Author Manuscript

HHS Public Access
Author manuscript
Science. Author manuscript; available in PMC 2023 June 12.

Published in final edited form as:
Science. 2022 November 04; 378(6619): 560–565. doi:10.1126/science.add4153.

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

Crystal M. Gigante[1], **Bette Korber**[2], **Matthew H. Seabolt**[1,3], **Kimberly Wilkins**[1], **Whitni Davidson**[1], **Agam K. Rao**[1], **Hui Zhao**[1], **Todd G. Smith**[1], **Christine M. Hughes**[1], **Faisal Minhaj**[1], **Michelle A. Waltenburg**[1], **James Theiler**[4], **Sandra Smole**[5], **Glen R. Gallagher**[5], **David Blythe**[6], **Robert Myers**[6], **Joann Schulte**[7], **Joey Stringer**[7], **Philip Lee**[8], **Rafael M. Mendoza**[9], **LaToya A. Griffin-Thomas**[10], **Jenny Crain**[11], **Jade Murray**[12], **Annette Atkinson**[12], **Anthony H. Gonzalez**[13], **June Nash**[13], **Dhwani Batra**[1], **Inger Damon**[1], **Jennifer McQuiston**[1], **Christina L. Hutson**[1], **Andrea M. McCollum**[1], **Yu Li**[1,*]

[1]National Center for Emerging and Zoonotic Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia, USA

[2]T-6: Theoretical Biology and Biophysics, Los Alamos National Laboratory, Los Alamos, New Mexico, USA; New Mexico Consortium, Los Alamos, New Mexico, USA.

[3]Leidos Inc., Reston, VA 20190, USA

[4]ISR-3: Space Data Science and Systems, Los Alamos National Laboratory, Los Alamos, New Mexico, USA

[5]Massachusetts Department of Public Health, Jamaica Plain, Massachusetts, USA.

[6]Infectious Disease Epidemiology and Outbreak Response Bureau, Maryland Department of Health, Baltimore, Maryland, USA

[7]Dallas County Health and Human Services Public Health Laboratory, Dallas, Texas, USA

[8]Florida Department of Health Bureau of Public Health Laboratories-Jacksonville, Jacksonville, Florida, USA

[9]Florida Department of Health in Broward County, Hollywood, Florida, USA

[10]Virginia Department of General Services - Division of Consolidated Laboratory Services, Richmond, Virginia, USA

[11]Virginia Department of Health, Richmond, Virginia, USA

[12]Utah Department of Health and Human Services, Salt Lake City, Utah, USA

---

[*]Corresponding author. yuli@cdc.gov.

Data and materials availability:
All data are available in the main text or public databases and GISAID. MPXV genomes were deposited to GenBank under accession numbers ON563414.3, ON674051, ON675438, and ON676703 – ON676708. Alignments and code have been made available (41–43).

Supplementary Materials
Materials and Methods
Tables S1 – S11
Figures S1 – S5
References 34–43

Author Manuscript
Author Manuscript
Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

[13]Sacramento County Public Health, Sacramento, California, USA

Abstract

Monkeypox is a viral zoonotic disease endemic in Central and West Africa. In May 2022, dozens of non-endemic countries reported hundreds of monkeypox cases, most with no epidemiological link to Africa. We identified two lineages of monkeypox virus (MPXV) among two 2021 and seven 2022 U.S. monkeypox cases: the major 2022 outbreak variant, B.1, and a minor contemporaneously sampled variant called A.2. Analyses of mutations among these two variants revealed an extreme preference for GA-to-AA mutations indicative of human APOBEC3 cytosine deaminase activity among Clade IIb MPXV (previously West African, Nigeria) sampled since 2017. Such mutations were not enriched within other MPXV clades. These findings suggest that APOBEC3 editing may be a recurrent and a dominant driver of MPXV evolution within the current outbreak.

