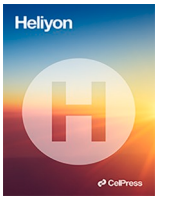
Heliyon 9 (2023) e12895

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Phylogenetic analysis of all available monkeypox virus strains shows the close relatedness of contemporary ones

Mária Benkő, Balázs Harrach, Győző L. Kaján *

Veterinary Medical Research Institute, Hungária krt. 21, 1143 Budapest, Hungary

ARTICLE INFO

Keywords:
Monkeypox
Phylogenetic analysis
RefSeq

ABSTRACT

The present research aimed to evaluate the diversity of all monkeypox virus strains with a special focus on recently isolated ones by a comprehensive phylogenetic analysis of all available sequences, based on the concatenate of four viral genes. Almost all current strains from 2022 showed a high level of similarity to each other on the analyzed stretches: 218 strains shared identical sequence. Among all analyzed strains, the highest number of differences was counted compared to a RefSeq strain (Zaire-96-I-16) on the whole concatenate. Our analysis supported the distinction between Clade I (formerly Congo Basin clade), IIa and IIb (together formerly West African clade) strains and classified all 2022 strains in the last one. The high number of differences and long branch observable concerning strain Zaire-96-I-16 is most probably caused by a sequencing error. As this strain represents one of the two available reference sequences in GenBank, it is recommendable to confirm or exclude the concerning mutation. The developed method, based on four gene sequences, reflected the established whole-genome-based intraspecies classification. Although this method provides significantly less information about the strains compared to whole genome analyses, since its resolution is much lower, it still enables the rapid subspecies classification of the strains into the established clades. The genes in the analyzed concatenate are so conserved that further differentiation of contemporary strains is impossible; these strains are identical in the analyzed sections. On the other hand, since whole genome analyses are compute-intensive, the described method offers a simpler and more accessible alternative for monitoring and preliminary typing of newly sequenced monkeypox virus strains.

1. Introduction

Monkeypox virus is closely related to variola virus and is classified in the genus *Orthopoxvirus*, subfamily *Chordopoxvirinae*. Its original host species is yet uncertain, but African squirrels and monkeys harbor the virus [1]. In humans, the disease is similar to smallpox but much milder – besides the rash, general viral symptoms are typical, like headache, fever, swollen lymph nodes and muscle aches – and the case fatality rate of the 2022 outbreak is about 0.04% [2]. The present research aimed to evaluate the diversity of all monkeypox virus strains with a special focus on recently isolated ones by a comprehensive phylogenetic analysis of all available sequences.

* Corresponding author.
E-mail address: kajan.gyozo@vmri.hu (G.L. Kaján).

https://doi.org/10.1016/j.heliyon.2023.e12895
Received 27 July 2022; Received in revised form 4 January 2023; Accepted 6 January 2023
Available online 10 January 2023

