kit (Oxford Nanopore Technologies SQK-LSK-109) following the manufacturer's protocol for genomic DNA. Libraries were sequenced (one sample per flow cell) on a MinION sequencer (MPXV_USA_2022_MA001, MPXV_USA_2021_MD, and MPXV_USA_2021_TX) or GridION sequencer (FL001, FL002, CA001, UT001, UT002, VA001) using a MIN109 R9.4.1 flow cell (Oxford Nanopore). For MA001, CA001, and VA001 data were combined from two independent swabs collected from two locations. For all other cases, data are from a single sample/swab. Basecalling was performed using guppy version 6.1.2 with high accuracy and qscore filtering (for MinION runs) or was performed on GridION using high accuracy basecalling.

Illumina: Extracted DNA (40 µl) was diluted to 130 µl in water and fragmented to 500bp on a Covaris S220 instrument using SonoLab 7 application software. Genomic DNA libraries were prepared using Swift Accel-NGS® 2S Plus DNA Library Kit (Cat# 21096) plus 2S Set A+B Indexing Kit (Cat#26396), following manufacturer's protocol using 15 cycles of PCR. Size selection was performed using Swift Accel-NGS® 2S Plus DNA Library Kit (Cat# 21096) to ensure a single peak of approximately 500bp. Libraries were prepared and normalized according to the manufacturer's recommendations to create a final DNA library with a concentration of 15 - 20nM/5ul which was then sequenced on a MiSeq instrument using MiSeq® Reagent kit v3 600 cycles (Cat# MS-102-3003, Illumina). DNA from the 2021_MD case underwent hybridization with a custom orthopoxvirus capture probe panel prior to processing and sequencing as described in (*3*).

Genome Assembly

Nanopore reads were trimmed to remove 55 bp from each end (seqtk 1.0, https://github.com/lh3/seqtk) and all reads below 50 bp were removed (trimmomatic 0.39, https://github.com/timflutre/trimmomatic) before mapping to MPXV Nigeria reference MT903344 with 6,000 bp removed from the left terminus using minimap2 2.16 (https://github.com/lh3/minimap2) to remove human and other non-MPXV reads. Illumina reads were trimmed using FaQCs 1.34 (https://github.com/LANL-Bioinformatics/FaQCs) using parameters -q 20 --5end 10 --3end 5 -n 5 --min_L 30 and then mapped to MPXV Nigeria reference MT903344 with 6,000 bp removed from the left terminus using bwa mem (https://github.com/lh3/bwa). A hybrid assembly was generated from mapped Illumina and Oxford Nanopore reads using Unicyler 0.4.7 (https://github.com/rrwick/Unicycler). Assemblies were polished by mapping reads back to draft genomes containing one complete ITR and one incomplete ITR with ~6,000 bases removed using bwa mem or minimap2 and generating a consensus

sequence using samtools 1.9 and ivar 1.0 (https://github.com/andersen-lab/ivar). Inverted Terminal Repeats (ITRs) were assembled manually by copying from one end to the other.

For quality control, separate assemblies were made with either nanopore or Illumina data since DNA extracted from separate lesions for the same patient were used for Nanopore and Illumina sequencing in three cases. Illumina-only assemblies were made using SPAdes/3.13.0 and CLC Genomics Workbench 22. Oxford Nanopore-only assemblies were made using flye 2.9 (https://github.com/fenderglass/Flye).

Annotations were transferred from MPXV Clade 3 Nigeria reference MT903345, then locus_tags were re-named with the strain ID. Alignments used to make trees were performed using MAFFT v.7.450 (*37*). Phylogenetic analysis (Figure 1) was performed using BEAST v. 1.8.3. All sites containing gaps were removed prior to phylogenetic analysis, and all sites containing gaps or ambiguities were removed prior to haplotype network analysis. Haplotype network analysis was generated using sequence alignments generated by whole genome alignment using MAFFT v.7.450 (*37*) followed by removal of alignment columns containing gaps or ambiguities with PopArt using the Median Joining method. An analysis of evolutionary rate was performed in BEAST v.1.8.3 using an alignment of 85 monkeypox virus genomes (Figure 2) after stripping all columns containing gaps. The GTR+g+i model, fixed local clock (uniform distribution 0,1), tip dates (year ± 0.5), and a coalescent Bayesian skyline tree prior with 10 groups were used. Clades I, IIa, and Lineage A were specified as taxon sets. A secondary analysis of evolutionary rate was performed for Clades I and IIa (54 monkeypox virus genomes, Figure 2) using the same alignment after removing Clade IIb sequences; GTR+g+i model, fixed local clock (uniform distribution 0,1), tip dates (year ± 0.5), and a coalescent Bayesian skyline tree prior with 10 groups were used. Lineage A sequences and KY642617 were extracted from the alignment used in Figure 2 and evolutionary rate was analyzed using both fixed local clock (for Lineage A only) and uncorrelated lognormal relaxed clock (uniform distribution 0,1). GTR+g+i model, tip dates (year ± 0.5), and a coalescent Bayesian skyline tree prior with 4 groups were used in both analyses. Sequences of all near complete MPXV genomes that were available at the time of analysis were used. Still, many genomes contained Ns and may contain errors caused by low coverage or sequencing technology bias. Availability of complete, final genomes is expected to improve resolution of future analyses.

APOBEC3 analysis

The sequence alignment we used began with foundational sequences from Mauldin et. al., a paper that described spread of MPXV beyond the African continent (*3*), adding additional sequences to create our baseline dataset of 85 MPXV-related sequences. Our initial alignment was generated using MAFFTv7 (*37*); ends were trimmed to exclude regions that were only sporadically sequenced. The alignment further edited to resolve poorly aligned regions; generally these were in regions with differing numbers of direct repeats or that were proximal to long gaps in sequences. A maximum likelihood tree used for defining clades of interest and reconstructing ancestral states for subsequent APOBEC3 analysis in was generated using the HIV database web interface for IQ-tree (https://www.hiv.lanl.gov/content/sequence/IQTREE/iqtree.html) (*38*), with ModelFinder (*39*); the Bayesian information criterion (BIC) best fit model was K3Pu+F+I, and a mid-point root was used. GenBank Accession numbers for these sequences are included in Tables S6 (Lineage A) and S8 (other Clade II and Clade I).

APOBEC analysis was performed by tracking all observed single nucleotide mutational events relative to the most common ancestor sequence within a given clade of interest (Figs. 2 an S2). This analysis was first performed on natural sequences (Tables S6-S8) to explore mutational patterns in sub-different lineages within the tree. To exclude redundancy due to shared internal branches, we also compressed all SNPS observed within a clade onto a single merged sequence that included all unique mutations found within that clade (mergesnps.py, https://zenodo.org/record/6970457#.YvAVPi-B1eg) (*36*). This merged sequence was then compared to the most recent common ancestor of the clade generated with IQ-tree as shown in Fig. S2, and such a merged comparisons are included in Fig. 2B and Fig S3, and Table S6 – S8.

The statistical strategies for resolving if G-to-A mutations within an APOBEC3 motif context are enhanced relative other G-to-A mutations are described in the HIV-database tool HYPERMUT (https://www.hiv.lanl.gov/content/sequence/HYPERMUT/hypermut.html). HYPERMT was written for RNA viruses, so for this study we adapted the approach to a double stranded DNA scenario and built improved visualization tools. The method makes pairwise comparisons between a reference sequence and each of a set of sequences aligned to that reference. It identifies and tallies all Gs that occur in the context of APOBEC3 motifs within a reference sequence and counts both unchanged Gs and G-to-A mutations that occur within the context of that motif. Similarly, all Gs that are not embedded in an APOBEC3 motif are tallied, and the number of G-to-A mutations that occur among those G's are counted. A Fisher's exact test is then used to determine if the G-to-A substitutions events are significantly enriched in the context of APOBEC3 motifs. Q-values were calculated to correct for multiple tests (*40*, *41*). Gaps and IUPAC ambiguity codes can be present in the input sequences, but these are excluded from the comparisons. The original HYPERMUT strategy was developed for HIV-1, so we adapted it for application to MPXV by enabling simultaneously tracking G-to-A mutations in the forward and reverse complement strand using apobec.py, apocount.py, and apoplot.py (available at https://zenodo.org/record/6970457#.YvAVPi-B1eg) (*36*). This was critical as what simply appears as C-to-T transition on the first strand can be embedded in the APOBEC3 motif in the reverse complement. Our original HYPERMUT enables the inclusion of more complex APOBEC3 motifs, and we also explored the use of the 3-base motif pattern including a cytosine in the +2 positions (GAC-to-AAC, or GGC-to-AGC), as pattern this can partially inhibit APOBEC3 activity (*42*); as this strategy did not substantively change our conclusions, we present the simpler version here. Also, as a cross-check we confirmed our results with ancestral reconstruction-based analysis by using natural sequences that were most proximal to ancestral nodes as a reference strain instead; theses analyses concurred. Our method distinguishes between GA-to-AA and GG-to-AG motifs, and almost all of the substitutions we observed were in the GA-to-AA context, indicative of APOBEC3D and 3F activity, not APOBEC3G (*43*).

Sensitivity of 2022 outbreak MPXV to TPOXX

A cytopathic effect (CPE) assay was used in a 96-well format. 0.0015 – 5 µM of tecovirimat (lot: SG-14G32-M) was added to a confluent monolayer of Vero E6 cells and incubated 1 hour at 37°C and 6.0% CO_2. After incubation MPXV was added (MOI = 0.01) and incubated 72 hours at 35.5°C and 6.0% CO_2. To avoid edge effects, outer wells were filled with an equivalent volume of medium except for virus-only control wells which were located on the edge of the plate. Plates were inactivated, fixed, and stained using formalinized crystal violet, washed with water, and the absorbance at 570 nm was measured. Half-maximal effective concentration (EC_{50}) was calculated using a non-linear curve fit with variable slope (four parameters) in GraphPad Prism v. 6.07 using 24 values from 4 statistical replicates from 2 biological replicates. Controls included virus-only, virus plus vehicle, cell-only, and cells plus vehicle wells as well as a drug-only plate to control for cytotoxicity.

.

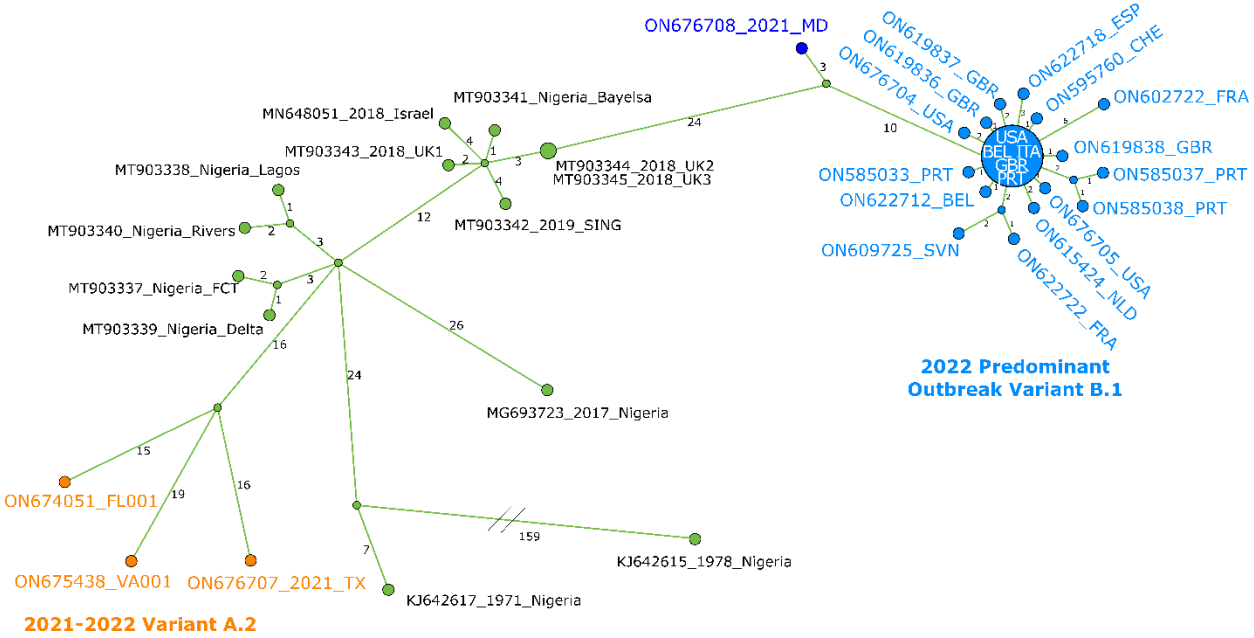


Fig. S1. Nucleotide changes among Clade IIb MPXV genome sequences. The predominant 2022 MPXV outbreak variant B.1 and outbreak variant A.2 are highlighted in blue and orange, respectively. The USA_2021_MD sequence is highlighted in dark blue. The large node at the center of variant B.1 represents 13 identical sequences; country abbreviations are given for sake of space (sequences used can be found in Table S1). GenBank accession numbers are given for all other samples. Sequence differences between nodes are indicated by the numbers on the branches. Unlabeled nodes represent hypothetical common ancestors, lines connecting nodes do not represent direct links between cases. GBR: United Kingdom; USA: United States of America; FRA: France; BEL: Belgium; PRT: Portugal; ITA: Italy; ESP: Spain; NLD: Netherlands; CHE: Switzerland; SVN: Slovenia.

