Received: 21 July 2022 | Revised: 1 September 2022 | Accepted: 4 September 2022

DOI: 10.1111/jam.15806

Journal of Applied Microbiology Applied Microbiology International

ORIGINAL ARTICLE

Comparison of Monkeypox virus genomes from the 2017 Nigeria outbreak and the 2022 outbreak

Trudy M. Wassenaar[1] | **Visanu Wanchai**[2] | **David W. Ussery**[2]

[1]Molecular Microbiology and Genomics Consultants, Zotzenheim, Germany

[2]Department of Biomedical Informatics, University of Arkansas for Medical Sciences, Little Rock, Arkansas, USA

Correspondence
Trudy M. Wassenaar, Molecular Microbiology and Genomics Consultants, Tannenstraße 7, D-55576 Zotzenheim, Germany.
Email: trudy@mmgc.eu

Abstract

Aims: The current Monkeypox virus (MPX) outbreak is not only the largest known outbreak to date caused by a strain belonging to the West-African clade, but also results in remarkably different clinical and epidemiological features compared to previous outbreaks of this virus. Here, we consider the possibility that mutations in the viral genome may be responsible for its changed characteristics.

Methods and Results: Six genome sequences of isolates from the current outbreak were compared to five genomes of isolates from the 2017 outbreak in Nigeria and to two historic genomes, all belonging to the West-African clade. We report differences that are consistently present in the 2022 isolates but not in the others. Although some variation in repeat units was observed, only two were consistently found in the 2022 genomes only, and these were located in intergenic regions. A total of 55 single nucleotide polymorphisms were consistently present in the 2022 isolates compared to the 2017 isolates. Of these, 25 caused an amino acid substitution in a predicted protein.

Conclusions: The nature of the substitution and the annotation of the affected protein identified potential candidates that might affect the virulence of the virus. These included the viral DNA helicase and transcription factors.

Significance: This bioinformatic analysis provides guidance for wet-lab research to identify changed properties of the MPX.

KEYWORDS

evolution, Monkeypox virus, outbreak, repeat, single nucleotide polymorphism

INTRODUCTION

Monkeypox virus (MPX, Poxviridae family) is a member of the genus Orthopoxvirus, to which Cowpox virus and Variola virus (smallpox) also belong. The genomes of these viruses consist of a linear double-strand DNA molecule of around 200 kilobasepairs (kbp), in the order of a small bacterial genome, and contains around 200 protein-coding genes. The genome is transcribed in the cytosol of an infected host cell, for which they use their own RNA polymerase.

Protein expression of orthopoxvirus species takes place in three phases. During the first phase, pre-packed enzymes remove the coat of the virus and assist to express early-phase genes, to be followed by expression of intermediate proteins that replicate the genome, and finally the late proteins are produced that are needed for the production of new virus particles (virions). Late-phase genes

are not all co-located on the genome but rather early and late genes can flank each other (Shchelkunov et al., 2002). As summarized elsewhere, the brick-shaped virions bind to glycosaminoglycans of their target host cells and enter via endocytosis (Kumar et al., 2022). The primary site of infection may be mucosal surfaces, the skin, or the respiratory tract, depending on the mode of transmission. The virus then spreads through the lymph system and causes viraemia; the typical pock lesions result from infection of the skin. Due to immune cross-reactivity, smallpox (Variola virus that could exclusively infect humans) could be eradicated with Vaccinia; the virus used for vaccination is closely related to horsepox and cowpox but its origin is not exactly known. Likewise, individuals immunized with vaccinia are protected against MPX. However, since vaccinia vaccination stopped in the 1970s there is now a considerable human population that is immunologically naïve to Orthopoxvirus species.

Monkeypox virus is endemic in Africa, where its natural hosts are small mammals. The virus can infect a variety of hosts. Based on phylogenetic analyses of their genomes, two clades of MPX are recognized: the Congo Basin (CB) clade and the West Africa (WA) clade (Likos et al., 2005; Nakazawa et al., 2015). The CB clade had caused a total of 760 human cases between 2005 and 2007 in the Democratic Republic of Congo (DRC) alone (Kugelman et al., 2014). The WA clade is less virulent, with a lower mortality and morbidity. For decades, human MPX cases were mostly restricted to zoonotic infection with limited human-to-human spread, although small-scale outbreaks and exported cases have been described, such as in 2003 in the US. However, in 2017 a large outbreak of the WA clade occurred in Nigeria, with a hotspot in Bayelsa state, that involved 122 confirmed or probable cases and 7 deaths (Alakunle et al., 2020; Yinka-Ogunleye et al., 2019). Even then, clinical symptoms remained typical, with pox marks on all parts of the body including the hands and in particular the face, together with fever, headache and malaise. A recent systematic review of the literature on MPX recognized the CB clade (called the Central African clade in that publication) as the source of most human cases, with DRC as a hotspot (Bunge et al., 2022).