One-Sentence Summary:

We report multiple introductions of monkeypox viruses of different origins into the U.S. and find evidence of APOBEC3 editing since 2017.

Monkeypox is a viral zoonotic disease caused by monkeypox virus (MPXV) that is endemic in West and Central Africa. There have been several reported cases of travel-associated monkeypox in non-African countries in recent years. In 2003, an outbreak of monkeypox in the United States (U.S.) was linked to imported African small mammals (1). In 2017, the largest monkeypox outbreak in western Africa occurred in Nigeria after decades of no identified cases (2), and during 2018 to 2021, eight cases were exported from Nigeria to non-endemic countries (2–5). In 2021, there were two U.S. monkeypox cases in travelers from Nigeria, in Maryland and Texas (4, 5). In May of 2022, this pattern of monkeypox cases in travelers from Nigeria shifted. As of September 28, 2022, 67,602 cases of monkeypox were reported in 99 non-endemic countries, most with no epidemiological link to Africa (6), and 25,509 cases have been confirmed in the U.S.A (7).

Comparison of MPXV sequences from nine U.S. monkeypox cases from 2021 and 2022 (ON563414.3, ON674051, ON675438, ON676703 – ON676708) revealed two distinct lineages (Fig. 1) within MPXV Clade IIb (formerly named West African MPXV found east of the Dahomey Gap). Neutral nomenclature was chosen to reduce stigmatization that can be associated with naming based on locations (8). We maintain the designation of two clades of MPXV (Clade I: formerly Congo Basin/Central African and Clade II: formerly West African) based on genetic distance and evidence-based clinical differences (9–12). Five of the seven May 2022 U.S. sequences formed a monophyletic clade with 2022 MPXV sequences from Europe (Fig. 1), with most genomes within this clade containing 0 – 2 nucleotide changes in non-repeat regions (Fig. S1, Table S1). This clade will be referred to as the current predominant 2022 MPXV outbreak variant, B.1 (based on proposed MPXV naming from Nextstrain (13)). MPXV from a 2021 travel-associated case from Nigeria to Maryland (USA_2021_MD) displayed high similarity to variant B.1 sequences, with approximately 13 nucleotide differences (Figs. 1 and S1, Table 1). The USA_2021_MD

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

and 2022 outbreak sequences shared many mutations that separated them from MPXV sequences from Nigeria and travel-associated cases from 2017 – 2019 (Table S2).

Two 2022 U.S. MPXV sequences, USA_2022_FL001 and USA_2022_VA001, and one 2021 U.S. sequence (USA_2021_TX) formed a monophyletic clade (variant A.2) that was polyphyletic to other 2022 MPXV sequences from the U.S. and Europe (Fig. 1). Each genome contained approximately 80 nucleotide changes relative to variant B.1 MPXV sequences (Fig. S1). The three variant A.2 genomes displayed approximately 30 unique nucleotide differences from each other (Fig. S1, Table S3). Each case also reported travel to different countries in the Middle East and West Africa (Table S4). Together, this suggests that although the three MPXVs share a common ancestor, they likely represent separate introductions to the USA.

Real-time PCR testing of USA_2021_MD and the five 2022 MPXV variant B.1 samples revealed a lower sensitivity (average Ct delay of 6.88, ranging from 5.3 to 8.3) in *Orthopoxvirus* generic OPX3 real-time PCR assay (14) compared to Monkeypox Clade II-specific real-time PCR assay (15) (Table S5, when performed as described in Methods). By contrast, similar Ct values were observed for USA_2022_FL001, USA_2022_VA001, or USA_2021_TX samples (average difference of −0.78 Ct, ranging from −1.4 to 0.5). By careful comparison of Ct values between the two assays laboratories can, in theory, differentiate between cases belonging to variant B.1 and cases from other Clade IIb MPXV without sequencing. Sequence examination revealed a SNP in the reverse primer binding site for the OPX3 real-time PCR assay (DNA polymerase gene, VACV-Cop E9L, C322T) in USA_2021_MD and variant B.1 sequences that was absent from other Clade IIb MPXV sequences. The decreased sensitivity in the OPX3 assay is unrelated to the U.S. Food and Drug Administration-cleared VAC1 assay (16) used for MPXV screening. Use of different commercial reagents, run parameters or PCR platform may result in different results.