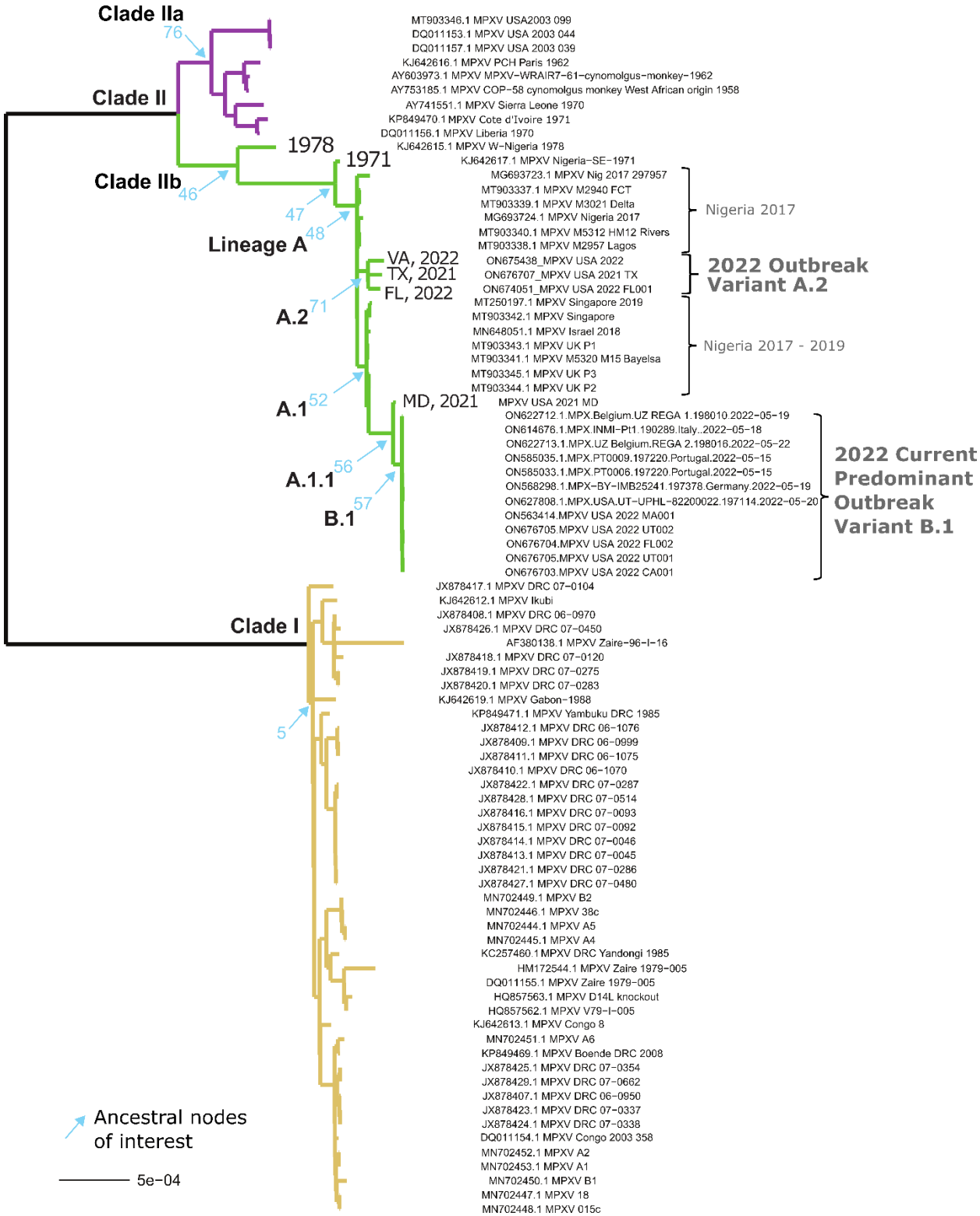
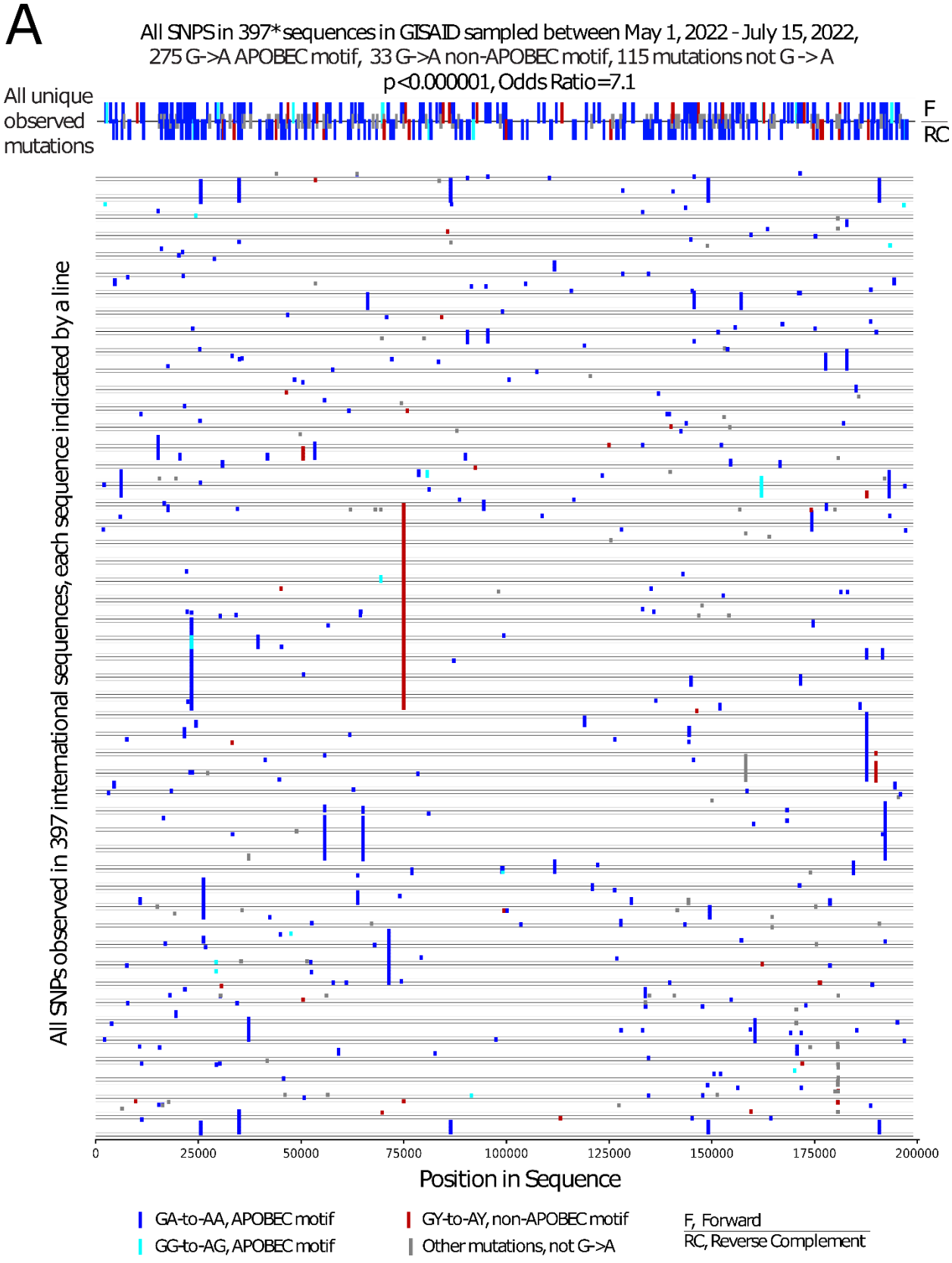
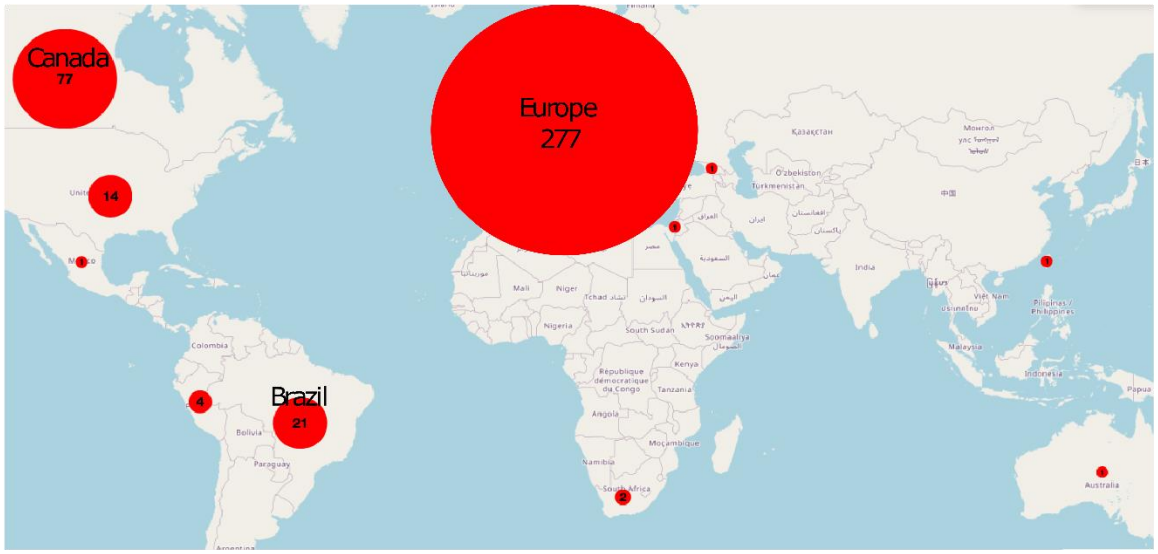


Fig. S2. A detailed version of the ML tree shown in main text Fig. 2A. Ancestral nodes of interest are noted here, and these are used to track the statistical exploration of G-to-A mutations in an APOBEC context through different lineages in the tree summarized in Fig. 2B in the main text and provided in detail in Tables S6 (for Lineage A) and S8 (for Clade I and II outside Lineage A). The sequence name of each of the taxa are shown, preceded by their GenBank accession number.


