















Letter to the editor:

ARTIFICIAL INTELLIGENCE (AI) TOOLS TO INVESTIGATE VIRULENCE AND EVALUATE THE INFECTIVITY OF MPOX

Rajiv Gandhi Gopalsamy¹, Athesh Kumaraswamy², Marina dos Santos Barreto^{3,*}, Ronaldy Santana Santos⁴, Eloia Emanuely Dias Silva⁵, Deise Maria Rego Rodrigues Silva³, Pedro Henrique Macedo Moura³, Pamela Chaves de Jesus³, Jessiane Bispo de Souza³, Adriana Kelly Santana Correa³, Allec Yuri Santos Martins³, Anne Gabriela de Freitas Almeida³, Lucas Alves da Mota Santana⁶, Lysandro Pinto Borges^{3,*}

¹ Division of Phytochemistry and Drug Design, Rajagiri College of Social Sciences (Autonomous), Kalamassery, Kochi, Kerala 683104, India

² School of Sciences, Bharata Mata College (Autonomous), Thrikkakara, Kochi, Kerala 682021, India

³ Department of Pharmacy, Federal University of Sergipe, São Cristóvão 49100-000, Sergipe, Brazil

⁴ University of São Paulo, Postgraduate Pharmacy Program (Pathophysiology and Toxicology), São Paulo 05508-000, São Paulo, Brazil

⁵ Department of Biology, Federal University of Sergipe, São Cristóvão 49100-000, Sergipe, Brazil

⁶ Graduate Program in Dentistry, Federal University of Sergipe, São Cristóvão 49100-000, Sergipe, Brazil

* **Corresponding authors:** Marina dos Santos Barreto and Lysandro Pinto Borges
Department of Pharmacy, Federal University of Sergipe, São Cristóvão 49100-000, Sergipe, Brazil. E-mails: sbarretomarina@outlook.com, lysandro.borges@gmail.com

<https://dx.doi.org/10.17179/excli2025-8129>

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>).

The zoonotic Mpox (formally known as Monkeypox) virus, from the same family as the smallpox virus, is a threat to global public health. This virus mainly affects animals but can be transmitted to humans through contact with infected individuals or animals. Recent global outbreaks and the rapid mutation rate of the virus raise concerns about its potential virulence and transmissibility (Martinez-Fernandez et al., 2023; Imani et al., 2024). Advanced genomic and bioinformatics tools can generate essential information on the evolution of the virus, as well as alerting us to transmission patterns and pathogenic potential (Nnaemeka et al., 2024). This knowledge is essential for developing effective strategies to prevent and control this pathology.

Next-generation bioinformatics tools of artificial intelligence (AI) allow researchers to dissect the intricacies of the Mpox genome. With the weighted gene co-expression network analysis (WGCNA), this enables *in silico* investigation of co-expression networks in various diseases, elucidating the molecular mechanisms of viral diseases and guiding the development of targeted therapies for them (Alarabi et al., 2022). WGCNA and host-pathogen interaction

(HPI) analysis helped to explore the complex interplay between the Mpox virus and the human host. WGCNA identified key gene modules significantly impacted by the virus, while HPI analysis revealed how viral proteins manipulate host cellular processes (Tang et al., 2024). The integrated results of the WGCNA and HPI tools highlighted the identification of possible antiviral therapeutic targets, such as kinase inhibitors and hormone-based therapies, which could modulate key signaling pathways such as nuclear factor kappa B (NF- κ B), interferons (IFNs), phosphatidylinositol-3 kinase (PI3K), and signal transducer and activator of transcription 3 (STAT3) (Imani et al., 2024).

MATLAB is another powerful AI platform revolutionizing genomic research, machine learning (ML) and deep learning. AI makes it possible to analyze large data sets efficiently, visualizing complex genomic data. This allows researchers to analyze patterns and understand the pathological mechanism of Mpox and other viruses (MathWorks, 2024). In addition to these tools, Nextclade is a bioinformatics tool vital for Mpox surveillance. Analyzing viral genomes classifies strains into clades (e.g., Clade IIb) and lineages (e.g., B.1), identifying mutations and tracking their evolution. This information aids in understanding genetic diversity and pinpointing the emergence and spread of distinct viral subtypes. While not directly assessing virulence or infectivity, Nextclade indirectly contributes by identifying mutations linked to specific lineages, potentially associated with changes in disease characteristics, such as signs and symptoms. The tool's output seamlessly integrates with Nextstrain for generating phylogenetic trees, visualizing evolutionary relationships, and informing public health strategies by tracking transmission patterns and identifying potential hotspots of viral evolution (Bosmeny et al., 2023; Choi et al., 2024; Aksamentov et al., 2021). In addition, MpoxRadar is a genomic surveillance tool created specifically for Mpox. MpoxRadar allows researchers and public health authorities to obtain information about the disease. It allows rapid analysis of virus variants, even down to mutations, providing information for public health. Its interface is easy to use, with customizable filters and intuitive visualizations, making it possible to access and interpret complex genomic data (Nasri et al., 2023).