When we compared the 2022 outbreak MPXV sequences in Fig. 1 to other MPVX sequences in Lineage A, we noticed a striking preponderance of G-to-A mutations, specifically 5′ GA-to-AA or 5′ GG-to-AG, indicative of host Apolipoprotein B mRNA Editing Catalytic Polypeptide-like3 (APOBEC3) activity (17, 18) (Fig. 2, Table 2). Looking at this systematically, we found a significant enrichment of APOBEC3 context G-to-A mutations was evident throughout Lineage A, among Clade IIb MPXV sequences sampled from 2017 to 2022 (Fig. 2). Overall, among unique mutations within Lineage A, 167 G-to-A mutations were in an APOBEC3 context, while nine G-to-A mutations were not in an APOBEC3 motif, and only 14 mutations were not G-to-A (Fig. 2, Table 2). The vast majority of the APOBEC3 context G-to-A mutations were specifically GA-to-AA (156 out of 167, 95%, Fig. 2), indicating that if these mutations were generated by APOBEC3 editing, APOBEC3G was not the major form, as it produces GG-to-AG changes (19, 20).

We explored the abundance of APOBEC3 motif G-to-A mutations along the evolutionary trajectory from the common ancestor of Lineage A to the variant A.2 and B.1 (Figs. 2, S2). The path from the estimated ancestor of Lineage A to the estimated ancestors of both variant B.1 and A.2, including each internal branch, revealed significantly increased APOBEC3 motif G-to-A changes, indicating sequential acquisition of mutations in distinct

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

individuals sampled historically since 2017 have contributed to the high proportions of GA-to-AA APOBEC3 (Fig. S2 and Table S6). Given the limited variation found in 2022 outbreak sequences in variant B.1, there were only a small number of SNPs relative to the ancestor for any one sequence. When considered together, however, all 8 SNPs observed among the 12 variant B.1 genomes available in early May 2022 were GA-to-AA APOBEC3 context mutations (Figs. 2, S2 and Table S6). Outbreak-related sequence data has increased very rapidly, enabling analysis of 397 outbreak genomes sampled between May 1 and July 15, 2022, available through the GISAID data sharing initiative (21). This greatly expanded data set confirmed the pattern observed among the first twelve sequences: 275/308 (89%) of observed unique G-to-A mutations occurred in an APOBEC3 context, and 261/275 (95%) of these were specifically 5' GA-to-AA (Figs. 2 and S3, Table S7). Beyond variant B.1, by partitioning Lineage A into sublineages (A.2, Nigeria 2017, and Nigeria 2017 – 2019; Fig. 2) we found that G-to-A mutations in A.2 were in an APOBEC3 context were enriched throughout Lineage A (Figs. 2, S2 and Table S6). Four additional MPXV sequences belonging to variant A.2 have now been detected in India and Thailand (22); 45/47 (96%) G-to-A mutations were in a APOBEC3 context, and all 45 of these were specifically 5' GA-to-AA (Table S7).

In contrast, APOBEC3 context G-to-A changes were lower than expected by chance relative to other G-to-A changes across Clade I and Clade IIa MPXV (Figs. 2 and S2, Table S8). To better resolve where within the phylogeny the switch in mutational patterns arose, we evaluated the mutational frequencies along the internal branches in Clade IIb leading into Lineage A (Figs. 2 and S2, Table S8); these were not found to be statistically enriched for APOBEC3 context G-to-A changes.