The picture changed in 2022, when human cases of MPX appeared in multiple developed countries that could no longer epidemiologically be directly related to each other and no longer had links to travels to Africa. The number of infections rapidly increased, suggesting a more effective human-to-human spread than previously recorded. Genome sequence analysis revealed that all cases are related to the Nigeria 2017 outbreak of the WA clade (Velavan & Meyer, 2022). However, the epidemiology and clinical symptoms no longer resemble what was typically observed prior to 2022. During the ongoing epidemic, cases are often (but not always) observed in men who have sex with men (MSM; Girometti et al., 2022; Minhaj et al., 2022; Tarín-Vicente et al., 2022). Involved cases showed that the typical rash is now often present in the genital region instead of face and hands (Venkatesan, 2022). This hints towards transmission during sexual activity due to close physical contact, a possibility first proposed during the Nigeria outbreak, where a number of patients were young men who also suffered from sexually transmitted diseases including syphilis and HIV, and where genital MPX lesions were observed (Ogoina et al., 2019). The authors of that publication speculated that sexual activity may enable transfer by close skin-to-skin contact, although they did not rule out that 'genital secretions' may have contributed. In the meantime, MPX virus has been detected in semen, but it remains unclear whether this is relevant to transmission (Kupferschmidt, 2022).

In July 2022 the WHO declared the MPX outbreak a global health emergency. The rapid spread might indicate a recent, unnoticed spread in the general population that is only now surfacing, but given the clinical symptoms of acute infection this is unlikely. Instead, the spread has been attributed to sexual promiscuity of affected individuals (Kupferschmidt, 2022). Alternatively, other changes are responsible for the novel features of this epidemic. It seems likely that at least some of the rapid spread now observed is due to asymptomatic infections.

Given the change in epidemiology, spread and clinical signs of MPX, we wondered if this new viral behaviour could be attributed to recent changes in the viral genome, a possibility also considered by others (Kupferschmidt, 2022). To address this question, 13 high-quality genomes were compared that all belonged to the WA clade of MPX, including five from the Nigerian 2017 outbreak and six from different countries from the current 2022 outbreak. Two WA-clade genomes of historical isolates were also included in the comparison. The findings reported here may provide a valuable guidance for wet-lab research.

MATERIALS AND METHODS

MPX genomes were selected from GenBank based on the following criteria: selected genomes had to be sized at least 197.5 kbp, belong to the WA clade and contain fewer than 70 ambiguous nucleotides. Based on the order of submission at GenBank, the first five genomes related to the 2017 Nigeria outbreak meeting these criteria were thus selected (GenBank accession numbers MK783028.1, MK783029.1, MK783030.1, MK783031.1 and MK783032.1), as were the first six genomes from the current outbreak that originated from different countries (ON563414.3 from the USA, ON568298.1 from Germany, ON585035.1 from Portugal,

ON602722.2 from France, ON609725.2 from Spain and ON619835.2 from the UK). Since a monophylogenetic origin of the current outbreak could be expected, and since the MPX genome mutates relatively infrequently, more genomes were not selected, though two historical isolates of the WA clade were also included: strain Liberia_1970_184 (DQ011156.1) isolated in Liberia in 1970 (Likos et al., 2005), which served as the reference genome and was used for nucleotide numbering, and a 1965 isolate sequenced in the Netherlands with an unidentified African origin (strain UTC, KJ642614.1) which was described as highly similar to strain Liberia_1970_184 (Nakazawa et al., 2015).

A phylogenetic tree (maximum likelihood) was constructed with the genomes by FastTree2.1 using the FFT-NS-I iterative refinement method (Price et al., 2010). The genome of strain Liberia_1970_184 was taken as the outgroup. A genome atlas was constructed for that same genome using CMG BioTools (Vesth et al., 2013). The fastA files of these genomes were compared by multiple alignments using ClustalW. Open reading frames were checked with 'translate' at https://web.expasy.org/translate. Locations of PFAM domains were checked at http://pfam.xfam.org.

RESULTS

A phylogenetic tree of the 13 selected genomes is shown in Figure 1. The tree was rooted with the 1970 isolate and shows that the two historical isolates are more similar to each than to the other genomes. All 2017 isolates from Liberia group together. The 2022 isolates are slightly more variable but are all found on one major branch. The observed slight variation indicates the need to include multiple genomes for each time period in order to identify consistent mutations rather than variation caused by micro-evolution during an ongoing outbreak.