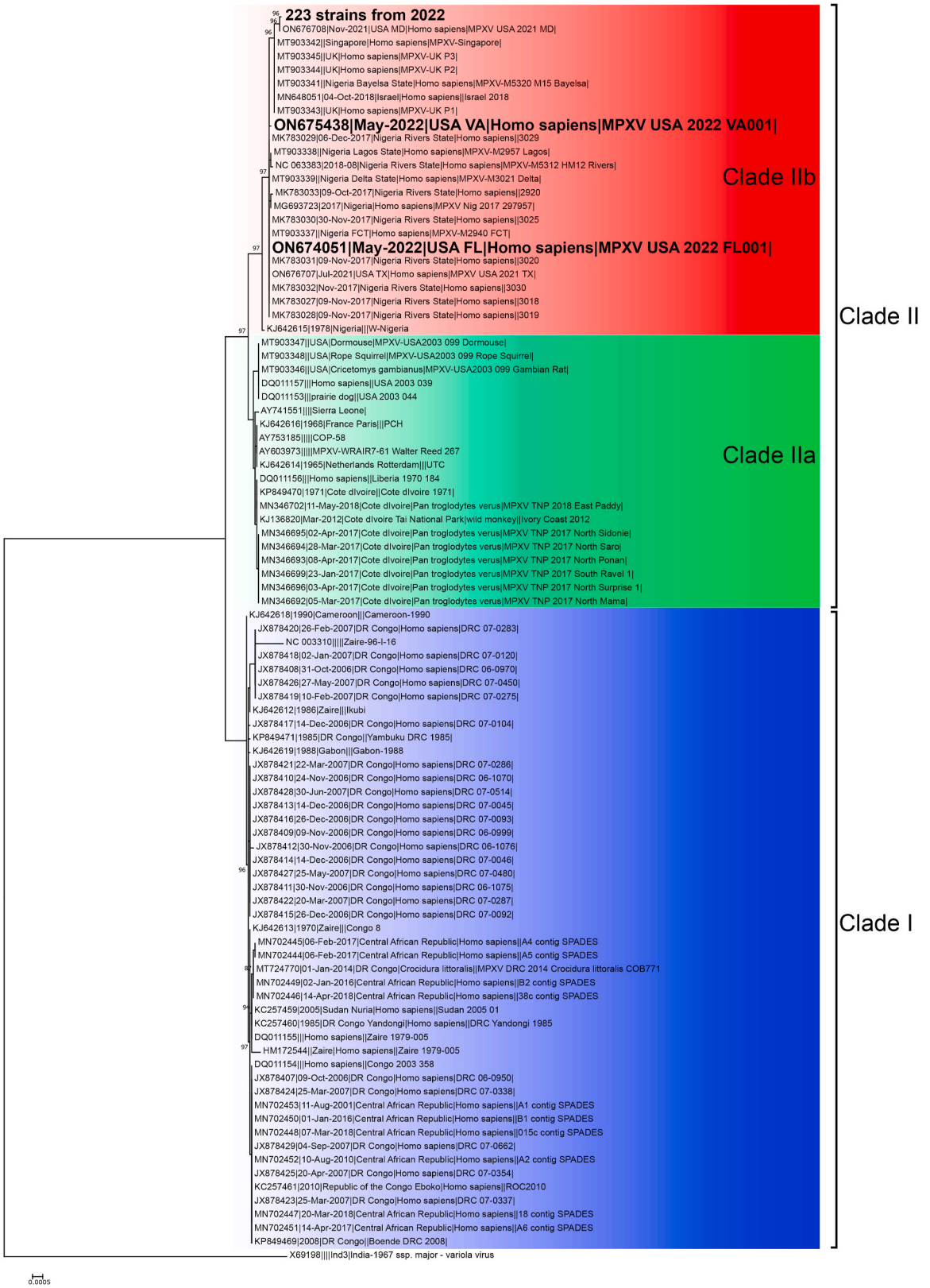


Fig. 1. Phylogenetic analysis of all available monkeypox virus strains. The tree was based on the concatenate of four gene alignments: early transcription factor (E6R), DNA-directed RNA polymerase (A25R), RNA polymerase-associated transcription-specificity factor (H4L), and DNA-

dependent RNA polymerase gene (L6R). The alignment length was 11,658 nucleotides, and the tree was rooted using the variola virus. If available, each strain is represented by the nucleotide accession number, collection date, country, host, isolate and strain name divided by vertical bars. Strains from the year 2022 are highlighted.

2. Materials and methods

Over the sequence length of 81,000 bp, all monkeypox virus (txid10244) nucleotide sequences were downloaded from the NCBI GenBank on July 5, 2022 (n = 402). Phylogenetic analysis was based on four viral genes, recommended by Yu et al. [3], as for the family *Poxviridae*, the comparison between the analysis of these four-genes-concatenates and that of whole genomic amino acid sequences or poxvirus core genes conveyed similar results; and the results of the concatenate phylogenetic analysis has mirrored the classification of the family *Poxviridae* accurately. Viral sequences were annotated according to strain Zaire-96-I-16 (NC_003310) using VAPiD 1.2 [4,5], and the early transcription factor (E6R, GeneID: 928933), DNA-directed RNA polymerase (A25R, 928976), RNA polymerase-associated transcription-specificity factor (H4L, 928896), and DNA-dependent RNA polymerase gene (L6R, 929036) nucleotide sequences were extracted from these using Geneious. Homologous variola virus (X69198) sequences were also included in the analysis. The gene sequences were translation aligned using the MAFFT G-INS-i algorithm, and alignments were concatenated in Geneious. Sequences containing ambiguous nucleotides or not containing all four genes were excluded (final n = 313). Evolutionary model selection was conducted using ModelTest-NG 0.1.6 [6], and the GTR + I model was applied in the phylogenetic tree reconstruction. Here, RAxML-NG 1.1.0 was used [7], and the best tree was chosen from 300 reconstructions with thousand bootstrap replicates, using transfer bootstrap expectations [8]. The tree was visualized in MEGA 7 [9] and rooted using the variola virus. The tree is depicted in Fig. 1.