Of note, in the July 15 GISAID sample there were three highly related sequences, one each from Italy, Spain and France, that were chimeric in that most of the genome was typical of variant B.1, but each ITR region carried 5 consecutive bases that were typically found among other Lineage A and Clade IIb MPXV but not among B.1 sequences. A detailed analyses of these sequences is provided in Fig. S4, and they were excluded from analysis shown in Fig S3. This pattern is potentially indicative of recombination, an important aspect of poxvirus evolution (23); however, assembly errors caused by reference-calling in low coverage regions could not be ruled out as a possible explanation.

Given the clear transition in the mutational pattern in Lineage A relative to Clades I and IIa (Fig. 2), we explored the hypothesis that there was a concomitant change in the evolutionary rate in Lineage A. First, we used the fixed local clock model as implemented in BEAST to compare the estimated mean evolutionary rate of Lineage A ($7.2\times10^{-6} \pm 8.9\times10^{-7}$ (standard deviation), to Clade I ($1.9\times10^{-6} \pm 3.1\times10^{-7}$) and Clade IIa ($3.9\times10^{-6} \pm 8.9\times10^{-7}$) (Table S9). We also separated Lineage A, retaining Nigerian-SE-1971 as an outgroup, from Clade I and IIa and repeated the analysis using two trees; we felt this to be the more appropriate alternative given the transition in the underlying evolutionary model. In this analysis we found Lineage A had an even higher evolutionary rate estimate of $2.8\times10^{-5} \pm 8.8\times10^{-6}$, while estimates for Clades I and IIa were essentially unchanged (Table S9). Thus, using either approach, we observed higher evolutionary rates within Lineage A, consistent with mutations being driven by host editing mediated through APOBEC3, over

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

what is expected by errors in the viral polymerase. An enriched evolutionary rate for clade B.1 outbreak viruses has been similarly reported by others (18, 22). A cautionary note regarding the robustness of estimates of rates of evolutionary change in this scenario is that if the mutations are indeed dominated by host-mediated mutations, different people may have different levels of APOBEC3 activity, and this may change over time. Such transient host-dependent increases in G-to-A ABOBEC3 motif enriched mutations in HIV genomes in single infected hosts are observed both in HIV-1 infection in humans and SHIV infections in monkeys (24, 25)

All publicly available MPXV genomes from the 2022 Monkeypox outbreak to date belong to Clade II, which may cause less severe disease and have a lower case fatality rate than Clade I MPXV (10–13). Genomes published during the 2022 outbreak share a common ancestor with MPXV sequences from Nigeria (Clade IIb); however, sequences from surrounding countries are limited and most of our understanding of these relationships comes from viruses linked to or identified in Nigeria. The high similarity among the current predominant 2022 MPXV outbreak variant B.1 is similar to the 0.4 to 1.5 SNPs reported between genomes from epidemiologically linked samples from the same transmission chain (3). Among Clade IIb sequences, many unique mutations were shared between USA_2021_MD and variant B.1 genomes, further indicating that any evolutionary force leading to these changes most likely preceded the 2022 outbreak and were found in a common ancestor shared between the variant B.1 and USA_2021_MD. In contrast, there were sufficient differences (>30 SNPs) among the three variant A.2 sequences to suggest they likely represent separate introductions into the U.S.

After the re-emergence of MPXV in Nigeria in 2017 and prior to 2022, there had only been two reported events of person-to-person transmission of MPXV outside of Africa (3, 26). Several factors may be contributing to this recent increased spread between humans, including behaviors involving close contact and failure to recognize or diagnose monkeypox to prevent spread. It is currently unclear if APOBEC3 motif or other mutations observed in the 2022 outbreak MPXV variants have affected or will affect viral transmissibility or pathogenicity. A mutation in F13L (which codes for the target of tecovirimat, an antiviral agent indicated for the treatment of smallpox) (E353K) was shared by all sequences in the predominant 2022 outbreak cluster; functional studies revealed no effect on the efficacy of tecovirimat (EC_{50} was 0.007731 ± 0.002 μM for 353E and 0.002860 ± 0.001 μM for 353K) (Fig S5).