A genome atlas of the genome of strain Liberia_1970_184 is shown in Figure 2. The red in the lane at the bottom (lane H) shows that the genome is AT rich (67% on average), and the top three lanes of this atlas show structural features of the DNA (Jensen et al., 1999; Pedersen et al., 2000). In combination, information on intrinsic curvature, stacking energy and position preference can provide information on how flexible the DNA is or how easily it melts, for a given position. The predicted protein-coding genes are shown in lane D of Figure 2. The open reading frames are found on both strands, but they dominate at the complementary strand for the first 50 kbp and at the other strand for the last 50 kbp. The genes in the centre of the genome are mostly essential to all orthopoxviruses, whereas the genes located towards the termini have more species- and strain-specific variation (Shchelkunov et al., 2002). Note that there are few large global direct repeats (absence of blue in lane E), but like all poxviridae members, the linear MPX genome contains inverted terminal repeats (ITRs), which are nearly 10 kbp long for this virus, and these shows up as dark red in lane F. The genes located in this repeat are of course found on opposite strands at both flanks. The GC-skew (lane G) shows a bias of guanine towards the coding strand of genes (where genes are oriented in the forward direction, shown as blue in the annotation lane D), and genes on the other strand (red in lane D) have a bias of more cytidine. Only one region combines a strong GC-skew with a low stacking energy, located around 135 kbp (Figure 2). In combination, these features indicate local low melting properties. This may be related to the fact that a number of proteins located here are expressed only in the late phase.

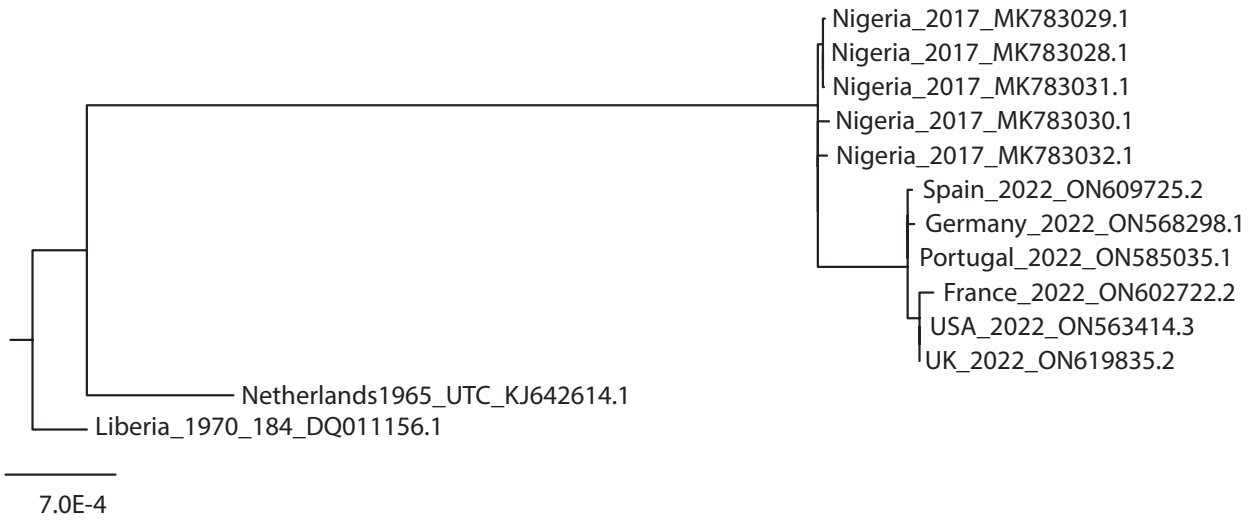

FIGURE 1 Phylogenetic tree (Maximum Likelihood) of the 13 Monkeypox virus genomes included in this study. The tree was rooted with the Liberia 1970 genome.