A
All SNPS in 397* sequences in GISAID sampled between May 1, 2022 - July 15, 2022,
275 G->A APOBEC motif, 33 G->A non-APOBEC motif, 115 mutations not G -> A
p<0.000001, Odds Ratio=7.1
All unique observed mutations
F
RC
All SNPs observed in 397 international sequences, each sequence indicated by a line
0
25000
50000
75000
100000
125000
150000
175000
200000
Position in Sequence
GA-to-AA, APOBEC motif
GG-to-AG, APOBEC motif
GY-to-AY, non-APOBEC motif
Other mutations, not G->A
F, Forward
RC, Reverse Complement

B Regional distribution of monkeypox B.1 sequences available at the time of sampling, July 15, 2022



European region expanded, by country

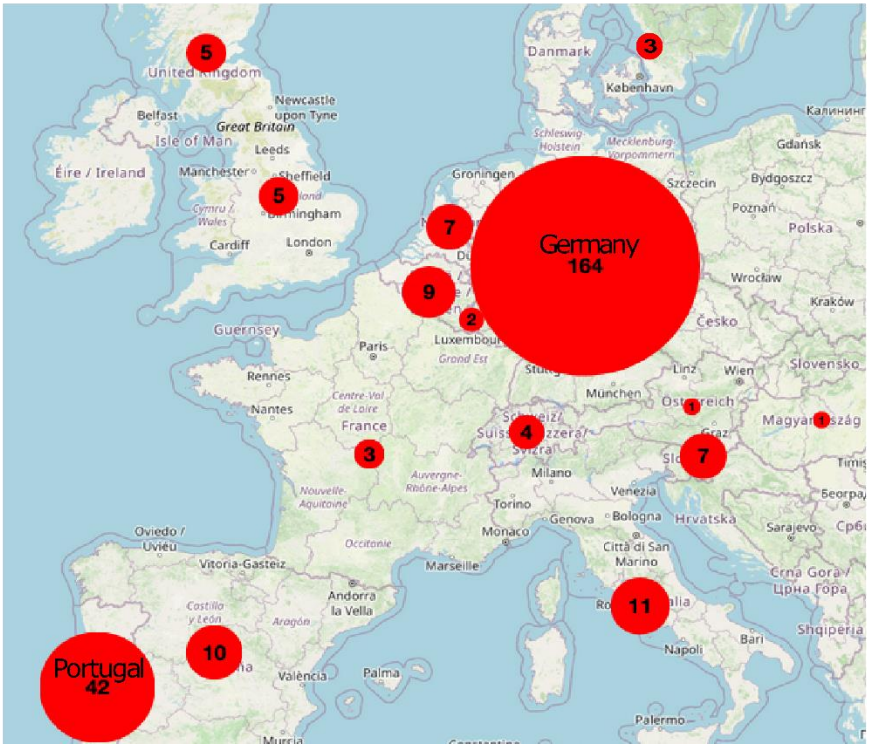
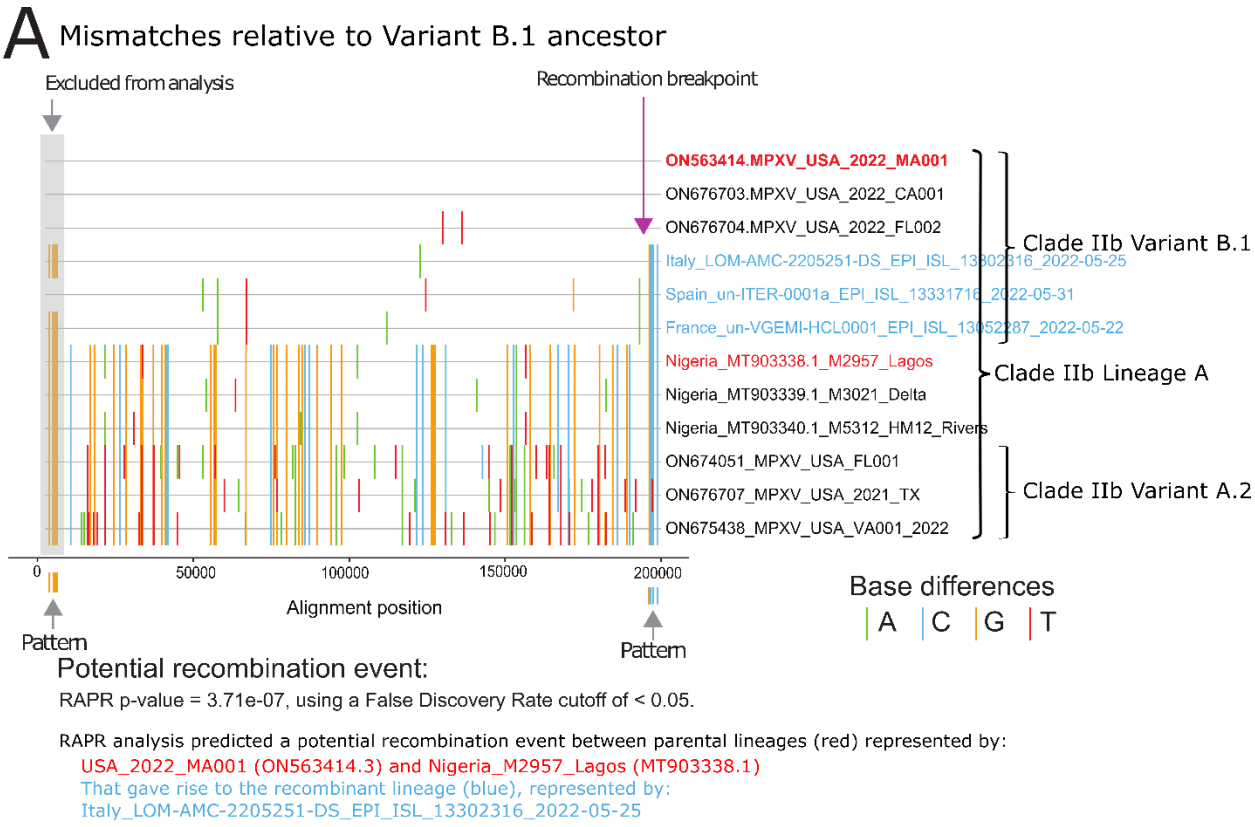


Figure S3. A. Details of mutational patterns among 397 variant B.1 sequences available through GISAID that were sampled between May 1, 2022, and July 15, 2022. Each horizontal line represents a unique sequence, and tick marks indicate all mutations relative to the ancestor of the 2022 outbreak variant B.1. Sequences that are phylogenetically similar based on shared mutations are proximal to each other along the Y-axis, and so vertical columns of mutations are indicative of shared mutations between sequences. One would expect roughly the same number of red and blue mutations if G-to-A mutations were randomly occurring; the preponderance of blue ticks indicates the extreme bias favoring G-to-A

mutations embedded in a 5' GA-to-AA context among these international 2022 outbreak strains. The GA-to-AA pattern dominates (dark blue) over the alternative APOBEC motif 5' GG-to-AG (cyan), as we have seen throughout Lineage A. 114 sequences among the 397 variant B.1 sequences had no SNPS relative to the ancestral form of the genome and these are not shown in the figure. GY-to-AY (red) indicates a G-to-A change followed by a C or T, the G-to-A mutations that are not in an APOBEC motif. **B. Geographic distribution of sampling of the 397 sequences used for part A, based on the sequences available from GISAID sampled between May 1, 2022, and July 15, 2022.** The relative area of a given red circle reflects the number of available sequences. This map captures the level of contributions of early monkeypox genomes to GISAID the geographic distribution of the available data and is not meant to reflect the confirmed cases. The GISAID acknowledgments tables are provided in Tables S10 and S11.



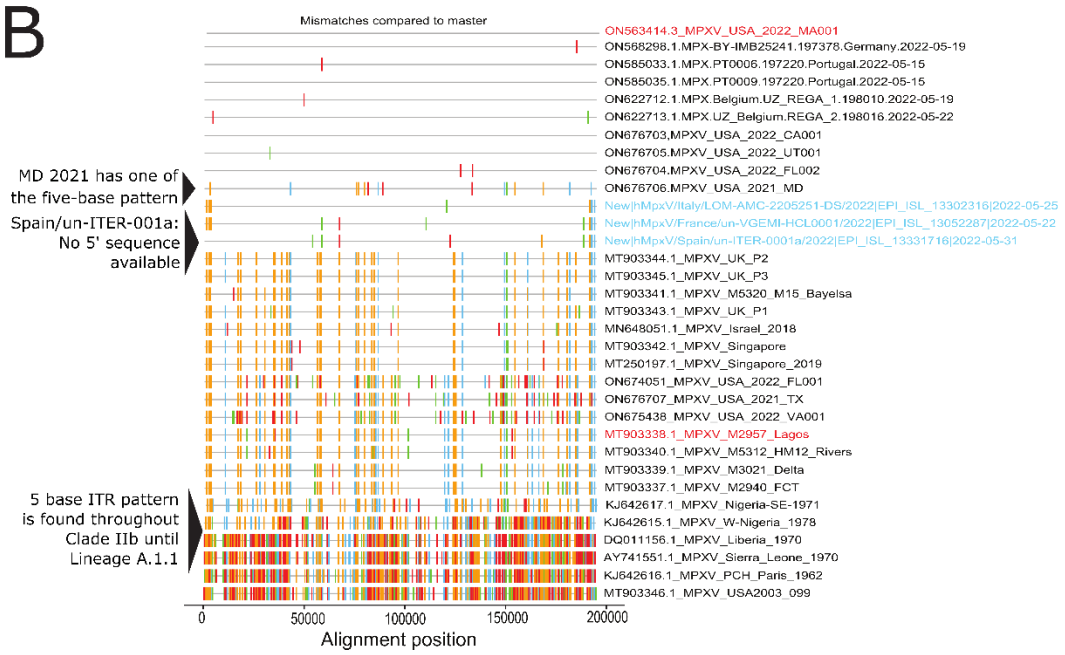
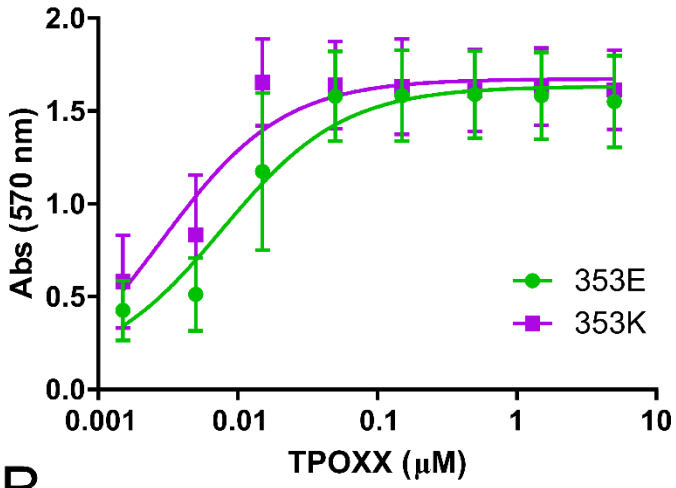


Fig. S4. Chimeric sequences within Clade IIb. A. We acquired 400 full genome MPXV sequences from GISAID sampled between May 1, 2022, and July 15, 2022, for APOBEC analysis. While confirming their lineage association, we identified 3 related sequences that were chimeric in that they carried 5

consecutive SNPs that were mirrored in each ITR that were shared with more ancestral sequences from Lineage A and Clade IIb but not found in variant B.1. Otherwise, these sequences were B.1-like throughout the remainder of the genome. These sequences originated from three different European laboratories*. The graphic is a "Highlighter" plot (a www.hiv.lanl.gov tool) that shows every SNP relative to the USA 2022 MA001 sequence for a small set of background sequences used to explore that hypothesis that the chimeric viruses were the result of recombination. Using the Recombination Analysis Program (RAPR) (PMID: 29765018 https://www.hiv.lanl.gov/content/sequence/RAP2017/rap.html), and excluding the 5' ITR from the analysis so we did not overcount the repeated mutations in both ITRs (marked in gray). The putative recombinant sequence from Spain was incomplete and did not span the 5' ITR region, therefore we used the 3' region for analysis to enable including all three related forms. We determined that a string of 5 mutations in series that differed from the representative variant B.1 sequences and that are shared with more ancestral Clade IIb sequences is unlikely to have occurred by chance alone (p-value 3.71e-07). Thus, the RAPR analysis supports this being a recombinant lineage, although alternatively this pattern may have emerged as a systematic sequencing artifact. In either case, the analysis indicated these sequences were chimeric and not simple variant B.1 viruses, and so we removed them from the subsequent analysis of variant B.1 sequences in Fig. S3. Red sequence names indicate sequences identified by RAPR as representative of candidate parental lineages, and blue the putative recombinant lineage. **B.** The pattern of interest in the ITR was consistently found in sequences throughout Clade IIb, including samples from the 1970s through contemporary samples excluding the B.1 variant. In the 2021_MD sequence, four of the five variant B.1 mutations were apparent. The GISAID acknowledgments tables are provided in Tables S10 and S11.

*__EPI_ISL_13302316__: Laboratory of Clinical Microbiology, Virology and Bioemergencies. ASST-Fatebenefratelli-Sacco, L.Sacco University Hospital, Milano, Italy; __EPI_ISL_13331716__: Genomics Division, Instituto Tecnologico y de Energias Renovables (ITER), Poligono Industrial de Granadilla Santa Cruz de Tenerife, Spain (the 3' ITR sequence was not available); __EPI_ISL_13052287__: Virology, GENomique EPIdemiologique des maladies Infectieuses, Lyon, France.

A.



B.