APOBEC3 proteins are an important component of the vertebrate innate immune system and restrict the replication of exogenous viruses through cytosine-to-uracil deaminase activity (19, 27). APOBEC3 proteins act mainly on single-stranded DNA and have been extensively studied in RNA viruses, including HIV (20) but can also act on DNA viruses (28, 29). It is unlikely the GA-to-AA substitutions we report here are due to sequencing error as the sequences used in the analyses were produced across many different laboratories using multiple methods, including both short and long read approaches, and both with and without PCR. The presence of mutations consistent with APOBEC3 editing in several branches of the Clade IIb MPXV will cause discordance in estimates of evolutionary rate and divergence time using many standard methods.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Specific enrichment of APOBEC3 motif mutations in Clade IIb MPXV since 2017 may suggest something different in virus-host interactions, facilitating this mutational pattern. Such a pattern might be caused by sustained transmission in a new host or a new route of infection. Primates, including humans, exhibit evolutionary expansion of the APOBEC3 family to seven members that are not seen in rodents (30), which are thought to be the primary reservoir of monkeypox virus in Africa. Additionally, APOBEC3 has been found in higher levels in mucosal tissues than in skin in humans (31), thus permucosal transmission may provide increased opportunity for APOBEC3 editing of MPXV. Prior to this outbreak, APOBEC3 editing was not recognized or appreciated as a mechanism of poxvirus mutation. Two studies in vaccinia virus found (1) overexpression of APOBEC3 had no effect on viral replication and (2) endogenous or overexpressed APOBEC3 was not degraded by vaccinia virus, as it is by other viruses (32, 33). However, neither of these studies performed sequencing. The fact that the enrichment can be observed at essentially all levels within Clade IIb MPXV since 2017 suggests it is a recurrent and dominant mutational effect in recent MPXV evolution. Furthermore, the majority of mutations detected among expanding outbreak samples were 5′ GA-to-AA (Fig. 2, Table S7), indicating that this mutational bias is continuing.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

We thank Brian Foley for providing a draft alignment of MPXV. We would like to thank Duncan MacCannell, Ketan Patel, Ying Tao, Sue Tong, Blake Cherney, and John Barnes from U.S. Centers for Disease Control and Prevention; Alexander Kim and Salimatu Lukula from the Maryland Department of Health Laboratory; Barbara Downes from Fairfax County, Virginia Department of Health, Catherine Brown, Larry Madoff, Mary DeMartino, Erika Buzby, Stephanie Ash, and Joshua Hall from the Massachusetts Department of Health; Juan Jaramillo from Dallas County LRN; Destiny Hairfield and Kimberly Stratton from Virginia Department of General Services - Division of Consolidated Laboratory Services. We gratefully acknowledge the authors from the originating laboratories and the submitting laboratories, who generated and shared via GISAID genetic sequence data on which this research is based and have included acknowledgements tables (Tables S10 and S11).

Funding:

No external funding was used for the CDC effort in this investigation. BK was supported by the Laboratory Directed Research and Development program of Los Alamos National Laboratory under project number 20220660ER.

References and Notes

1. CDC, MMWR 52, 561–564 (2003). [PubMed: 12816106]
2. Yinka-Ogunleye A et al., Lancet Infect Dis 19, 872–879 (2019). [PubMed: 31285143]
3. Mauldin MR et al., J Infect Dis 225, 1367–1376 (2022). [PubMed: 32880628]
4. Costello V et al.,. Emerg Infect Dis 28, 1002–1005 (2022). [PubMed: 35263559]
5. Rao AK et al., MMWR Morb Mortal Wkly Rep 71, 509–516 (2022). [PubMed: 35389974]
6. CDC,https://www.cdc.gov/poxvirus/monkeypox/response/2022/world-map.html (2022).
7. CDC,https://www.cdc.gov/poxvirus/monkeypox/response/2022/us-map.html (2022).
8. Happi C et al., PLoS Biol 20, e3001769 (2022). [PubMed: 35998195]
9. Likos AM et al., J Gen Virol 86, 2661–2672 (2005). [PubMed: 16186219]
10. Brown K, Leggat PA, Trop Med Infect Dis 1, 1:8 (2016). [PubMed: 30270852]