Journal of Applied Microbiology Applied Microbiology International

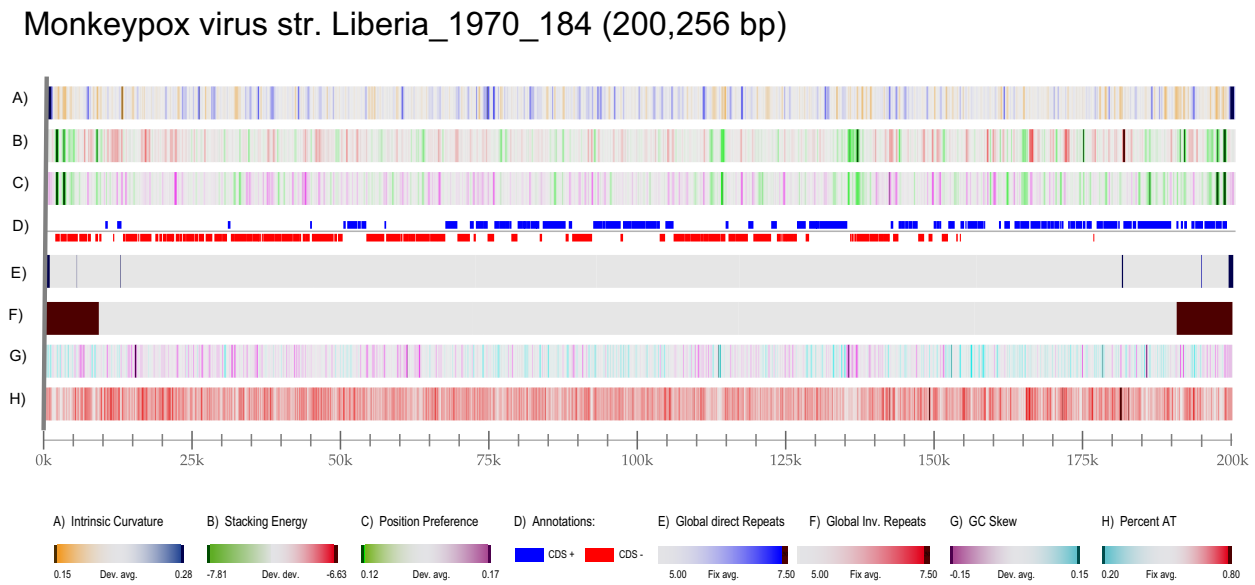

FIGURE 2 Genome atlas of Monkeypox virus strain Liberia_1970_184. The top three lanes (A–C) show DNA structural features (intrinsic curvature, stacking energy and position preference), followed by protein-coding genes on the positive (blue) and negative (red) strand. Lane (E) shows global direct repeats are mostly absent, while Lane (F) indicates the long inverted repeats at both flanks. Lanes (G) and (H) show the base composition features GC-skew and AT-content, respectively.

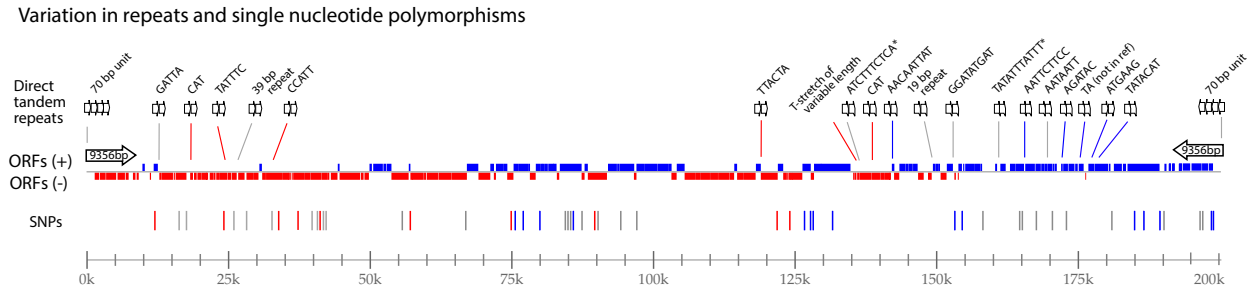

FIGURE 3 Genome organization of Monkeypox virus based on strain Liberia_1970_184. The protein-coding genes are shown in blue for the positive strand and in red for the negative strand. Above these, the location of direct tandem repeats that vary in repeat units in the 13 analysed genomes is shown. The repeat unit for each location is indicated (or, when >20 bp, its size) and the connecting line is coloured red or blue when the repeat occurs in an open reading frame, and grey when present in intergenic regions. The asterisks indicate consistent repeat unit numbers in 2022 genomes that differ from 2017 genomes. The TA-repeat around 153 kb is absent in the represented reference genome but was found in isolates from the 2017 and 2022 outbreaks, with variable length. Below the genes the position of single-nucleotide polymorphisms (SNPs) is shown as lines, again coloured for the gene orientation in which they occur. SNPs not affecting protein sequences are shown as grey lines.

The last lane shows that the AT content has relatively little variation along the genome.

A multiple alignment of the 13 genomes served as the basis of further analyses. Important features are graphically summarized in Figure 3. We concentrated our analysis on variable repeat units and on single nucleotide polymorphisms (SNPs).

Repeat analysis

The most extensive repeat of this virus is formed by the already-mentioned ITR. The very ends of these ITRs contain 12 direct repeats of units sized 70 bp per unit. The ITRs are presumably conserved between all MPX genomes, although the sequences analysed here did not cover the 12 direct repeats at the very ends, with the exception of the reference genome. Variation in the number of other tandem direct repeat units was a common reason for gaps in the alignment. Those tandem direct repeats that were present in variable numbers in the 13 genomes are summarized at the top of Figure 3. The graphic includes all variations in repeat unit numbers, not only those consistently found in the 2022 genomes. Direct tandem repeats that varied between the strains are notoriously absent in the central region of the genome, between 35

and 120 kbp (Figure 3). In most repeats summarized in the figure, the number of repeat units in a particular location was shared between the 2017 and 2022 genomes. The two repeats with consistently conserved numbers in the 2022 genomes were both located in intergenic regions.