F13L variant	Isolate	EC_{50} (µM)
353E	Aggregate	0.007731 ± 0.002
353K	Aggregate	0.002860 ± 0.001
353E	MPXV Clade IIa	0.01754 ± 0.007
353E	MPXV_USA_2022_FL001	0.008412 ± 0.004
353E	MPXV_USA_2022_VA001	0.006001 ± 0.002
353E	MPXV_USA_2021_MD	0.003621 ± 0.003
353K	MPXV_USA_2022_MA001	0.002680 ± 0.002
353K	MPXV_USA_2022_FL002	0.003391 ± 0.002
353K	MPXV_USA_2022_UT001	0.002739 ± 0.002
353K	MPXV_USA_2022_UT002	0.002658 ± 0.002

Fig. S5. Sensitivity of 2022 outbreak MPXVs to tecovirimat (TPOXX). A. Aggregate results of cytopathic effect (CPE) assay showing cell growth in the presence of MPXV after treatment with different doses of TPOXX. Error bars indicate 95% confidence intervals based on four statistical and two biological replicates at each dose per group. **B.** Half-maximal effective concentration EC_{50} of TPOXX for different MPXV isolates. Average plus 95% confidence intervals were based on 24 values from four statistical and two biological replicates.

Accession numbers for Figure S1	
ON676703.1	ON676708.1
ON563414.3	ON585037.1
ON622713.1	ON585032.1
ON585035.1	MT903345.1
ON568298.1	MT903344.1
ON614676.1	MT903343.1
ON563414.2	MT903341.1
ON627808.1	MT903342.1
ON585033.1	MN648051.1
ON622712.1	MT903340.1
ON676705.1	MT903338.1
ON676706.1	MT903339.1
ON595760.1	MT903337.1
ON676704.1	MG693723.1
ON585031.1	ON676707.1
ON615424.1	ON674051.1
ON622722.1	ON675438.1
ON609725.1	KJ642617.1
ON622718.1	ON619835.1
ON585038.1	ON619836.1
ON602722.1	ON619837.1
ON585034.1	ON619838.1
	KJ642615.1

Table S1. List of GenBank accession numbers for sequences used in haplotype network analysis (Fig. S1).

Position	reference	MA001-2022	MD-2021	TYPE	FTYPE	STRAND	EFFECT	NT CHANGE	AA CHANGE ALT	GENE	PRODUCT	Genbank Note
1271	G	A	A	snp	CDS	-	missense_variant	314C>T	Ser105Leu	MPXV-UK_P2-001	MPXVgp001	Chemokine binding protein (Cop-C23L); J1L; similar to Vaccinia virus strain Copenhagen C23L; most abundant secreted protein; CC-chemokine binding
2600	G	A	A	snp	CDS	-	missense_variant	161C>T	Ser54Phe	MPXV-UK_P2-002	MPXVgp002	TNF receptor (CrmB) (Cop-C22L); J2L
3120	G	A	.	snp	CDS	-	synonymous_variant	1497C>T	Ile499Ile	MPXV-UK_P2-003	MPXVgp003	Ankyrin (Cop-C19L); J3L
3531	G	A	A	snp	CDS	-	synonymous_variant	1086C>T	Ile362Ile	MPXV-UK_P2-003	MPXVgp003	Ankyrin (Cop-C19L); J3L
3827	C	T	T	snp	CDS	-	missense_variant	790G>A	Asp264Asn	MPXV-UK_P2-003	MPXVgp003	Ankyrin (Cop-C19L); J3L
7780	C	T	T	snp	CDS	+	synonymous_variant	192C>T	Ile64Ile	MPXV-UK_P2-006	MPXVgp006	Secreted EGF-like protein (Cop-C11R); D3R; similar to Vaccinia virus strain Copenhagen C11R; secreted growth factor
14009	G	T	T	snp	CDS	-	missense_variant	1268C>A	Ala423Asp	MPXV-UK_P2-012	MPXVgp012	Ankyrin; Type I IFN resistance (Cop-C9L); D9L; similar to Vaccinia virus strain Copenhagen C9L; ankyrin-like
15437	G	A	A	snp			intergenic_variant					
21732	G	A	A	snp	CDS	-	synonymous_variant	711C>T	Phe237Phe	MPXV-UK_P2-021	MPXVgp025	ANK-containing protein; apoptosis inihibitor (Cop-M1L); O1L; similar to Vaccinia virus strain Copenhagen M1L; ankyrin-like
30376	G	A	A	snp	CDS	-	synonymous_variant	828C>T	Leu276Leu	MPXV-UK_P2-031	MPXVgp035	Kelch-like protein (Cop-F3L); C9L; similar to Vaccinia virus strain Copenhagen F3L; kelch-like
31062	G	A	A	snp	CDS	-	missense_variant	142C>T	Arg48Cys	MPXV-UK_P2-031	MPXVgp035	Kelch-like protein (Cop-F3L); C9L; similar to Vaccinia virus strain Copenhagen F3L; kelch-like
34468	G	A	A	snp	CDS	-	missense_variant	232C>T	Pro78Ser	MPXV-UK_P2-037	MPXVgp041	S-S bond formation pathway protein substrate (Cop-F9L); C15L; similar to Vaccinia virus strain Copenhagen F9L
37211	G	A	A	snp	CDS	-	synonymous_variant	1833C>T	Phe611Phe	MPXV-UK_P2-040	MPXVgp044	EEV maturation protein (Cop-F12L); C18L; similar to Vaccinia virus strain Copenhagen F12L; actin tail formation
38369	G	A	A	snp	CDS	-	synonymous_variant	675C>T	Ile225Ile	MPXV-UK_P2-040	MPXVgp044	EEV maturation protein (Cop-F12L); C18L; similar to Vaccinia virus strain Copenhagen F12L; actin tail formation
38671	C	T	T	snp	CDS	-	missense_variant	373G>A	Glu125Lys	MPXV-UK_P2-040	MPXVgp044	EEV maturation protein (Cop-F12L); C18L; similar to Vaccinia virus strain Copenhagen F12L; actin tail formation
39148	C	T	.	snp	CDS	-	missense_variant	1057G>A	Glu353Lys	MPXV-UK_P2-041	MPXVgp045	C19L similar to Vaccinia virus strain Copenhagen F13L major envelope antigen of EEV wrapping of IMV to form IEV phospholipase D-like
52894	G	A	A	snp	CDS	-	synonymous_variant	1554C>T	Val518Val	MPXV-UK_P2-055	MPXVgp057	DNA polymerase (Cop-E9L); F8L; similar to Vaccinia virus strain Copenhagen E9L; DNA polymerase, catalytic subunit
54126	G	A	A	snp	CDS	-	missense_variant	322C>T	Leu108Phe	MPXV-UK_P2-055	MPXVgp057	DNA polymerase (Cop-E9L); F8L; similar to Vaccinia virus strain Copenhagen E9L; DNA polymerase, catalytic subunit
54644	G	A	A	snp	CDS	+	missense_variant	166G>A	Asp56Asn	MPXV-UK_P2-056	MPXVgp058	Sulfhydryl oxidase (FAD-linked) (Cop-E10R); F9R; similar to Vaccinia virus strain Copenhagen E10R; protein disulfide bond-forming enzyme
64306	G	A	A	snp	CDS	-	synonymous_variant	420C>T	Ile140Ile	MPXV-UK_P2-066	MPXVgp068	Virion core cysteine protease (Cop-I7L); I7L; similar to Vaccinia virus strain Copenhagen I7L; virion core protein; similar to DNA topoisomerase II
73075	C	T	T	snp	CDS	+	missense_variant	89C>T	Ser30Leu	MPXV-UK_P2-076	MPXVgp078	VLTF-1 (late transcription factor 1) (Cop-G8R); G9R; similar to Vaccinia virus strain Copenhagen G8R; late gene transcription factor, VLTF-1
73248	G	A	.	snp	CDS	+	missense_variant	262G>A	Asp88Asn	MPXV-UK_P2-076	MPXVgp078	VLTF-1 (late transcription factor 1) (Cop-G8R); G9R; similar to Vaccinia virus strain Copenhagen G8R; late gene transcription factor, VLTF-1
74214	G	A	.	snp	CDS	+	missense_variant	426G>A	Met142Ile	MPXV-UK_P2-077	MPXVgp079	Entry/fusion complex component, myristylprotein (Cop-G9R); G10R; similar to Vaccinia virus strain Copenhagen G9R; myristylated protein
77392	G	A	.	snp	CDS	+	missense_variant	484G>A	Glu162Lys	MPXV-UK_P2-081	MPXVgp083	ss/dsDNA binding protein (VP8) (Cop-L4R); M4R; similar to Vaccinia virus strain Copenhagen L4R; virion core protein; ssDNA binding; stimulation of I8R helicase activity
79357	C	.	T	snp	CDS	+	synonymous_variant	264C>T	Ile88Ile	MPXV-UK_P2-085	MPXVgp087	L3R similar to Vaccinia virus strain Copenhagen J3R poly-A polymerase stimulatory subunit simultaneously cap-specific mRNA (nucleoside-O2'-)-methyltransferase
81284	G	A	A	snp	CDS	+	synonymous_variant	150G>A	Lys50Lys	MPXV-UK_P2-088	MPXVgp090	RNA polymerase subunit (RPO147) (Cop-J6R); L6R; similar to Vaccinia virus strain Copenhagen J6R; RNA polymerase, 147 kDa subunit
82382	C	T	T	snp	CDS	+	synonymous_variant	1248C>T	Phe416Phe	MPXV-UK_P2-088	MPXVgp090	RNA polymerase subunit (RPO147) (Cop-J6R); L6R; similar to Vaccinia virus strain Copenhagen J6R; RNA polymerase, 147 kDa subunit
82460	G	A	A	snp	CDS	+	synonymous_variant	1326G>A	Thr442Thr	MPXV-UK_P2-088	MPXVgp090	RNA polymerase subunit (RPO147) (Cop-J6R); L6R; similar to Vaccinia virus strain Copenhagen J6R; RNA polymerase, 147 kDa subunit
84596	C	T	.	snp	CDS	+	synonymous_variant	3462C>T	Phe1154Phe	MPXV-UK_P2-088	MPXVgp090	RNA polymerase subunit (RPO147) (Cop-J6R); L6R; similar to Vaccinia virus strain Copenhagen J6R; RNA polymerase, 147 kDa subunit
86988	C	.	T	snp	CDS	-	synonymous_variant	81G>A	Glu27Glu	MPXV-UK_P2-091	MPXVgp093	H3L similar to Vaccinia virus strain Copenhagen H3L IMV surface membrane protein
95043	G	A	A	snp	CDS	+	synonymous_variant	375G>A	Val125Val	MPXV-UK_P2-098	MPXVgp100	Virion core (Cop-D3R); E3R; similar to Vaccinia virus strain Copenhagen D3R; virion core protein
124139	G	A	A	snp	CDS	+	missense_variant	184G>A	Glu62Lys	MPXV-UK_P2-128	MPXVgp129	DNA helicase, transcript release factor (Cop-A1 8R); A19R; similar to Vaccinia virus strain Copenhagen A1 8R; virion core associated DNA helicase; post-replicative negative transcription elongation factor
124683	G	A	A	snp	CDS	+	missense_variant	728G>A	Arg243Gln	MPXV-UK_P2-128	MPXVgp129	DNA helicase, transcript release factor (Cop-A1 8R); A19R; similar to Vaccinia virus strain Copenhagen A1 8R; virion core associated DNA helicase; post-replicative negative transcription elongation factor
128707	C	T	T	snp	CDS	+	missense_variant	920C>T	Ser307Leu	MPXV-UK_P2-133	MPXVgp134	VITF-3 45kda subunit (Cop-A23R); A24R; similar to Vaccinia virus strain Copenhagen A23R; intermediate transcription factor, VITF-3, 45 kDa
133174	GCAATCTTTCTCAATCTTTCTCAA	GCAATCTTTCTCAA	GCAATCTTTCTCAA	indel			intergenic_variant					
133725	C	.	T	snp			intergenic_variant					
150480	C	T	.	snp	CDS	+	missense_variant	661C>T	His221Tyr	MPXV-UK_P2-157	MPXVgp157	IL-1/TLR signaling inhibitor (Cop-A46R); A47R; similar to Vaccinia virus strain Copenhagen A46R
150561	GGATATGATGGATATGATGGATATGATGATGGATGATG	GGATATGATGGATATGATGATGGATGATG	.	indel			intergenic_variant					
151472	A	C	.	snp			intergenic_variant					
155806	G	A	.	snp			intergenic_variant					
162342	C	T	T	snp	CDS	+	synonymous_variant	736C>T	Leu246Leu	MPXV-UK_P2-164	MPXVgp165	Schlafen (Cop-B2R); B4R; schlafen-like
169732	TCAGATACAGATACAGATACAGATACAGATACAGATACAGAT	.	TCAGATACAGATACAGATACAGATACAGATACAGATACAGAT	indel	CDS	+	conservative_inframe_deletion	100_105delACAGAT	Thr34_Asp35del	MPXV-UK_P2-171	Hypothetical pro	Hypothetical protein (Cop-B11R)
170273	G	A	.	snp	CDS	+	synonymous_variant	225G>A	Arg75Arg	MPXV-UK_P2-172	MPXVgp172	Ser/Thr Kinase (Cop-B12R); B11R; similar to Vaccinia virus strain Copenhagen B12R; protein kinase-like
178220	G	A	A	snp			intergenic_variant					
179163	complex	complex	complex	indel			intergenic_variant					
181995	G	A	A	snp	CDS	+	missense_variant	625G>A	Asp209Asn	MPXV-UK_P2-182	MPXVgp182	Surface glycoprotein; B21R; putative membrane-associated glycoprotein; cadherin-like domain
183534	C	T	.	snp	CDS	+	missense_variant	2164C>T	Pro722Ser	MPXV-UK_P2-182	MPXVgp182	Surface glycoprotein; B21R; putative membrane-associated glycoprotein; cadherin-like domain
186593	G	A	A	snp	CDS	+	missense_variant	5223G>A	Met1741Ile	MPXV-UK_P2-182	MPXVgp182	Surface glycoprotein; B21R; putative membrane-associated glycoprotein; cadherin-like domain
192418	complex	complex	complex	indel			intergenic_variant					
193407	G	A	A	snp	CDS	+	missense_variant	790G>A	Asp264Asn	MPXV-UK_P2-188	MPXVgp189	Ankyrin (Cop-C19L); J1R; ankyrin-like
193703	C	T	T	snp	CDS	+	synonymous_variant	1086C>T	Ile362Ile	MPXV-UK_P2-188	MPXVgp189	Ankyrin (Cop-C19L); J1R; ankyrin-like
194114	C	T	.	snp	CDS	+	synonymous_variant	1497C>T	Ile499Ile	MPXV-UK_P2-188	MPXVgp189	Ankyrin (Cop-C19L); J1R; ankyrin-like (ITR)
194634	C	T	T	snp	CDS	+	missense_variant	161C>T	Ser54Phe	MPXV-UK_P2-189	MPXVgp190	TNF receptor (CrmB) (Cop-C22L); J2R; secreted TNF binding protein
195963	C	T	T	snp	CDS	+	missense_variant	314C>T	Ser105Leu	MPXV-UK_P2-190	MPXVgp191	Chemokine binding protein (Cop-C23L); J3R; similar to Vaccinia virus strain Copenhagen B29R; CC-chemokine binding