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

11. Reynolds MG et al., Emerg Infect Dis 13, 1332–1339 (2007). [PubMed: 18252104]
12. Hutson et al., BioMed Research International 2015, 965710 (2015). [PubMed: 26380309]
13. Hadfield et al., Bioinformatics 34, 4121–4123 (2018). [PubMed: 29790939]
14. Reynolds MG et al., Am J Trop Med Hyg 82, 746–754 (2010). [PubMed: 20348530]
15. Li Y et al., J Virol Methods 169, 223–227 (2010). [PubMed: 20643162]
16. Li Y, et al., J Clin Virol 36, 194–203 (2006). [PubMed: 16731033]
17. O'Toole AR, Rambaut A, https://virological.org/t/initial-observations-about-putative-apobec3-deaminase-editing-driving-short-term-evolution-of-mpxv-since-2017/830 (2022).
18. Isidro J et al., Nature Medicine (2022)
19. Harris RS et al.,Virology 479–480, 131–145 (2015).
20. Liddament MT et al., Curr Biol 14, 1385–1391 (2004). [PubMed: 15296757]
21. www.gisaid.org
22. Jolly B and Scaria V, J Infect 22 00476–5 (2022).
23. Esposito JJ et al., Science 313, 807–812 (2006). [PubMed: 16873609]
24. Giorgi EE et al., mBio 11, e00324–20 (2020). [PubMed: 32209678]
25. Keele BF et al., Proc Natl Acad Sci USA 21, 7552–7 (2008).
26. Hobson G et al., Euro Surveill 26, 32:2100745 (2021).
27. Salter JD et al., Trends Biochem Sci 41, 578–594 (2016). [PubMed: 27283515]
28. Vartanian JP et al., Science 320, 230–233 (2008). [PubMed: 18403710]
29. Vartanian JP et al., PLoS Pathog 6, e1000928 (2010). [PubMed: 20523896]
30. Jarmuz A et al., Genomics 79, 285–96 (2002) [PubMed: 11863358]
31. Warren CJ et al., Virus Evolution 1, 1 vev015 (2015) [PubMed: 27774275]
32. Kremer M et al., Virol J 3, 86 (2006). [PubMed: 17052331]
33. Soday L et al., Cell Reports 27, 1920–1933 (2019). [PubMed: 31067474]
34. Katoh K et al., Nucleic Acids Res 30, 3059–3066 (2002). [PubMed: 12136088]
35. Nguyen LT et al., Mol Biol Evol 32, 268–274 (2015). [PubMed: 25371430]
36. Kalyaanamoorthy S et al., Nat Methods 14, 587–589 (2017). [PubMed: 28481363]
37. Storey JD et al., PNAS 100, 9440–9445 (2003). [PubMed: 12883005]
38. Storey JD, False discovery rates. In International Encyclopedia of Statistical Science (2011).
39. Armitage AE et al., J Virol 88, 12882–12894 (2014). [PubMed: 25165112]
40. Refsland EW, et al., PLoS Pathog 8, e1002800 (2012). [PubMed: 22807680]
41. 10.5061/dryad.jq2bvq8cd
42. 10.55876/gis8.220715bk
43. 10.5281/zenodo.6970457

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 1. Phylogenetic analysis of Clade IIb MPXV genome sequences.

Variant B.1 sequences are shown in blue text; variant A.2 is shown in orange; green branches indicate Clade IIb. Phylogenetic analysis was performed in BEAST v1.8.3 using HKY+G model and constant coalescent prior on complete genome alignments after removing all sites containing gaps. Scale bar is in substitutions per site; posterior support values are shown at branch points.