In 13 occasions, repeat units were present within a predicted open reading frame, resulting in low-complexity amino acid sequences. The T-stretch of variable length around 135,530 bp truncated the protein annotated as 'cowpox A-type inclusion protein', in most of the analysed genomes other than in the reference strain (AAY97535.1, gene locus MPXV_LIB1970_184_142). A start codon downstream of the repeat TATACAT in a serine protease inhibitor (AAY97588.1, LIB1970_184_195, around 181,600 kbp) might avoid incorporation of long Ile–Ile–Tyr stretches in this protein. Possibly, that start codon is also used by the other strains; protein annotation of virus genomes is mostly based on inference of other virus genomes (Ehlers et al., 2002; Likos et al., 2005), and few of the annotated genes have been experimentally verified for MPX. Likewise, the dinucleotide (TA) repeat around position 175,650 kbp shifted the start of the open reading frame of AAY97583.1 in some of the strains (it is absent in the reference strain). Interestingly, this protein is annotated as "bifunctional IL-1-β-inhibitor/prevents febrile response in VAC-mouse intranasal model" but Likos et al. described this protein as an IL-1-β-receptor (not IL-1-β-inhibitor) that inhibits IL1-induced febrile responses, citing previous mouse work performed with vaccinia (Alcami & Smith, 1992; Spriggs et al., 1992). Likos et al. assumed that this protein was absent in WA-clade strains and took this as further explanation for the decreased virulence of WA strains. However, an open reading frame with the telltale immunoglobulin domain #3 is present in all WA strains analysed here.

In conclusion, the repeat regions were not considered to be responsible for the changed epidemiological features, although it cannot be entirely ruled out that some longer or shorter repeat units affect protein function or gene expression, for instance by varying the distance between a promoter and its gene. This was not further assessed.

SNP analysis

The multiple alignment of the 13 genomes was also used to identify single-nucleotide mutations that were consistently present in the 2022 isolates but not in the 2017 genomes. Their positions are indicated in the lower half of Figure 3. Of note is that some regions of the genome are devoid of any mutations, e.g., between 97 and 120 kb and also between 133 and153 kb. Singleton mutations that were only detected once in the dataset were ignored, as these results of microevolution were unlikely to contribute to the differences in epidemiology that are observed during the current outbreak as opposed to the 2017 outbreak or the historic cases. A total of 55 SNPs were consistently identified between the 2017 and 2022 isolates, and these are summarized in Table S1. One of these was polymorphic within the four 2017 isolates, and another was polymorphic within the six 2022 isolates. Of the 55 SNPs summarized in Table S1, 8 were in intragenic regions and 22 were synonymous so they did not change the amino acid in the deduced protein sequence. The positions of these 30 SNPs are shown in grey in Figure 3. The remaining 25 mutations lead to an amino acid substitution and these are listed in Table 1; they are represented as red or blue lines in Figure 3, depending on the orientation of the gene in which they reside.

A number of amino acid substitutions result in a relatively minor protein change, but there are three glutamic acid (E) to lysine (K) substitutions, which means a negative charge is changed to a positive charge at that location. A substitution of arginine (R) with threonine (T) also involves a change in charge, now replacing a positive with a neutral/polar amino acid, and going from aspartic acid (D) to asparagine (N) replaces a negative change with a neutral/polar side residue. Two substitutions changed proline (P) to serine (S), which may have structural consequences for a protein, as proline can hamper the formation of alpha-helices. A change of serine (S) to phenylalanine (F) introduces a bulky, hydrophobic amino acid at a position where a polar small side chain was present.

It was checked if these amino acid substitutions were located in a PFAM domain, and if so, whether the mutation involved an amino acid crucial to that domain. This information is added in the last column of Table 1. Substitutions that affected the PFAM domain at a crucial amino acid were identified for the Kelch-like protein AAY97432.1, the virion structural protein AAY97475.1 and the intermediate gene transcription factor VITF-3 (AAY97533.1) (Table 1).

PFAM domains are based on conservation of amino acids in a diverse set of proteins that all have the same function. Such conservation can be based on relatively few but crucial amino acids, whose position is structure-dependent. It is therefore no surprise that, although many mutations fall within a recognized PFAM domain, they do not involve crucial amino acids. Interestingly, the DNA helicase protein (AAY97528_) did not align to a PFAM domain with significance, but it contained three amino acid substitutions, two of which involved a change in charge.

Any of the non-synonymous single-nucleotide mutations could in principle contribute to changes in pathogenicity or transferability/infectivity. We note that three protein genes contained more than one non-synonymous substitution: the DNA helicase contained three (E62K,

TABLE 1 Non-synonymous mutations that were consistently found in six MPX isolates originating from various countries in 2022 compared to five 2017 isolates from Nigeria