Table S2. Variant table showing differences in genomes of USA_2022_MA001 (ON563414.3) and USA_2021_MD (ON676708.1) compared to MT903344.1 UK-P2. Changes shared among USA_2022_MA001 and USA_2021_MD are highlighted in green. Position and annotation information is based on reference MT903344.1.

Position	reference	TX-2021	FL001-2022	VA001-2022	TYPE	FTYPE	STRAND	EFFECT	NT CHANGE	AA CHANGE AL	GENE	PRODUCT	Genbank Note
2761	G	A	.	.	snp			intergenic_variant					
11341	G	.	.	A	snp	CDS	-	synonymous_variant	1569C>T	Ile523Ile	MPXV-UK_P2-010	MPXVgp010	D7L host range ankyrin-like
12108	G	.	.	A	snp	CDS	-	stop_gained	802C>T	Arg268*	MPXV-UK_P2-010	MPXVgp010	D7L host range ankyrin-like
13307	C	T	T	T	snp			intergenic_variant					
13657	C	.	.	T	snp	CDS	-	missense_variant	1620G>A	Met540Ile	MPXV-UK_P2-012	MPXVgp012	D9L similar to Vaccinia virus strain Copenhagen C9L ankyrin-like
15114	C	.	.	T	snp	CDS	-	missense_variant	163G>A	Asp55Asn	MPXV-UK_P2-012	MPXVgp012	D9L similar to Vaccinia virus strain Copenhagen C9L ankyrin-like
16236	C	.	.	T	snp	CDS	-	missense_variant	151G>A	Asp51Asn	MPXV-UK_P2-013	MPXVgp013	D10L similar to Vaccinia virus strain Copenhagen C7L host range
18801	C	T	T	T	snp	CDS	-	missense_variant	37G>A	Asp13Asn	MPXV-UK_P2-016	MPXVgp016	D13L similar to Vaccinia virus strain Copenhagen C4L
23573	T	C	C	C	snp	CDS	-	synonymous_variant	555A>G	Ser185Ser	MPXV-UK_P2-023	MPXVgp027	C1L similar to Vaccinia virus strain Copenhagen K1L host range ankyrin-like
25081	C	.	T	.	snp	CDS	-	missense_variant	407G>A	Arg136Gln	MPXV-UK_P2-024	MPXVgp028	SPI-3 C2L similar to Vaccinia virus strain Copenhagen K2L serine protease inhibitor-like SPI-3 inhibition of the ability of infected cells to fuse
25670	A	G	G	G	snp	CDS	-	missense_variant	107T>C	Phe36Ser	MPXV-UK_P2-025	MPXVgp029	C3L
29896	C	.	.	T	snp	CDS	-	synonymous_variant	1308G>A	Pro436Pro	MPXV-UK_P2-031	MPXVgp035	C9L similar to Vaccinia virus strain Copenhagen F3L kelch-like
30645	C	T	T	T	snp	CDS	-	missense_variant	559G>A	Asp187Asn	MPXV-UK_P2-031	MPXVgp035	C9L similar to Vaccinia virus strain Copenhagen F3L kelch-like
34481	C	T	T	T	snp	CDS	-	synonymous_variant	219G>A	Val73Val	MPXV-UK_P2-037	MPXVgp041	C15L similar to Vaccinia virus strain Copenhagen F9L
34587	C	T	T	T	snp	CDS	-	missense_variant	113G>A	Gly38Glu	MPXV-UK_P2-037	MPXVgp041	C15L similar to Vaccinia virus strain Copenhagen F9L
36907	G	.	A	.	snp	CDS	-	synonymous_variant	186C>T	Phe62Phe	MPXV-UK_P2-039	MPXVgp043	C17L similar to Vaccinia virus strain Copenhagen F11L
39128	T	C	C	C	snp	CDS	-	synonymous_variant	1077A>G	Glu359Glu	MPXV-UK_P2-041	MPXVgp045	C19L similar to Vaccinia virus strain Copenhagen F13L major envelope antigen of EEV wrapping of IMV to form IEV phospholipase D-like
39182	G	A	.	.	snp	CDS	-	synonymous_variant	1023C>T	Val341Val	MPXV-UK_P2-041	MPXVgp045	C19L similar to Vaccinia virus strain Copenhagen F13L major envelope antigen of EEV wrapping of IMV to form IEV phospholipase D-like
42110	C	.	.	T	snp	CDS	+	missense_variant	155C>T	Ser52Leu	MPXV-UK_P2-046	MPXVgp049	C23R similar to Vaccinia virus strain Copenhagen F17R virion core DNA-binding
42192	G	.	A	.	snp	CDS	+	missense_variant	237G>A	Met79Ile	MPXV-UK_P2-046	MPXVgp049	C23R similar to Vaccinia virus strain Copenhagen F17R virion core DNA-binding
42912	C	.	T	.	snp	CDS	-	missense_variant	786G>A	Met262Ile	MPXV-UK_P2-047	MPXVgp050	F1L similar to Vaccinia virus strain Copenhagen E1L poly-A polymerase catalytic
50355	G	.	A	.	snp	CDS	+	missense_variant	361G>A	Asp121Asn	MPXV-UK_P2-053	MPXVgp055	F6R similar to Vaccinia virus strain Copenhagen E7R soluble myristylated protein
54277	C	.	T	.	snp	CDS	-	synonymous_variant	171G>A	Ala57Ala	MPXV-UK_P2-055	MPXVgp057	F8L similar to Vaccinia virus strain Copenhagen E9L DNA polymerase catalytic subunit
57293	C	T	.	.	snp	CDS	-	missense_variant	215G>A	Arg72Lys	MPXV-UK_P2-059	MPXVgp061	Q2L similar to Vaccinia virus strain Copenhagen O2L virion-associated glutaredoxin
61853	G	A	.	.	snp	CDS	-	synonymous_variant	175C>T	Leu59Leu	MPXV-UK_P2-063	MPXVgp065	I4L similar to Vaccinia virus strain Copenhagen I4L ribonucleotide reductase large
64225	G	.	.	A	snp	CDS	-	synonymous_variant	501C>T	Ala167Ala	MPXV-UK_P2-066	MPXVgp068	I7L similar to Vaccinia virus strain Copenhagen I7L virion core protein similar to DNA
72371	T	C	C	C	snp	CDS	-	missense_variant	586A>G	Asn196Asp	MPXV-UK_P2-075	MPXVgp077	G8L similar to Vaccinia virus strain Copenhagen G7L virion protein
73625	C	.	T	.	snp	CDS	+	synonymous_variant	639C>T	Leu213Leu	MPXV-UK_P2-076	MPXVgp078	G9R similar to Vaccinia virus strain Copenhagen G8R late gene transcription factor VLTF-
74235	C	T	.	.	snp	CDS	+	synonymous_variant	447C>T	Ile149Ile	MPXV-UK_P2-077	MPXVgp079	G10R similar to Vaccinia virus strain Copenhagen G9R myristylated protein
75623	G	.	.	A	snp	CDS	+	missense_variant	28G>A	Asp10Asn	MPXV-UK_P2-079	MPXVgp081	M2R similar to Vaccinia virus strain Copenhagen L2R
79415	G	.	A	.	snp	CDS	+	missense_variant	322G>A	Asp108Asn	MPXV-UK_P2-085	MPXVgp087	L3R similar to Vaccinia virus strain Copenhagen J3R poly-A polymerase stimulatory subunit simultaneously cap-specific mRNA (nucleoside-O2'-)-methyltransferase
80108	G	A	A	A	snp	CDS	+	synonymous_variant	99G>A	Leu33Leu	MPXV-UK_P2-086	MPXVgp088	L4R similar to Vaccinia virus strain Copenhagen J4R RNA polymerase 22 kDa subunit
83335	T	C	C	C	snp	CDS	+	missense_variant	2201T>C	Leu734Ser	MPXV-UK_P2-088	MPXVgp090	L6R similar to Vaccinia virus strain Copenhagen J6R RNA polymerase 147 kDa subunit
87239	A	G	G	G	snp	CDS	-	missense_variant	2218T>C	Tyr740His	MPXV-UK_P2-092	MPXVgp094	H4L similar to Vaccinia virus strain Copenhagen H4L virion core RNA polymerase-
87306	A	G	G	G	snp	CDS	-	synonymous_variant	2151T>C	Phe717Phe	MPXV-UK_P2-092	MPXVgp094	H4L similar to Vaccinia virus strain Copenhagen H4L virion core RNA polymerase-
91737	A	G	G	G	snp			intergenic_variant					
93396	G	A	A	A	snp	CDS	+	missense_variant	1657G>A	Glu553Lys	MPXV-UK_P2-096	MPXVgp098	E1R similar to Vaccinia virus strain Copenhagen D1R mRNA capping enzyme large subunit RNA 5' triphosphatase and RNA guanylyl transferase activities
95945	T	.	A	.	snp	CDS	+	missense_variant	576T>A	Asp192Glu	MPXV-UK_P2-099	MPXVgp101	E4R similar to Vaccinia virus strain Copenhagen D4R uracil DNA glycosylase
100649	C	T	.	.	snp	CDS	+	synonymous_variant	255C>T	Val85Val	MPXV-UK_P2-102	MPXVgp104	E7R similar to Vaccinia virus strain Copenhagen D7R RNA polymerase 18 kDa subunit
105755	G	.	A	.	snp	CDS	-	synonymous_variant	223C>T	Leu75Leu	MPXV-UK_P2-107	MPXVgp109	E12L similar to Vaccinia virus strain Copenhagen D12L mRNA capping enzyme small subunit mRNA (guanine-N7-)-methyl-transferase transcription initiation factor
112553	C	.	T	.	snp	CDS	-	missense_variant	997G>A	Asp333Asn	MPXV-UK_P2-115	MPXVgp117	A7L similar to Vaccinia virus strain Copenhagen A6L
114668	G	A	A	A	snp	CDS	-	synonymous_variant	1038C>T	Ala346Ala	MPXV-UK_P2-116	MPXVgp118	A8L similar to Vaccinia virus strain Copenhagen A7L early transcription factor VETF large subunit needed for morphogenesis of the virion core
117128	C	.	.	T	snp	CDS	-	synonymous_variant	2469G>A	Ala823Ala	MPXV-UK_P2-119	MPXVgp121	A11L similar to Vaccinia virus strain Copenhagen A10L major virion core protein p4a
118748	G	A	.	.	snp	CDS	-	synonymous_variant	849C>T	Phe283Phe	MPXV-UK_P2-119	MPXVgp121	A11L similar to Vaccinia virus strain Copenhagen A10L major virion core protein p4a
119305	T	C	C	C	snp	CDS	-	missense_variant	292A>G	Asn98Asp	MPXV-UK_P2-119	MPXVgp121	A11L similar to Vaccinia virus strain Copenhagen A10L major virion core protein p4a
121329	T	C	C	C	snp	CDS	-	missense_variant	49A>G	Thr17Ala	MPXV-UK_P2-122	MPXVgp124	A14L similar to Vaccinia virus strain Copenhagen A13L IMV inner and outer membrane
124040	G	A	A	A	snp	CDS	+	missense_variant	85G>A	Asp29Asn	MPXV-UK_P2-128	MPXVgp129	A19R similar to Vaccinia virus strain Copenhagen A1 8R virion core associated DNA helicase post-replicative negative transcription elongation factor
