Genomic diversity of mpox virus in Paris area (France) during the 2022 outbreak

Géraldine Piorkowski\textsuperscript{1}, Jade Ghosn\textsuperscript{2}, Romain Coppée\textsuperscript{2}, Morgane Mailhe\textsuperscript{2}, Valentine M. Ferré\textsuperscript{2}, Nadhira Houhou-Fidouh\textsuperscript{2}, Yazdan Yazdanpanah\textsuperscript{2}, Quentin Le Hingrat\textsuperscript{2}, Hervé Raouï\textsuperscript{3}, Charlotte Charpentier\textsuperscript{2}, Diane Descamps\textsuperscript{2}, and Xavier de Lamballerie\textsuperscript{1}

\textsuperscript{1}Aix-Marseille Université
\textsuperscript{2}Hôpital Bichat - Claude-Bernard
\textsuperscript{3}Agence Nationale de Recherches sur le SIDA et les Hepatites Virales

April 15, 2023

Abstract

In May 2022, several countries reported mpox cases from patients without history of traveling to endemic areas. France was one of the most affected European countries. In this study, we described the clinical characteristics of mpox cases in France, and studied the genetic diversity of the virus. Patients diagnosed with mpox infection (qPCR ct<28) between 21\textsuperscript{th} May and 4\textsuperscript{th} July 2022 and between 16\textsuperscript{th} August and 10\textsuperscript{th} September 2022 were included to this study. Twelve amplicons corresponding to the most polymorphic regions of the mpox genome and covering \textasciitilde30,000 nucleotides were generated and sequenced using the S5 XL Ion Torrent technology to evaluate the genetic diversity of mpox sequences. One hundred and forty-eight patients were diagnosed with mpox-infection. 95\% were men, 5\% transgender (M-to-F), 50\% were taking HIV pre-exposure prophylaxis, and 25\% were HIV seropositive. One hundred and sixty-two samples were sequenced and compared to GenBank sequences. Thirty-two distinct mutational patterns were identified. Overall, low genetic diversity of mpox sequences was found compared with pre-epidemic Western-African sequences, with 32 distinct mutational patterns. Our results provide a first glance at the mutational landscape of early mpox 2022 circulating strains in Paris (France).
Abstract

In May 2022, several countries reported mpox cases from patients without history of traveling to endemic areas. France was one of the most affected European countries. In this study, we described the clinical characteristics of mpox cases in France, and studied the genetic diversity of the virus.

Patients diagnosed with mpox infection (qPCR ct<28) between 21th May and 4th July 2022 and between 16th August and 10th September 2022 were included to this study. Twelve amplicons corresponding to the most polymorphic regions of the mpox genome and covering ~30,000 nucleotides were generated and sequenced using the S5 XL Ion Torrent technology to evaluate the genetic diversity of mpox sequences.

One hundred and forty-eight patients were diagnosed with mpox-infection. 95% were men, 5% transgender (M-to-F), 50% were taking HIV pre-exposure prophylaxis, and 25% were HIV seropositive. One hundred and sixty-two samples were sequenced and compared to GenBank sequences. Thirty-two distinct mutational patterns were identified. Overall, low genetic diversity of mpox sequences was found compared with pre-epidemic Western-African sequences, with 32 distinct mutational patterns.

Our results provide a first glance at the mutational landscape of early mpox 2022 circulating strains in Paris (France).

Keywords: mpox, genetic diversity, sequencing, 2022 outbreak, phylogenomics, France, co-circulation

Material and methods

Ethic statement
The Research Ethics Committee of the North Parisian Academic Hospital Group approved the study. All the participants included in the study gave consent for the use of their anonymized medical data.

Sample collection
The study was conducted in Bichat-Claude Bernard university hospital, one of the mpox referral centers in the Paris Metropolitan area (France), between 21st May and 4th July 2022 and between 16th August and 10th September 2022. Patients were diagnosed with mpox infection by quantitative polymerase chain reaction (qPCR). Only patients associated with a cycle threshold < 28 were included for genomic analyses. Samples consisted of skin, pharyngeal or anal swabs.

Nucleic acid extraction and mpox q-PCR
After a heat inactivation step (12 minutes at 70°C), nucleic acids (500μL of samples) were extracted using MagNA Pure LC 2.0 Instrument (Roche, Meylan, France) and eluted in 100μL. Mpox virus specific real-time PCR assay (qPCR) validated by both Centers for Disease Control & Prevention (CDC) and French Orthopoxvirus national reference center was performed. An exogenous internal extraction and amplification control (IEAC) was added to each sample before extraction (Simplexa Extraction and Amplification Control Set, DiaSorin, Saluggia, Italy). Negative controls were added to each extraction batch and a positive control (DNA extracted from a MXPV cell culture given by the French Orthopoxvirus national reference center) was tested in each PCR run.

Design of primers and PCR protocol
Primers were designed after alignment of 311 sequences from clades 2 and 3 (published up to 29th June 2022) and sequences published by Isidro and colleagues, using MAFFT 7 online software. Twelve amplicons (~30,000 bp) were realized to detect 26/35 of the characteristic mutations which differentiated 2022 outbreak clade 3 and pre-epidemic clade 3 2018 UK P2 (GenBank: MT903344.1). The nine uncovered mutations were all synonymous or non-coding.