Position[a]	Mutation	Locus tag[a]	Protein function annotation and accession number	Amino acid change[b]	In PFAM domain[c]
12,059	G to A (polymorphic within 2017 genomes)	MPXV_LIB1970_184_013	Bifunctional zinc finger-like protein/E3 ubiquitin ligase AAY97410.1	A66T	[KIIA-N]
33,479	G to A	MPXV_LIB1970_184_038	Kelch-like protein (protein–protein and DNA interactions) AAY97432.1	R48C Pos to polar	BTB/POZ
36,879	G to A	MPXV_LIB1970_184_044	Product unknown AAY97438.1	P78S Pro removed	[In L1R F9L]
41,080	C to T	MPXV_LIB1970_184_047	Product unknown AAY97441.1	E125K Neg to Pos	[Pox F12L]
57,056	C to A	MPXV_LIB1970_184_063	Sulfhydryl oxidase AAY97456.1	D56E	[Pox E10]
74,784	C to T	MPXV_LIB1970_184_082	Virion structural protein AAY97475.1	D196N Neg to polar	Pox G7
75,488	C to T	MPXV_LIB1970_184_083	Late gene transcription VLTF-1 AAY97476.1	S30L	[Pox TAP]
75,661	G to A			D88N Neg to polar	
76,628	G to A	MPXV_LIB1970_184_084	Myristyl protein AAY97477.1	M142N	n.a.
79,805	G to A	MPXV_LIB1970_184_088	Core protein vp8 AAY97481.1	E162K Neg to Pos	[Pox VP8 L4R]
85,750	C to T	MPXV_LIB1970_184_095	DNA-dependent RNA polymerase subunit rpo147 AAY97488.1	S734L	[RNA pol Rpb1_5]
89,654	G to A	MPXV_LIB1970_184_099	Protein RAP94 (associated with RNA pol, early-stage transcription) AAY97492.1	Y739H	[Pox Rap94]
121,731	C to T	MPXV_LIB1970_184_126	Precursor p4a of core protein 4a AAY97519.1	D98N Neg to polar	[Pox P4A]
123,756	C to T	MPXV_LIB1970_184_129	IMV membrane protein AAY97522.1	A17T	n.a.
126,566	G to A	MPXV_LIB1970_184_135	DNA helicase AAY97528.1	E62K Neg to Pos	n.a.
127,110	G to A			R243Q Pos to polar	n.a.
127,685	G to A			E435K Neg to Pos	n.a.

(Continues)

TABLE 1 (Continued)

Position[a]	Mutation	Locus tag[a]	Protein function annotation and accession number	Amino acid change[b]	In PFAM domain[c]
131,134	C to T	MPXV_LIB1970_184_140	45 kDa large subunit of intermediate gene transcription factor VITF-3 AAY97533.1	S307L	Pox int trans
152,888	C to T	MPXV_LIB1970_184_166	Toll/IL1-receptor [TIR]-like protein AAY97559.1	H221Y	n.a.
153,874	A to C	MPXV_LIB1970_184_168	Unknown AAY97561.1	L41R Hydrophobic to pos	n.a.
184,395	G to A	MPXV_LIB1970_184_197	Putative membrane-associated glycoprotein AAY97590.1	D209N Neg to polar	n.a.
185,934	C to T			P722S Pro removed	
188,993	G to A			M1741S	
197,040	C to T	MPXV_LIB1970_184_205	NF-alpha-receptor-like protein AAY97597.1	S54F Small to bulky	n.a.
198,388	C to T (polymorphic within 2022 genomes)	MPXV_LIB1970_184_206	Chemokine-binding protein AAY97598.1	S111L	[Orthopox 35kD]

[a]Nucleotide positions and locus tags are given with reference to monkeypox virus strain Liberia_1970_184, GenBank accession number DQ011156.1.

[b]The nature of significant changes due to the indicated amino acid substitutions is indicated, with pos: positively charged, neg: negatively charged.

[c]Substitutions in the indicated PFAM domain not affecting the match are shown between parentheses. n.a.: not applicable as no PFAM domain match was identified that covered the position of the mutation.

R243Q and E435K), the late gene transcription factor VLTF-1a contained two (S30L and D88N) and a putative membrane-associated glycoprotein also contained two substitutions (P722S and M1741S). However, the latter gene contained a two-nucleotide insert in the French 2022 genome which would introduce a frameshift following amino acid L725. This suggests that (the rest of) this protein is not essential for the virus. Nevertheless, membrane-associated glycoproteins are often involved in host-cell interactions and may play a role in tissue tropism. Mutations in transcription factors can have downstream effect on the expression of other proteins. For proteins interacting with negatively charged DNA, a charge change towards less negative or more positive might improve protein-DNA affinity. This applies to both VLTF-1 and DNA helicase. If production of late proteins, in which VLTF-1 is involved, would be a rate-limiting step in virion production, higher expression of this transcription factor could potentially result in higher viral loads. Likewise, mutations in DNA helicase, a protein essential for viral replication, can affect the production of virions, when its activity is rate limiting during replication. MPX contains two helicase genes, one coding for a bifunctional DNA/RNA helicase (AAY97467.1) and the DNA helicase in which the three non-synonymous SNPs were identified (AAY97528.1). If any of the noted mutations would result in more efficient virion production in the human host, it could in principle increase infectivity, although this remains hypothetical at this stage.