125258	A	G	G	G	snp	CDS	+	missense_variant	1303A>G	Lys435Glu	MPXV-UK_P2-128	MPXVgp129	A19R similar to Vaccinia virus strain Copenhagen A1 8R virion core associated DNA helicase post-replicative negative transcription elongation factor
128543	C	.	.	T	snp	CDS	+	synonymous_variant	756C>T	Ile252Ile	MPXV-UK_P2-133	MPXVgp134	A24R similar to Vaccinia virus strain Copenhagen A23R intermediate transcription
130463	G	.	.	A	snp	CDS	+	missense_variant	1531G>A	Glu511Lys	MPXV-UK_P2-134	MPXVgp135	A25R similar to Vaccinia virus strain Copenhagen A24R RNA polymerase 132 kDa
134453	C	.	.	T	snp			intergenic_variant					
136523	complex	complex	complex	complex	indel	CDS	-	conservative_inframe	1114_1122delGA	Asp372_Asp374	MPXV-UK_P2-137	MPXVgp138	A28L major component of IMV surface tubules p4c
140503	A	.	C	.	snp	CDS	-	missense_variant	583T>G	Ser195Ala	MPXV-UK_P2-144	MPXVgp144	A34L similar to Vaccinia virus strain Copenhagen A32L DNA packaging into virion NTP-
142509	C	.	T	.	snp	CDS	+	synonymous_variant	297C>T	Ile99Ile	MPXV-UK_P2-147	MPXVgp147	A37R similar to Vaccinia virus strain Copenhagen A35R
142647	G	A	.	.	snp	CDS	+	synonymous_variant	435G>A	Leu145Leu	MPXV-UK_P2-147	MPXVgp147	A37R similar to Vaccinia virus strain Copenhagen A35R
142929	C	.	.	T	snp	CDS	+	synonymous_variant	141C>T	Ile47Ile	MPXV-UK_P2-148	MPXVgp148	A38R similar to Vaccinia virus strain Copenhagen A36R IEV but not CEV envelope protein plays critical role for actin tail formation interacts with VAC A33R and A34R
144457	G	.	.	A	snp			intergenic_variant					
146234	C	T	.	.	snp	CDS	-	missense_variant	556G>A	Asp186Asn	MPXV-UK_P2-151	MPXVgp151	A41L similar to Vaccinia virus strain Copenhagen A41L secreted protein reducing infiltration of inflammatory cells into the infected area
148423	A	G	G	G	snp	CDS	-	synonymous_variant	984T>C	Ile328Ile	MPXV-UK_P2-155	MPXVgp158	A45L similar to Vaccinia virus strain Copenhagen A44L 3-b-Hydroxy-delta5-steroid
148438	G	A	.	.	snp	CDS	-	synonymous_variant	969C>T	Phe323Phe	MPXV-UK_P2-155	MPXVgp158	A45L similar to Vaccinia virus strain Copenhagen A44L 3-b-Hydroxy-delta5-steroid
148539	TCATATCA	T	T	T	indel	CDS	-	frameshift_variant	861_867delTGA	Asn287fs	MPXV-UK_P2-155	MPXVgp158	A45L similar to Vaccinia virus strain Copenhagen A44L 3-b-Hydroxy-delta5-steroid
149276	G	A	A	A	snp	CDS	-	missense_variant	131C>T	Ser44Leu	MPXV-UK_P2-155	MPXVgp158	A45L similar to Vaccinia virus strain Copenhagen A44L 3-b-Hydroxy-delta5-steroid
149883	C	T	T	T	snp	CDS	+	stop_gained	64C>T	Gln22*	MPXV-UK_P2-157	MPXVgp157	A47R similar to Vaccinia virus strain Copenhagen A46R
150811	G	A	.	.	snp			intergenic_variant					
154026	G	A	A	A	snp	CDS	+	missense_variant	1324G>A	Asp442Asn	MPXV-UK_P2-159	MPXVgp160	A50R similar to Vaccinia virus strain Copenhagen A50R DNA ligase
156440	C	.	.	T	snp			intergenic_variant					
157748	C	.	T	.	snp			intergenic_variant					
161163	C	.	T	.	snp	CDS	+	synonymous_variant	531C>T	Asn177Asn	MPXV-UK_P2-163	MPXVgp164	B3R similar to Vaccinia virus strain Copenhagen B1R serine
161964	C	T	T	T	snp	CDS	+	missense_variant	358C>T	His120Tyr	MPXV-UK_P2-164	MPXVgp165	B4R schlafen-like
162254	A	G	G	G	snp	CDS	+	synonymous_variant	648A>G	Leu216Leu	MPXV-UK_P2-164	MPXVgp165	B4R schlafen-like
163551	G	.	A	.	snp	CDS	+	missense_variant	208G>A	Glu70Lys	MPXV-UK_P2-165	MPXVgp166	B5R similar to Vaccinia virus strain Copenhagen B4R ankyrin-like
164843	T	C	C	.	snp	CDS	+	synonymous_variant	1500T>C	Leu500Leu	MPXV-UK_P2-165	MPXVgp166	B5R similar to Vaccinia virus strain Copenhagen B4R ankyrin-like
165793	C	T	T	T	snp	CDS	+	missense_variant	661C>T	Pro221Ser	MPXV-UK_P2-166	MPXVgp167	B6R similar to Vaccinia virus strain Copenhagen B5R palmitated 42 kDa glycoprotein located both on the membranes of infected cells and on EEV envelope complement
168131	T	C	C	C	snp	CDS	+	missense_variant	787T>C	Phe263Leu	MPXV-UK_P2-169	MPXVgp170	B9R similar to Vaccinia virus strain Copenhagen B8R secreted IFN-g binding protein
168216	G	A	A	A	snp			intergenic_variant					
168376	G	A	.	.	snp	CDS	+	missense_variant	136G>A	Glu46Lys	MPXV-UK_P2-170	MPXVgp171	B10R shope fibroma virus T4 protein-like
168543	C	.	.	T	snp	CDS	+	synonymous_variant	303C>T	Phe101Phe	MPXV-UK_P2-170	MPXVgp171	B10R shope fibroma virus T4 protein-like
169732	complex	complex	complex	.	indel	CDS	+	conservative_inframe	94_105delACAG	Thr32_Asp35de	MPXV-UK_P2-171	Hypothetical pr	Hypothetical protein (Cop-B11R)
170328	G	A	.	.	snp	CDS	+	missense_variant	280G>A	Asp94Asn	MPXV-UK_P2-172	MPXVgp172	B11R similar to Vaccinia virus strain Copenhagen B12R protein kinase-like
172538	G	A	.	.	snp	CDS	+	synonymous_variant	378G>A	Leu126Leu	MPXV-UK_P2-174	MPXVgp174	B13R similar to Vaccinia virus strain Copenhagen B15R
174605	complex	complex	complex	complex	indel	CDS	+	disruptive_inframe_d	11_16delAGATC	Lys4_Met5del	MPXV-UK_P2-177	MPXVgp177	B16R similar to Vaccinia virus strain Copenhagen B19R cell surface antigen and
174697	G	.	.	A	snp	CDS	+	missense_variant	79G>A	Glu27Lys	MPXV-UK_P2-177	MPXVgp177	B16R similar to Vaccinia virus strain Copenhagen B19R cell surface antigen and
175846	C	T	.	.	snp	CDS	+	missense_variant	104C>T	Ala35Val	MPXV-UK_P2-178	MPXVgp178	B17R ankyrin-like
177807	C	T	T	T	snp	CDS	+	missense_variant	2065C>T	Arg689Cys	MPXV-UK_P2-178	MPXVgp178	B17R ankyrin-like
179163	complex	complex	complex	complex	indel	CDS	+	intergenic_variant					
179552	C	.	T	.	snp	CDS	+	missense_variant	272C>T	Ser91Leu	MPXV-UK_P2-180	MPXVgp180	R1R
179643	A	.	.	G	snp	CDS	+	synonymous_variant	363A>G	Thr121Thr	MPXV-UK_P2-180	MPXVgp180	SPI-1 B19R serine protease inhibitor-like SPI-1 apoptosis inhibition
179868	C	.	.	T	snp	CDS	+	synonymous_variant	588C>T	Phe196Phe	MPXV-UK_P2-180	MPXVgp180	SPI-1 B19R serine protease inhibitor-like SPI-1 apoptosis inhibition
186119	C	T	.	.	snp	CDS	+	synonymous_variant	4749C>T	Phe1583Phe	MPXV-UK_P2-182	MPXVgp182	B21R putative membrane-associated glycoprotein cadherin-like domain
187443	T	C	C	C	snp			intergenic_variant					
188559	G	.	.	A	snp			intergenic_variant					
189432	G	A	.	.	snp	CDS	+	synonymous_variant	219G>A	Ser73Ser	MPXV-UK_P2-184	MPXVgp185	N1R similar to Vaccinia virus strain Copenhagen B22R VAC B15R-like
189489	C	T	.	.	snp	CDS	+	synonymous_variant	276C>T	Ile92Ile	MPXV-UK_P2-184	MPXVgp185	N1R similar to Vaccinia virus strain Copenhagen B22R VAC B15R-like
194473	C	T	.	.	snp			intergenic_variant					

Table S3. Variant table showing differences in genomes of U.S. variant A.2 USA_2022_FL001 (ON674051.1), USA_2022_VA001 (ON675438.1) and USA_2021_TX (ON676707.1) compared to MT903344.1 UK-P2. Changes shared among all three A.2 sequences are highlighted in green. Position and annotation information is based on reference MT903344.1.

Sample	Year	Month	Country	State	Specimen	Travel	WA_MPXV	PCR	Clade	Variant
MD	2021	November	USA	MD	swab	Nigeria	19.2	positive	IIb	A.1.1
TX	2021	July	USA	TX	swab	Nigeria	18.6	positive	IIb	A.2
MA001	2022	May	USA	MA	swab	Canada	22.7	positive	IIb	B.1
UT001	2022	May	USA	UT	swab	Europe	24.1	positive	IIb	B.1
UT002	2022	May	USA	UT	swab	Europe	24.9	positive	IIb	B.1
FL001	2022	May	USA	FL	swab	Middle East	21.5	positive	IIb	A.2
FL002	2022	May	USA	FL	swab	Domestic USA	23.4	positive	IIb	B.1
VA001	2022	May	USA	VA	swab	Western Africa	18.8	positive	IIb	A.2
CA001	2022	May	USA	CA	swab	Europe	21.7	positive	IIb	B.1

	Oxford Nanopore					Illumina				
Sample	platform	reads	mapped	perc_map	avg_cov	platform	reads	mapped	perc_map	avg_cov
MD	MinION	9,555,644	20,924	0.22%	91	MiSeq	4075122	2009295	49.31	1705
TX	MinION	9,806,213	57,432	0.59%	64	MiSeq	3,392,364	181,357	5.35	59
MA001	MinION	1,843,299	10,548	0.57%	35	MiSeq	59,677,436	17501	0.03	1.5
UT001	GridION	2,979,999	8,855	0.30%	42	MiSeq	18896008	87807	0.46	92
UT002	GridION	1,132,000	4,728	0.42%	23	MiSeq	12570222	63578	0.51	61
FL001	GridION	2,555,998	24,451	0.96%	150	MiSeq	18434660	253654	1.38	258
FL002	GridION	1,456,000	17,786	1.22%	61	MiSeq	11192710	90699	0.81	97
VA001	GridION	5,119,998	11,446	0.22%	28	MiSeq	17421364	39206	0.23	37
CA001	GridION	3,020,000	34,874	1.15%	58	MiSeq	11538620	25569	0.22	21

Table S4. Details for U.S. MPXV genomes from nine 2021 – 2022 cases.