Sequencing of mpox fragments
After Qubit quantification using Qubit® dsDNA HS Assay Kit and Qubit 4.0 fluorometer (Thermo Fisher) amplicons were fragmented (Bioruptor, Diagenode) into fragments of 250 bp long. Libraries were built adding barcode, for sample identification, and primers to fragmented DNA with Ion Plus Fragment Library Kit using AB Library Builder System (Thermo Fisher Scientific). To pool equimolarly the barcoded samples, a quantification step by real time PCR using Ion Library TaqMan Quantitation Kit (Thermo Fisher) was realized. An emulsion PCR of the pools and loading on 530 Chip were performed using the automated Ion Chef instrument (Thermo Fisher). Sequencing was performed using the S5 XL Ion torrent technology (Thermo Fisher) following manufacturer’s instructions. Consensus sequence was obtained after trimming of reads (reads with quality score < 0.99, and length < 100 bp were removed and the 30 first and 30 last nucleotides were removed from the reads) and mapping on the MT903344.1 mpox reference sequence using CLC genomics workbench software v.21.0.5 (Qiagen). Parameters for reference-based assembly consisted of match score = 1, mismatch cost = 2, length fraction = 0.5, similarity fraction = 0.8, insertion cost = 3, and deletion cost = 3. A de novo contig was also produced to ensure that the consensus sequence was not affected by the reference sequence.

Phylogenetic analysis and genetic diversity
A phylogenetic tree was produced with MEGA 6 software using the maximum likelihood method (options: Tamura-Nei 93 + G model, 1,000 replicates). For ease of representation, only one representative sequence was for each pattern of mutation. The sequences produced were compared to pre-epidemic and epidemic sequences retrieved on GenBank (Supplementary Methods S1).

Results

Patient characteristics

A total of 148 patients were positive to mpox by qPCR (ct < 28). Available medical records showed that 95% were men and 5% transgender (M-to-F). Ninety percent of patients were MSM, of whom 50% were taking HIV pre-exposure prophylaxis, and 25% lived with HIV. A quarter of the patients had travelled within 3 weeks prior to mpox diagnosis (>85% to Europe, none to Africa). Only, 5% of the patients were hospitalized, mostly for severe pharyngeal/tonsils.

Genetic variability of mpox sequences

A total of 162 sequences (SRA, BioProject: PRJNA944928) were produced from 148 patients (some patients had two samples). Among them, 32 different mutational patterns were distinguished – defined in comparison to the ON602722.2 reference sequence. Pattern #1 was identical to ON602722.2, while patterns #2 to #32 harbored at least one mutation (Supplementary Table S2). Pattern #1 included most of the sequences (n=120, 74.1%), and pattern #9 encompassed 10 sequences (6.2%). All other mutational patterns included only one or two sequences. Fourteen patterns were already reported in the GenBank database. The most common ones are patterns #9 and 13 observed in 6.16% and 1% of the GenBank sequences, respectively (Supplementary Table S2). Others were observed in less than 1% of GenBank sequences. Interestingly, all of the sequences produced in this study included a 17-nucleotide deletion (position 150,621, intergenic). This deletion was a succession of 17-nucleotides repeated sequence and was present in only 30% of the 5,661 GenBank mpox sequences (15th March 2023), and 32% of 5,306 described during the 2022 outbreak. One sequence included an unreported 4 nt-deletion, pattern #8. One sequence included an unreported 4 nt-deletion (position 14,509, ORF OPG25). Overall, 51% of mutations were non-synonymous, 37% synonymous, and 12% non-coding. A majority of transitions (17 G-to-A; 17 C-to-T) was observed, confirming a highly specific mutational topology.

Of note, a comparison of pre-epidemic strain MT903344.1 and the French MPXV_FRA_2022_TLS67 outbreak strain (ON602722.2) identified more mutations (58 substitutions and 27 indels along the genome) (Supplementary Table S3).

Phylogenetic reconstruction

The 162 mpox sequences of this study were compared with representative sequences of clades. Phylogenetic reconstruction was realized using maximum Likelihood method and, Tamura-Nei 93 + G model, 1,000 replicates; (Supplementary Method S1). All patterns from this study, were closely related to other 2022 epidemic strains (ON745225_Spain_2022, ON637938_Germany_2022; Figure 1). However, for one patient returning from Asia, pattern #32 was identified, closely related to clade 3 pre-epidemic and USA-2022_FL001 (ON674051) strains. This suggests a possible low-level circulation of non-epidemic mpox virus in Western countries, probably underdiagnosed and detected here due to attention generated by the current outbreak.

Discussion

To conclude, this rapid parsimonious mpox sequencing strategy in 148 mpox-infected patients provides a unique insight into the genomic variability of an epidemic DNA virus. The existence among the mpox strains identified in France of a mutational pattern specific to the 2022 epidemic was confirmed, and showed the overall low genomic variability of circulating strains. The fact that the majority of the substitutions were transitions (G>A and C>T) suggests a role of the human APOBEC3 enzyme editing in mpox sequences, known to target G and C in specific nucleotide contexts for other viruses as an innate immune response.
This study also highlights that, although rare, mpox can circulate in Europe far from epidemics and should be included in differential diagnosis of other sexually transmitted infections in particular in MSMs.

References


Figure 1. Phylogeny of mpox based on the analysis of selected 30,000 nucleotides from 80 sequences. The tree was resolved by maximum likelihood using the TN93 + G model of evolution. Bootstrap values are
indicated at nodes. The blue circle in front of an identifier indicates the sequences resulting of this study.

This work was presented at the Conference of retroviruses and Opportunistic Infections (19th to 22th February 2023), abstract ndeg: 0236

Funding: This work was supported by ANRS MIE, Project APP PRI 22275, Emergen-2021

ACKNOWLEDGMENTS This work was supported by the Agence Nationale de la Recherche sur le SIDA et les Maladies Infectieuses Emergentes (ANRS MIE), Medical Virology and pharmacology Network