DISCUSSION

Here, we report differences in the genomes of MPX strains isolated during the 2022 global outbreak when compared to closely related isolates from the 2017 outbreak that took place in Nigeria. The comparisons identified variation in the number of repeat sequences, in particular closer to the ends of the genome, but we consider it unlikely that variation in repeat numbers has resulted in behavioural changes of the virus. If changes in the genetic makeup of MPX are responsible for the more frequent human-to-human spread observed during the current outbreak, single nucleotide changes are more likely responsible. We noted 25 mutations that lead to an amino acid substitution, 14 of which represented a change that potentially would result in changes of protein function. In an attempt to predict such functional alterations, conservation of the PFAM domains was checked, which identified three proteins in which a PFAM domain had changed: a Kelch-like protein, a virion structural protein and the intermediate gene transcription factor VITF-3. In particular changes in transcription factor activity can result in downstream effects, although it needs to be experimentally demonstrated whether the single-nucleotide variation we identified actually results in changed protein activity. Of interest is the finding of three amino acid changes, which involve a change in charge, in the helicase gene of MPX. This can affect the efficiency of replication and hence virus loads, but again, this would need to be tested.

Limited information is available on the proteins that determine the degree of pathogenicity of MPX strains. Early work where MPX strains belonging to the CB and WA clades were compared to each other and to vaccinia had identified a number of differences between the two MPX clades that could explain the lesser virulence of WA strains (Likos et al., 2005). Our analysis suggests that none of those proteins had undergone mutations between WA isolates of 2017 and 2022. The comparison of MPX genomes presented here is based on the assumption that one or more changes in the viral genome might be responsible for the changes in epidemiology that were recently seen: whereas MPX was typically a zoonosis with limited human-to-human spread and was mostly restricted to regions in Africa, the ongoing outbreak has spread to multiple countries on other continents, with different clinical presentations (Kumar et al., 2022; Kupferschmidt, 2022; Velavan & Meyer, 2022; Venkatesan, 2022). The virus responsible for the current outbreak is genetically closely related to the 2017 outbreak of the WA clade strain that occurred in Nigeria, with—up to 2022—limited spill over to other countries, although travel-related cases had been recorded from the UK, Israel and Singapore (Mauldin et al., 2022). The current outbreak is the result of increased human-to-human spread and currently involves a high number of cases in MSM (Girometti et al., 2022; Minhaj et al., 2022; Tarín-Vicente et al., 2022). However, it is not proven whether changes in the virus are responsible for these new characteristics.

The rapid spread between and within countries during the current outbreak is worrying. The initial rapid spread might be explained by a limited number of high-spreading events in which multiple human interactions took place. Networks describing intimate human relationships are like any other network, where multiple nodes connect a few individuals only, but some individuals function as 'hubs' that connect to many individuals. However, the sudden increase of MPX is more likely an indicator of more rapid human-to-human spread as a result of some genomic change, and the virus may continue to spread beyond particular sub-populations.

Scientists should keep an open mind to possible changes in the virus' behaviour that may be the result of accumulated mutations. For this reason, we took the effort to list those mutations that may be likely candidates to partly explain the changed epidemiology. The approach

followed here is not suitable to provide conclusive evidence for mutations involved in behavioural changes of the virus, but the reported findings can serve as a basis for future research.

ACKNOWLDGEMENT

This work was supported in part by the National Science Foundation, Award No. OIA-1946391.

CONFLICT OF INTEREST

None of the authors have a conflict of interest to be declared.

ORCID

Trudy M. Wassenaar https://orcid.org/0000-0002-7024-1139
Visanu Wanchai https://orcid.org/0000-0003-2243-3838
David W. Ussery https://orcid.org/0000-0003-3632-5512

REFERENCES

Alakunle, E., Moens, U., Nchinda, G. & Okeke, M.I. (2020) Monkeypox Virus in Nigeria: Infection Biology, Epidemiology, and Evolution. *Viruses.*, 5;12(11):1257.

Alcami, A. & Smith, G.L. (1992) A soluble receptor for interleukin-1 beta encoded by vaccinia virus: a novel mechanism of virus modulation of the host response to infection. *Cell*, 71, 153–167.

Bunge, E.M., Hoet, B., Chen, L., Lienert, F., Weidenthaler, H., Baer, L.R. et al. (2022) The changing epidemiology of human monkeypox-A potential threat? A systematic review. *PLoS Neglected Tropical Diseases*, 16(2), e0010141. Available from: https://doi.org/10.1371/journal.pntd.0010141

Ehlers, A., Osborne, J., Slack, S., Roper, R.L. & Upton, C. (2002) Poxvirus orthologous clusters (POCs). *Bioinformatics*, 18, 1544–1545.