Sample	Clade	Variant	OPX3	MPXV-CladeII	diff
ON563414_MPXV_USA_2022_MA001	Clade IIb	B.1	31.0	22.7	8.3
ON676704_MPXV_USA_2022_FL002	Clade IIb	B.1	30.6	23.4	7.2
ON676705_MPXV_USA_2022_UT001	Clade IIb	B.1	29.4	24.1	5.3
ON676706_MPXV_USA_2022_UT002	Clade IIb	B.1	32.1	24.6	7.5
ON676703_MPXV_USA_2022_CA001	Clade IIb	B.1	29.1	21.7	7.4
ON676708_MPXV_USA_2021_MD	Clade IIb	A.1.1	24.8	19.2	5.6
ON674051_MPXV_USA_2022_FL001	Clade IIb	A.2	21.7	21.2	0.5
ON675438_MPXV_USA_2022_VA001	Clade IIb	A.2	17.8	18.8	-1.0
ON676707_MPXV_USA_2021_TX	Clade IIb	A.2	17.2	18.6	-1.4
MT903337_MPXV_Nigeria_2017_FCT	Clade IIb	A	18.0	19.2	-1.2
Positive control			23.4	23.2	0.2

Table S5. Summary PCR results for OPX3 (*15*) and Clade II-specific (*14*) real-time PCR assay. Average Ct value is based on triplicate testing for each assay. Difference (diff) was calculated by subtracting Clade II-specific average Ct from OPX3 average Ct value.

Data subsets description and sequence names	All other mutations	G-to-A in an APOBEC context	APOBEC context G, no change	G-to-A non-APOBEC context	Non-APOBEC context G, no change	p-value	q-value	Odds Ratio	Clade or Lineage
All mutations in Lineage A, from Node 48									
All unique SNPs in lineage A merged	14	167	36415	9	30669	<0.000001	0.000006	15.63	Lineage A
ON627808.1.MPX.USA.UT-UPHL-82200022.197114.2022-05-20	2	58	36524	2	30676	<0.000001	0.000006	24.36	A B.1
ON568298.1.MPX-BY-IMB25241.197378.Germany.2022-05-19	2	59	36523	2	30676	<0.000001	0.000006	24.78	A B.1
ON585033.1.MPX.PT0006.197220.Portugal.2022-05-15	2	59	36523	2	30676	<0.000001	0.000006	24.78	A B.1
ON585035.1.MPX.PT0009.197220.Portugal.2022-05-15	2	58	36524	2	30676	<0.000001	0.000006	24.36	A B.1
ON563414.MPXV_USA_2022_MA001	2	58	36524	2	30676	<0.000001	0.000006	24.36	A B.1
ON676703.MPXV_USA_2022_CA001	2	58	36524	2	30676	<0.000001	0.000006	24.36	A B.1
ON676705.MPXV_USA_2022_UT001	2	59	36523	2	30676	<0.000001	0.000006	24.78	A B.1
ON676704.MPXV_USA_2022_FL002	2	60	36522	2	30676	<0.000001	0.000006	25.2	A B.1
ON676706.MPXV_USA_2022_UT002	2	58	36524	2	30676	<0.000001	0.000006	24.36	A B.1
ON622712.1.MPX.Belgium.UZ_REGA_1.198010.2022-05-19	2	59	36523	2	30676	<0.000001	0.000006	24.78	A B.1
ON622713.1.MPX.UZ_Belgium.REGA_2.198016.2022-05-22	2	60	36522	2	30676	<0.000001	0.000006	25.2	A B.1
ON614676.1.MPX.INMI-Pt1.190289.Italy..2022-05-18	2	53	36529	2	30676	<0.000001	0.000006	22.25	A B.1
ON676708.MPXV_USA_2021_MD	2	49	36533	3	30675	<0.000001	0.000006	13.71	A A.1.1
MT903344.1_MPXV_UK_P2	0	15	36567	1	30677	0.001513	0.004144	12.58	A A.1
MT903345.1_MPXV_UK_P3	0	15	36567	1	30677	0.001513	0.004144	12.58	A A.1
MT903341.1_MPXV_M5320_M15_Bayelsa	0	13	36569	1	30677	0.005002	0.011892	10.91	A A.1
MT903343.1_MPXV_UK_P1	0	14	36568	2	30676	0.009786	0.022019	5.872	A A.1
MN648051.1_MPXV_Israel_2018	1	16	36566	1	30677	0.000829	0.002678	13.42	A A.1
MT903342.1_MPXV_Singapore	0	16	36566	1	30677	0.000829	0.002678	13.42	A A.1
MT250197.1_MPXV_Singapore_2019	2	16	36566	1	30677	0.000829	0.002678	13.42	A A.1
ON674051_MPXV_USA_2022_FL001	2	28	36554	2	30676	0.000008	0.000042	11.75	A A.2
ON676707_MPXV_USA_2021_TX	0	32	36550	2	30676	0.000001	0.000006	13.43	A A.2
ON675438_MPXV_USA_2022_VA001	1	34	36548	2	30676	<0.000001	0.000006	14.27	A A.2
MT903338.1_MPXV_M2957_Lagos	0	4	36578	0	30678	0.130775	0.189398	inf	A
MT903340.1_MPXV_M5312_HM12_Rivers	0	5	36577	0	30678	0.067325	0.114635	inf	A
MG693724.1_MPXV_Nigeria_2017	3	5	36577	0	30678	0.067325	0.114635	inf	A
MT903339.1_MPXV_M3021_Delta	0	4	36578	0	30678	0.130775	0.189398	inf	A
MT903337.1_MPXV_M2940_FCT	1	3	36579	1	30677	0.630746	0.749755	2.516	A
MG693723.1_MPXV_Nig_2017_297957	1	11	36571	0	30678	0.001409	0.004144	inf	A
Internal Branches within Lineage A									
Node 52-56	1	35	36535	1	30676	<0.000001	0.000006	29.39	A
Node 56-57	1	11	36524	0	30675	0.001403	0.004144	inf	A
Node 48-52	0	12	36570	1	30677	0.004944	0.011892	10.07	A
Node 48-71	0	14	36568	1	30677	0.002755	0.007232	11.74	A
All mutations in Variant B.1, from Node 57									
All unique SNPs in variant B.1 merged	0	8	36517	0	30675	0.009498	0.021759	inf	Lineage A B.1
ON627808.1.MPX.USA.UT-UPHL-82200022.197114.2022-05-20	0	0	36525	0	30675	1	1.000000		A B.1
ON568298.1.MPX-BY-IMB25241.197378.Germany.2022-05-19	0	1	36524	0	30675	1	1.000000	inf	A B.1
ON585033.1.MPX.PT0006.197220.Portugal.2022-05-15	0	1	36524	0	30675	1	1.000000	inf	A B.1
ON585035.1.MPX.PT0009.197220.Portugal.2022-05-15	0	0	36525	0	30675	1	1.000000		A B.1
ON563414.MPXV_USA_2022_MA001	0	0	36525	0	30675	1	1.000000		A B.1
ON676703.MPXV_USA_2022_CA001	0	0	36525	0	30675	1	1.000000		A B.1
ON676705.MPXV_USA_2022_UT001	0	1	36524	0	30675	1	1.000000	inf	A B.1
ON676704.MPXV_USA_2022_FL002	0	2	36523	0	30675	0.503782	0.616338	inf	A B.1
ON676706.MPXV_USA_2022_UT002	0	0	36525	0	30675	1	1.000000		A B.1
ON622712.1.MPX.Belgium.UZ_REGA_1.198010.2022-05-19	0	1	36524	0	30675	1	1.000000	inf	A B.1
ON622713.1.MPX.UZ_Belgium.REGA_2.198016.2022-05-22	0	2	36523	0	30675	0.503782	0.616338	inf	A B.1
ON614676.1.MPX.INMI-Pt1.190289.Italy..2022-05-18	0	0	36525	0	30675	1	1.000000		A B.1
All unique mutations in extended variant B1 data, including 397 international GISAID sequences sampled between May 1 - July 15, 2022, details in Sup. Fig. S2 and S5.									
All unique SNPs in extended lineage B.1 merged	115	275	35064	33	29703	<0.000001	0.000006	7.06	A B.1
All mutations in variant A.2, from Node 71									
All unique SNPs in variant A.2 merged	3	52	36516	3	30674	<0.000001	0.000006	14.56	Lineage A A.2
ON674051_MPXV_USA_2022_FL001	2	14	36554	1	30676	0.002755	0.007232	11.75	A A.2
ON676707_MPXV_USA_2021_TX	0	18	36550	1	30676	0.000248	0.000977	15.11	A A.2
ON675438_MPXV_USA_2022_VA001	1	20	36548	1	30676	0.000074	0.000322	16.79	A A.2
All mutations in the 2018/2019 imports set, from Node 52									
All unique SNPs in the 2018/2019 import set merged	3	16	36554	1	30676	0.000829	0.002678	13.43	Lineage A A.1
MT903344.1_MPXV_UK_P2	0	3	36567	0	30677	0.255748	0.354113	inf	A A.1.1
MT903345.1_MPXV_UK_P3	0	3	36567	0	30677	0.255748	0.354113	inf	A A.1.1
MT903341.1_MPXV_M5320_M15_Bayelsa	0	2	36568	0	30677	0.503832	0.616338	inf	A A.1.1
MT903343.1_MPXV_UK_P1	0	2	36568	1	30676	1	1.000000	1.678	A A.1.1
MN648051.1_MPXV_Israel_2018	1	4	36566	0	30677	0.130756	0.189398	inf	A A.1.1
MT903342.1_MPXV_Singapore	0	4	36566	0	30677	0.130756	0.189398	inf	A A.1.1
MT250197.1_MPXV_Singapore_2019	2	4	36566	0	30677	0.130756	0.189398	inf	A A.1.1
All mutations in the Lineage A Nigerian set, from Node 48									
All unique SNPs in the Lineage A Nigerian set merged	5	24	36558	1	30677	0.000007	0.000038	20.14	Lineage A
MT903338.1_MPXV_M2957_Lagos	0	4	36578	0	30678	0.130775	0.189398	inf	A
MT903340.1_MPXV_M5312_HM12_Rivers	0	5	36577	0	30678	0.067325	0.114635	inf	A
MG693724.1_MPXV_Nigeria_2017	3	5	36577	0	30678	0.067325	0.114635	inf	A
MT903339.1_MPXV_M3021_Delta	0	4	36578	0	30678	0.130775	0.189398	inf	A
MT903337.1_MPXV_M2940_FCT	1	3	36579	1	30677	0.630746	0.749755	2.516	A
MG693723.1_MPXV_Nig_2017_297957	1	11	36571	0	30678	0.001409	0.004144	inf	A

Table S6. Details and statistical analysis of mutational patterns within Lineage A, based on the trees shown in Fig. 2 and Fig. S2. Here we compare the frequencies of G-to-A substitutions in APOBEC3 and non-APOBEC3 contexts in Lineage A and each relevant sub-lineage within Lineage A. This table details the mutational patterns throughout the entire Lineage A, and G-to-A mutations in an APOBEC3 context are highly enriched relative to other mutations. First, to determine the overall level of enrichment for these mutations in Lineage A, we compressed every unique SNP mutation that arose within Lineage A onto a single sequence we call "All unique SNPs in Lineage A merged"; the analysis of the merged data is highlighted in gray at the top of the table, and this summary was also used in Figure 2 in the main text. Within Lineage A, 167 G-to-A mutations arose in an APOBEC3 context (5' GA-to-AA or GG-to-AG, in blue); 36,415 Gs in a APOBEC context, GA or GG in the ancestral sequence, did not change. In contrast, only 9 G-to-A mutations arose out of *outside* of an APOBEC3 context (GY-to-AY, where Y is C or T, in red), while 30,669 GY's remained unchanged. A simple two-sided Fisher's exact text was used to test the null hypothesis that the distributions between the two mutational patterns was random; q-values (*37, 38*) were calculated based on all p-values included Tables S6 and S8 to correct for multiple tests. The G-to-A mutations are obviously highly enriched. When the odds ratio is greater than one (highlighted in light blue) it is indicative more G-to-A mutations in an APOBEC3 context. We also tallied all other SNPs that were not G-to-A (in gray); these were rare, there were only 14 of these in all of Lineage A. To resolve if the enrichment pattern for G-to-A mutations in an APOBEC context was observed throughout Lineage A, we compared the mutational patterns on the internal branches within Lineage A, and we found all partitions of the data throughout Lineage A were enriched for G-to-A in an APOBEC3 context. While variant B.1, the predominant 2022 outbreak, had too few changes between the ancestor and any one leaf for a single sequence to be significant, all of the 8 mutations observed in the 12 earliest sequences from variant B.1 available for our initial analysis were G-to-A in an APOBEC3 context, and this enrichment was significant when considered together as a merged sequence. There were enough mutations in the longer branches in variant A.2 for statistical significance of the pattern to be evident for each of the 3 taxa independently. G-to-A in an APOBEC3 context was also enriched among the Nigerian sequences, in 2018/2019 export set, and on each of 4 major internal branches within Lineage A.