3. Results

Almost all current strains from 2022 (n = 225) showed a high level of similarity to each other: 218 strains shared identical sequence, whereas further five strains differed in a single nucleotide on the analyzed stretches. Three of these five mutations were silent. The number of differences was counted among all analyzed strains on the complete concatenate, and the highest count, 59, was measured compared to strain Zaire-96-I-16 (RefSeq: NC_003310). The tree was divided into two major clades, and the second of these could be divided further into two subclades. The first major clade was dominated by African strains from 2006 to 2007 and 2016–2018, whereas the second by contemporary strains and African isolates from 2017 to 18.

4. Discussion

Poxviruses are double-stranded DNA viruses with a low genomic variability; thus, the low diversity of the contemporary strains is reasonable. Our analysis supported the distinction between Clade I (formerly Congo Basin clade), IIa and IIb (together formerly West African clade) strains [10,11] and classified all 2022 strains in the last one. The high number of differences and long branch observable concerning strain Zaire-96-I-16 is most probably caused by a sequencing error: in its L6R gene (929036) a homopolymer adenine triplet is observable (NC_003310: nt 84,241–84,243) instead of the common quadruplet, and this causes the shift of the next nine nucleotides, where an adenine is inserted. As this strain represents one of the two available reference sequences in GenBank, and this gene is optional for phylogenetic analyses, it is recommendable to confirm or exclude this mutation.

The developed method, based on four gene sequences, reflected the established whole-genome-based intraspecies classification [11-13]. Although this method provides significantly less information about the strains compared to whole genome analyses, since its resolution is much lower, it still enables the rapid subspecies classification of the strains into the established clades. The genes in the analyzed concatenate are so conserved that further differentiation of contemporary strains is impossible; these strains are identical in the analyzed sections. Higher resolution and further differentiation requires whole genome sequencing [11]. On the other hand, since whole genome analyses are compute-intensive, the described method offers a simpler and more accessible alternative for monitoring and preliminary typing of newly sequenced monkeypox virus strains.

Author contribution statement

Mária Benkő: Conceived and designed the experiments; Analyzed and interpreted the data; Wrote the paper.

Balázs Harrach: Conceived and designed the experiments; Performed the experiments; Contributed reagents, materials, analysis tools or data.

Győző László Kaján: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Wrote the paper.

Funding statement

Balázs Harrach was supported by National Research, Development and Innovation Office [NN140356].

Győző László Kaján was supported by Magyar Tudományos Akadémia [János Bolyai Research Scholarship].

Data availability statement

Data included in article/supp. material/referenced in article.

Declaration of interest's statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Phylogenetic calculations were performed using computing resources provided by KIFÜ, Hungary. The research of GLK is supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