Author Manuscript

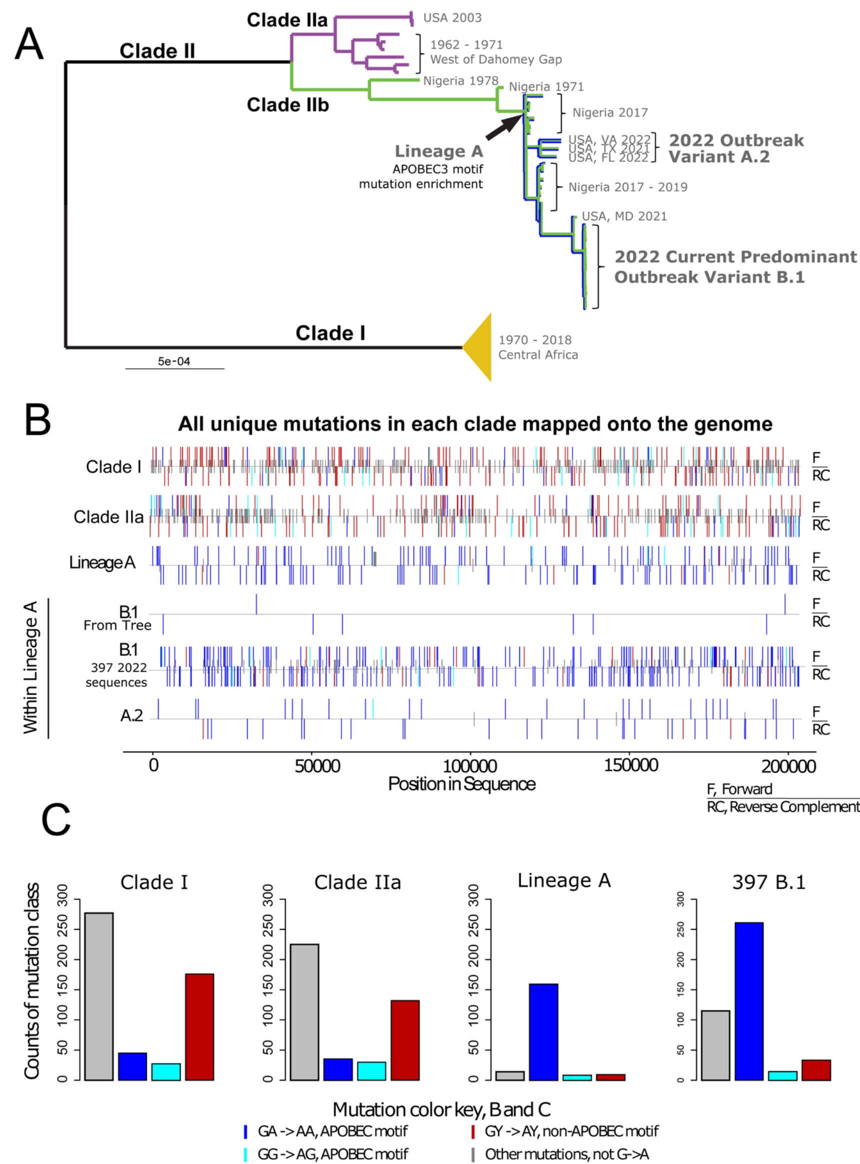

Fig. 2. Analysis of APOBEC3 motif mutations in MPXV.