Table S7. G-A substitutions are enriched in an APOBEC3 context among 397 variant B.1 sequences available in GISAID through mid-July and among 4 additional variant A.2 sequences*. To confirm the enrichment for G-to-A substitutions was retained in the rapidly expanding available outbreak sequence data, we extended our original analysis to include an additional 397 variant B.1 sequences that were available in GISAID as of July 15, 2022; these data supported our initial findings. A summary of all merged unique mutations in this set of 397 sequences is provided in this table, and the merged data is presented in Fig. 2B in the main text; a detailed graphic displaying the details of the analyses is provided in the supplemental Fig. S3. We also identified four additional variant A.2 sequences, two from Thailand and two from India, that became available through GISAID later in July. These too were highly enriched for G-to-A substitutions, and the statistics are provided in this table. The column headers are described in Table S6. *__EPI_ISL_14011193:__ Bangkok Hospital Phuket, Thai Red Cross Emerging Infectious Diseases Clinical Center and Faculty of Medicine, Chulalongkorn University; **EPI_ISL_13953611**: Indian Council of Medical Research-National Institute of Virology; **EPI_ISL_14049245**, Indian Council of Medical Research-National Institute of Virology; **EPI_ISL_13983888**: Bangkok Hospital Phuket, National Institute of Health, Department of Medical Sciences, Ministry of Public Health, Thailand**.** The GISAID acknowledgments tables for these two data sets are provided in Tables S10 and S11.

Data subset descriptions and sequence names	All other mutations	G-to-A in an APOBEC context	APOBEC context G, no change	G-to-A non-APOBEC context	Non-APOBEC context G, no change	p-value	q-value	Odds Ratio	Clade or Lineage
All mutations in Clade IIa, from Node 76									
Merged-SNPs	225	65	36548	132	30578	<0.000001	0.000006	0.412	Clade IIa
KJ642615.1_MPXV_W-Nigeria_1978	112	17	36596	48	30662	0.000005	0.000029	0.2967	IIa
DQ011156.1_MPXV_Liberia_1970	33	14	36599	21	30689	0.092181	0.150842	0.559	IIa
AY741551.1_MPXV_Sierra_Leone_1970	28	20	36593	24	30686	0.289194	0.391811	0.6988	IIa
AY753185.1_MPXV_COP-58_cynomolgus_monkey_West_African_origin_1958	26	12	36601	15	30695	0.336987	0.442295	0.6709	IIa
AY603973.1_MPXV_MPXV-WRAIR7-61-cynomolgus-monkey-1962	27	10	36603	17	30693	0.082367	0.136556	0.4933	IIa
KJ642616.1_MPXV_PCH_Paris_1962	39	8	36605	22	30688	0.002826	0.007267	0.3049	IIa
MT903346.1_MPXV_USA_2003_099	40	9	36604	32	30678	0.000028	0.000141	0.2357	IIa
DQ011157.1_MPXV_USA_2003_039	40	10	36603	32	30678	0.000071	0.000320	0.2619	IIa
DQ011153.1_MPXV_USA_2003_044	40	10	36603	32	30678	0.000071	0.000320	0.2619	IIa
Mutations in Clade IIb, prior to the emergence of Lineage A									
Node 60 to KJ642615.1_MPXV_W-Nigeria_1978	29	4	36604	19	30693	0.00047	0.001745	0.1765	Clade IIb
Node 46-47	60	25	36583	52	30660	0.00013	0.000546	0.4029	IIb
Node 47-48	11	11	36584	8	30675	0.821212	0.967035	1.153	Iib
Mutations in Clade I, from Node 5: Clade I									
Merged-SNPs	277	72	36552	176	30580	<0.000001	0.000006	0.3423	Clade I
JX878428.1_MPXV_DRC_07-0514	1	0	36624	1	30755	0.456456	0.575264	0	I
JX878422.1_MPXV_DRC_07-0287	3	1	36623	1	30755	1	1.000000	0.8398	I
JX878413.1_MPXV_DRC_07-0045	0	0	36624	0	30756	1	1.000000		I
JX878427.1_MPXV_DRC_07-0480	1	0	36624	0	30756	1	1.000000		I
JX878421.1_MPXV_DRC_07-0286	1	0	36624	0	30756	1	1.000000		I
JX878414.1_MPXV_DRC_07-0046	0	0	36624	0	30756	1	1.000000		I
JX878415.1_MPXV_DRC_07-0092	0	0	36624	0	30756	1	1.000000		I
JX878416.1_MPXV_DRC_07-0093	0	1	36623	0	30756	1	1.000000	inf	I
JX878410.1_MPXV_DRC_06-1070	9	0	36624	0	30756	1	1.000000		I
JX878411.1_MPXV_DRC_06-1075	17	4	36620	7	30749	0.245298	0.347276	0.4798	I
JX878409.1_MPXV_DRC_06-0999	17	4	36620	7	30749	0.245298	0.347276	0.4798	I
JX878412.1_MPXV_DRC_06-1076	18	3	36621	9	30747	0.046645	0.087720	0.2799	I
KP849471.1_MPXV_Yambuku_DRC_1985	17	6	36618	8	30748	0.430146	0.553045	0.6298	I
MN702448.1_MPXV_015c	36	7	36617	15	30741	0.051532	0.094102	0.3918	I
MN702447.1_MPXV_18	36	6	36618	15	30741	0.026145	0.051473	0.3358	I
MN702450.1_MPXV_B1	43	7	36617	17	30739	0.022071	0.044854	0.3457	I
MN702453.1_MPXV_A1	34	5	36619	14	30742	0.019292	0.041200	0.2998	I
MN702452.1_MPXV_A2	35	6	36618	14	30742	0.040971	0.078217	0.3598	I
DQ011154.1_MPXV_Congo_2003_358	38	5	36619	16	30740	0.007076	0.016511	0.2623	I
JX878429.1_MPXV_DRC_07-0662	39	7	36617	17	30739	0.022071	0.044854	0.3457	I
JX878425.1_MPXV_DRC_07-0354	39	7	36617	17	30739	0.022071	0.044854	0.3457	I
JX878424.1_MPXV_DRC_07-0338	41	5	36619	15	30741	0.011744	0.025513	0.2798	I
JX878423.1_MPXV_DRC_07-0337	43	6	36618	15	30741	0.026145	0.051473	0.3358	I
JX878407.1_MPXV_DRC_06-0950	42	5	36619	17	30739	0.004225	0.010647	0.2469	I
KP849469.1_MPXV_Boende_DRC_2008	41	7	36617	14	30742	0.077307	0.129876	0.4198	I
MN702451.1_MPXV_A6	37	8	36616	17	30739	0.027567	0.053438	0.3951	I
KJ642613.1_MPXV_Congo_8	33	5	36619	12	30744	0.050168	0.092958	0.3498	I
MN702449.1_MPXV_B2	34	10	36614	13	30743	0.304321	0.403626	0.6459	I
MN702446.1_MPXV_38c	37	11	36613	13	30743	0.41972	0.545203	0.7105	I
MN702445.1_MPXV_A4	36	10	36614	16	30740	0.117126	0.184473	0.5247	I
MN702444.1_MPXV_A5	36	10	36614	16	30740	0.117126	0.184473	0.5247	I
DQ011155.1_MPXV_Zaire_1979-005	37	6	36618	13	30743	0.063366	0.114059	0.3875	I
HQ857562.1_MPXV_V79-I-005	45	10	36614	13	30743	0.304321	0.403626	0.6459	I
HQ857563.1_MPXV_D14L_knockout	51	10	36614	12	30744	0.52176	0.632132	0.6997	I
HM172544.1_MPXV_Zaire_1979-005	54	11	36613	35	30721	0.000039	0.000189	0.2637	I
KC257460.1_MPXV_DRC_Yandongi_1985	39	7	36617	9	30747	0.456559	0.575264	0.6531	I
KJ642619.1_MPXV_Gabon-1988	40	7	36617	13	30743	0.114327	0.184473	0.4521	I
KJ642612.1_MPXV_Ikubi	42	8	36616	12	30744	0.261696	0.358410	0.5598	I
JX878426.1_MPXV_DRC_07-0450	43	4	36620	21	30735	0.000147	0.000597	0.1599	I
JX878408.1_MPXV_DRC_06-0970	40	4	36620	15	30741	0.004551	0.011244	0.2239	I
JX878420.1_MPXV_DRC_07-0283	43	4	36620	17	30739	0.001484	0.004144	0.1975	I
JX878419.1_MPXV_DRC_07-0275	43	4	36620	17	30739	0.001484	0.004144	0.1975	I
JX878418.1_MPXV_DRC_07-0120	45	5	36619	21	30735	0.000501	0.001804	0.1998	I
AF380138.1_MPXV_Zaire-96-I-16	109	16	36608	30	30726	0.011006	0.024329	0.4476	I
JX878417.1_MPXV_DRC_07-0104	48	4	36620	19	30737	0.000471	0.001745	0.1767	I

Table S8. Details and statistical analysis of mutational patterns outside of Lineage A, based on the trees shown in Fig. 2 and Fig. S2. Here we compare Clade I, Clade IIa and Clade IIb data prior to the emergence of Lineage A. In these clades, there was no evident enrichment for G-to-A changes in APOBEC3 context. Instead, there was a modest, but overall significant, enrichment for G-to-A substitutions to occur *not* in a APOBEC3 context (the two-sided p-values were for the merged data for Clades I and IIa were significant) and the Odds Ratios were less than one throughout these clades indicating G-to-A changes in APOBEC3 context were relatively diminished compared to other G-to-A mutations.

Alignment	Full	Clades I and IIa	LineageA + Nigerian-SE-1971	LineageA + Nigerian-SE-1971
taxa	Clade I, Clade IIa, Lineage A	Clade I, Clade IIa	LineageA	LineageA
model	GTR+g+i	GTR+g+i	GTR+g+i	GTR+g+i
clock	fixed local	fixed local	fixed local	Uncorrelated lognormal
tree	Bayesian skyline	Bayesian skyline	Bayesian skyline	Bayesian skyline
LineageA rate	7.1817E-6 (5.5355E-6, 8.9951E-6)	NA	8.7045E-6 (6.4756E-6, 1.108E-5)	2.8395E-5 (1.4592E-5, 4.604E-5)
CladeIIa rate	3.8539E-6 (2.1937E-6, 5.5976E-6)	3.4138E-6 (1.4092E-6, 5.5292E-6)	NA	NA
CladeI rate	1.8768E-6 (2.1937E-6, 5.5976E-6)	1.8471E-6 (1.209E-6, 2.4765E-6)	NA	NA
meanrate	3.2883E-6 (2.339E-6, 4.217E-6)	2.3403E-6 (1.5014E-6, 3.1729E-6)	8.026E-6 (6.6458E-6, 1.1203E-5)	1.4876E-5 (1.2073E-5, 1.7897E-5)

Table S9. Rate estimates for Lineage A, Clade I and Clade IIa. Analysis was performed in BEAST v1.8.3 using the alignment from Figure 2. Mean rates are shown with 95% highest posterior density intervals in parenthesis.

Table S10. GISAID acknowledgement table for variant B.1 sequences

Table S11. GISAID acknowledgement table for variant A.2 sequences