References

[1] A.I. Kabuga, M.E. el Zowalaty, A review of the monkeypox virus and a recent outbreak of skin rash disease in Nigeria, J. Med. Virol. 91 (2019) 533–540, https://doi.org/10.1002/jmv.25348.
[2] M. Kozlov, How deadly is monkeypox? What scientists know, Nature 609 (2022) 663, https://doi.org/10.1038/D41586-022-02931-1.
[3] Z. Yu, W. Zhang, H. Fu, X. Zou, M. Zhao, S. Liang, C. Gu, Q. Yang, M. He, Q. Xiao, W. Xiao, L. He, M. Lü, Genomic analysis of *Poxviridae* and exploring qualified gene sequences for phylogenetics, Comput. Struct. Biotechnol. J. 19 (2021) 5479–5486, https://doi.org/10.1016/J.CSBJ.2021.09.031.
[4] S.N. Shchelkunov, A.V. Totmenin, I.V. Babkin, P.F. Safronov, O.I. Ryazankina, N.A. Petrov, V.V. Gutorov, E.A. Uvarova, M.V. Mikheev, J.R. Sisler, J.J. Esposito, P.B. Jahrling, B. Moss, L.S. Sandakhchiev, Human monkeypox and smallpox viruses: genomic comparison, FEBS Lett. 509 (2001) 66–70, https://doi.org/10.1016/s0014-5793(01)03144-1.
[5] R.C. Shean, N. Makhsous, G.D. Stoddard, M.J. Lin, A.L. Greninger, VAPiD: a lightweight cross-platform viral annotation pipeline and identification tool to facilitate virus genome submissions to NCBI GenBank, BMC Bioinf. 20 (2019) 1–8, https://doi.org/10.1186/S12859-019-2606-Y.
[6] D. Darriba, D. Posada, A.M. Kozlov, A. Stamatakis, B. Morel, T. Flouri, ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models, Mol. Biol. Evol. 37 (2020) 291–294, https://doi.org/10.1093/molbev/msz189.
[7] A.M. Kozlov, D. Darriba, T. Flouri, B. Morel, A. Stamatakis, RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference, Bioinformatics 35 (2019) 4453–4455, https://doi.org/10.1093/bioinformatics/btz305.
[8] F. Lemoine, J.-B. Domelevo Entfellner, E. Wilkinson, D. Correia, M. Dávila Felipe, T. de Oliveira, O. Gascuel, Renewing Felsenstein's phylogenetic bootstrap in the era of big data, Nature 556 (2018) 452–456, https://doi.org/10.1038/s41586-018-0043-0.
[9] S. Kumar, G. Stecher, K. Tamura, MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets, Mol. Biol. Evol. 33 (2016) 1870–1874, https://doi.org/10.1093/molbev/msw054.
[10] N. Chen, G. Li, M.K. Liszewski, J.P. Atkinson, P.B. Jahrling, Z. Feng, J. Schriewer, C. Buck, C. Wang, E.J. Lefkowitz, J.J. Esposito, T. Harms, I.K. Damon, R. L. Roper, C. Upton, R.M.L. Buller, Virulence differences between monkeypox virus isolates from West Africa and the Congo basin, Virology 340 (2005) 46–63, https://doi.org/10.1016/J.VIROL.2005.05.030.
[11] C. Happi, I. Adetifa, P. Mbala, R. Njouom, E. Nakoune, A. Happi, N. Ndodo, O. Ayansola, G. Mboowa, T. Bedford, R.A. Neher, C. Roemer, E. Hodcroft, H. Tegally, Á. O'Toole, A. Rambaut, O. Pybus, M.U.G. Kraemer, E. Wilkinson, J. Isidro, V. Borges, M. Pinto, J.P. Gomes, L. Freitas, P.C. Resende, R.T.C. Lee, S. Maurer-Stroh, C. Baxter, R. Lessells, A.E. Ogwell, Y. Kebede, S.K. Tessema, T. de Oliveira, Urgent need for a non-discriminatory and non-stigmatizing nomenclature for monkeypox virus, PLoS Biol. 20 (2022), e3001769, https://doi.org/10.1371/JOURNAL.PBIO.3001769.
[12] L. Wang, J. Shang, S. Weng, S.R. Aliyari, C. Ji, G. Cheng, A. Wu, Genomic annotation and molecular evolution of monkeypox virus outbreak in 2022, J. Med. Virol. 95 (2023), e28036, https://doi.org/10.1002/JMV.28036.
[13] M. Vandenbogaert, A. Kwasiborski, E. Gonofio, S. Descorps-Declère, B. Selekon, A.A. Nkili Meyong, R.S. Ouilibona, A. Gessain, J.C. Manuguerra, V. Caro, E. Nakoune, N. Berthet, Nanopore sequencing of a monkeypox virus strain isolated from a pustular lesion in the Central African Republic, Sci. Rep. 12 (2022) 1–13, https://doi.org/10.1038/s41598-022-15073-1.