Girometti, N., Byrne, R., Bracchi, M., Heskin, J., McOwan, A., Tittle, V. et al. (2022) Demographic and clinical characteristics of confirmed human monkeypox virus cases in individuals attending a sexual health centre in London, UK: an observational analysis. *The Lancet Infectious Diseases*, 22, 1321–1328.

Jensen, L.J., Friis, C. & Ussery, D.W. (1999) Three views of microbial genomes. *Research in Microbiology*, 150(9–10), 773–777.

Kugelman, J.R., Johnston, S.C., Mulembakani, P.M., Kisalu, N., Lee, M.S., Koroleva, G. et al. (2014) Genomic variability of monkeypox virus among humans, Democratic Republic of the Congo. *Emerging Infectious Diseases*, 20, 232–239.

Kumar, N., Acharya, A., Gendelman, H.E. & Byrareddy, S.N. (2022) The 2022 outbreak and the pathobiology of the monkeypox virus. *Journal of Autoimmunity*, 131, 102855. Available from: https://doi.org/10.1016/j.jaut.2022.102855

Kupferschmidt, K. (2022) Why monkeypox is mostly hitting men who have sex with men. *Science*, 376(6600), 1364–1365.

Likos, A.M., Sammons, S.A., Olson, V.A., Frace, A.M., Li, Y., Olsen-Rasmussen, M. et al. (2005) A tale of two clades: monkeypox viruses. *The Journal of General Virology*, 86(Pt 10), 2661–2672.

Mauldin, M.R., McCollum, A.M., Nakazawa, Y.J., Mandra, A., Whitehouse, E.R., Davidson, W. et al. (2022) Exportation of monkeypox virus from the African continent. *Journal of Infectious Diseases*, 225(8), 1367–1376.

Minhaj, F.S., Ogale, Y.P., Whitehill, F., Schultz, J., Foote, M., Davidson, W. et al. (2022) Monkeypox outbreak - nine states, may 2022. *Morbidity and Mortality Weekly Report*, 71(23), 764–769.

Nakazawa, Y., Mauldin, M.R., Emerson, G.L., Reynolds, M.G., Lash, R.R., Gao, J. et al. (2015) A phylogeographic investigation of African monkeypox. *Viruses*, 7(4), 2168–2184.

Ogoina, D., Izibewule, J.H., Ogunleye, A., Ederiane, E., Anebonam, U., Neni, A. et al. (2019) The 2017 human monkeypox outbreak in Nigeria-report of outbreak experience and response in the Niger Delta University Teaching Hospital, Bayelsa State, Nigeria. *PLoS One*, 14(4), e0214229. Available from: https://doi.org/10.1371/journal.pone.0214229

Pedersen, A.G., Jensen, L.J., Brunak, S., Staerfeldt, H.H. & Ussery, D.W. (2000) A DNA structural atlas for *Escherichia coli*. *Journal of Molecular Biology*, 299(4), 907–930.

Price, M.N., Dehal, P.S. & Arkin, A.P. (2010) FastTree 2-approximately maximum-likelihood trees for large alignments. *PLoS One*, 5(3), e9490.

Shchelkunov, S.N., Totmenin, A.V., Safronov, P.F., Mikheev, M.V., Gutorov, V.V., Ryazankina, O.I. et al. (2002) Analysis of the monkeypox virus genome. *Virology*, 297(2), 172–194.

Spriggs, M.K., Hruby, D.E., Maliszewski, C.R., Pickup, D.J., Sims, J.E., Buller, R.M. et al. (1992) Vaccinia and cowpox viruses encode a novel secreted interleukin-1-binding protein. *Cell*, 71, 145–152.

Tarín-Vicente, E.J., Alemany, A., Agud-Dios, M., Ubals, M., Suñer, C., Antón, A. et al. (2022) Clinical presentation and virological assessment of confirmed human monkeypox virus cases in Spain: a prospective observational cohort study. *Lancet*, 400(10353), 661–669.

Velavan, T.P. & Meyer, C.G. (2022) Monkeypox 2022 outbreak: an update. *Tropical Medicine & International Health*, 7(7), 604–605.

Venkatesan, P. (2022) Global monkeypox outbreak. *The Lancet Infectious Diseases*, 22(7), 950.

Vesth, T., Lagesen, K., Acar, Ö. & Ussery, D. (2013) CMG-biotools, a free workbench for basic comparative microbial genomics. *PLoS One*, 8(4), e60120.

Yinka-Ogunleye, A., Aruna, O., Dalhat, M., Ogoina, D., McCollum, A., Disu, Y. et al. (2019) Outbreak of human monkeypox in Nigeria in 2017–18: a clinical and epidemiological report. *The Lancet Infectious Diseases*, 19(8), 872–879.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Wassenaar, T.M., Wanchai, V. & Ussery, D.W. (2022) Comparison of Monkeypox virus genomes from the 2017 Nigeria outbreak and the 2022 outbreak. *Journal of Applied Microbiology*, 133, 3690–3698. Available from: https://doi.org/10.1111/jam.15806