A. Maximum Likelihood phylogenetic tree using IQ-TREE. Enrichment for G-to-A APOBEC context mutations (shown by blue line) was found in Lineage A, within Clade IIb (green branches). A detailed version of this tree with taxa names is included in Fig. S2. The scale bar indicates the number of substitutions per sequence site. B. Mutational patterns found among Clade I, Clade IIa, and Lineage A in the phylogeny above. All unique mutations in a clade relative to the most recent common ancestor of that clade are shown in a single panel, and the class of each mutation is shown on either the forward or reverse complement strand. The increased number of APOBEC3 context mutations, indicated by blue ticks, in Lineage A, versus red and grey captures the dominance of GA-to-AA APOBEC3 context mutations in this lineage and contrasts with Clades I and IIa. The exact numbers and statistics are provided in Table 2. The A.2 and B.1 variants within Lineage A in the tree, and an updated analysis of 397 variant B.1 sequences from GISAID,

demonstrate the continuing dominance of GA-to-AA APOBEC3 context mutations among currently sampled outbreak sequences. C. Bar charts showing the total number of each class of unique mutations within each clade in Fig. 2B, Clade I, Clade IIa, Lineage A, as well as for 397 sequences from lineage B.1 from Fig. S3A.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 1.

Unique coding nucleotide changes in 2022 predominant MPXV outbreak variant B.1.

Mutations listed were shared among all MPXV Clade IIb variant B.1 sequences examined in Figure 1 and were not present in USA_2021_MD or UK-P2 (MT903344.1). Gene homologs in Vaccinia virus Copenhagen (VACV-Cop) are given for each gene. Pos is position in MT903344.1. Additional information can be found in Table S2.

Pos	MD 2021	2022 Outbreak	TYPE	EFFECT	AA CHANGE	NOTE
3120, 194114	G	A	snp	synonymous		VACV-Cop C19L Ankyrin-like
39148	C	T	snp	missense	Glu353Lys	VACV-Cop F13L major envelope antigen of EEV
73248	G	A	snp	missense	Asp88Asn	VACV-Cop G8R VLTF-1 (late transcription factor 1)
74214	G	A	snp	missense	Met142Ile	VACV-Cop G9R Entry/fusion complex component
77392	G	A	snp	missense	Glu162Lys	VACV-Cop L4R ss/dsDNA binding protein
84596	C	T	snp	synonymous		VACV-Cop J6R RNA polymerase subunit (RPO147)
150480	C	T	snp	missense	His221Tyr	VACV-Cop A46R IL-1/TLR signaling inhibitor
170273	G	A	snp	synonymous		VACV-Cop B12R Ser/Thr Kinase
183534	C	T	snp	missense	Pro722Ser	VACV-Cop B21R Membrane-associated glycoprotein

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 2.

Summary of the counts of different classes of all unique mutations that were observed in different clades shown in Fig. 2.

C-to-T mutations in the forward strand are included as G-to-A in the reverse complement, as a simplified way to tally of all mutations that occurred in the context of an APOBEC3 motif in the viral genome relative to those that did not. The p-values are based on a Fisher's exact test of a contingency table that considered all G's either in an APOBEC context (GA or GG) or not in an APOBEC context (GC or GT), and then determining the number of G-to-A substitutions that occurred within each of the two contexts. ABOBEC motif Odds Ratio >1 indicates APOBEC context enrichment.

Contingency Table										
	Clade	All other mutations	G-to-A APOBEC context	G in APOBEC context, no change	G-to-A non APOBEC context	G not in APOBEC context, no change	p-value 2-sided	q-value	Odds Ratio	GA-to-AA or GG-to-AG
Non-APOBEC context	I	277	72	36552	176	30580	<0.000001	0.000006	0.3423	45 + 27
	IIa	225	65	36548	132	30578	<0.000001	0.000006	0.412	35 + 30
APOBEC context enriched	**Lineage A**	14	167	36415	9	30669	<0.000001	0.000006	**15.63**	159 + 8
	B.1, N=12	0	8	36517	0	30675	0.0095	0.022	**inf**	8 + 0
	B.1, N=397	115	275	35064	33	29705	<0.000001	0.000006	**7.06**	261 + 14
	A.2	3	52	36516	3	30674	<0.000001	0.000006	**14.56**	51 + 1