Public databases such as NCBI, GISAID and virological.org provide a wealth of viral sequences, while BLAST facilitates in-depth sequence comparisons to identify potentially virulent factors. Parsnp and Muscle allow robust alignments of multiple sequences, revealing patterns of genetic variation between strains. Powerful visualization tools such as Jalview and MEGA reveal evolutionary relationships and potential drug targets. Analyzing the structure of proteins using PDB's web tools, exemplified by the A42R protein (PDB entry 4QWO), makes it possible to identify structural changes that can influence infectivity. Therefore, several existing tools can help in the fight against Mpox (RCSB PDB, 2014). Differential expression analysis using limma and pathway enrichment analysis using GSEA with MSigDB collections (KEGG, WikiPathways, Gene Ontology) reveal virus-host interactions and potential virulence mechanisms. The R statistical software with the *corto* package allows the creation of informative graphs to effectively communicate these findings (Giorgi et al., 2024).

Rigorous quality control and assembly of raw sequencing data are facilitated by tools like FastQC, Fastp, Flye, and SPADES. Reference-based assembly and variant calling are streamlined using Bowtie2, Pilon, iVar, LoFreq and BCFtools. Variant annotation and impact assessment are achieved with SnpEff and SnpSift. Phylogenetic relationships are elucidated through IQ-Tree 2 and visualized with iTOL. PopArt enables exploration of genetic diversity through haplotype networks. MEGA_X_10.0.2 constructs reference sequences for subsequent analyses. Furthermore, tools like SWISS-MODEL facilitate the visualization and comparison of protein conformation models, providing crucial insights into structural variations (Zhu et al., 2023). CheckV v1.0.1, utilizing the checkv-db v1.2 databases, assesses the quality of each Mpox genome assembly by comparing them to reference genomes using a diamond database.

This method provides an estimate of genome completeness on a scale of 0-100 %, ensuring the reliability of subsequent analyses by eliminating incomplete sequences (Agarwal et al., 2024). Kraken2 and NCBI SRA human-scrubber remove human reads. Minimap2, BWA-MEM, and Bowtie2 align Illumina unclassified reads to the MPXV genome. PICARD generates duplicate metrics, while Samtools and mosdepth provide coverage metrics. Samtools and iVar generate a consensus sequence. Unicycler produces a hybrid de novo assembly visualized with Bandage and refined with Kraken2. Samtools, BCFtools, and seqtk generate the consensus of this assembly. QUAST compares the consensus sequences against the reference genome (Munoz-Barrera et al., 2023). AGAT, Geneious Prime, and MetaLogo provide valuable insights into functional annotation and the visualization of sequence motifs within specific genomic regions (Monzon et al., 2024).

Advanced bioinformatics techniques are crucial for understanding the Mpox virus and developing effective control strategies. By analyzing the virus's genome, researchers can gain insights into its evolution, transmission patterns, and virulence. Tools like WGCNA, HPI analysis, and MATLAB enable the identification of key genes, host-pathogen interactions, and potential therapeutic targets. Nextclade facilitates viral classification and mutation tracking, aiding in surveillance efforts. A wide array of software tools, from sequence alignment and phylogenetic analysis to protein structure prediction and differential expression analysis, empowers researchers to comprehensively investigate the Mpox virulence and develop effective strategies for containment and mitigation of this virus.

This integrated approach using AI has led to the development of regulatory networks involving transcription factors and protein kinases, making it possible to understand the mechanisms responsible for viral evasion of the immune system and disease progression (Loganathan et al., 2024). With the AI tool and ML algorithms, it is possible to predict viral mutations, model evolutionary pathways and assess the impact of these changes on transmissibility. This is important for responding to emerging infectious diseases such as Mpox, as it allows for the development of effective and agile targeted interventions, minimizing the impacts of these endemics. This integration of AI and ML presents an important advance in the understanding of Mpox and other infectious diseases, making MATLAB an essential tool for future scientific research (Nnaemeka et al., 2024). In conclusion, the application of advanced AI techniques is instrumental in unraveling the complexities of the Mpox virus. By leveraging advanced sequencing analyses using these tools to unravel the genetic intricacies of Mpox, we can pave the way for optimized research, enabling more effective strategies for containment and mitigation.

Conflict of interest

The authors declare no conflicts of interest related to this research.

Using artificial intelligence (AI)

AI was used for writing assistance in this manuscript. The authors take full responsibility for the content.

REFERENCES

- Agarwal P, Shukla N, Bhatia A, Mahfooz S, Narayan, J. Comparative genome analysis reveals driving forces behind monkeypox virus evolution and sheds light on the role of ATC trinucleotide motif. *Virus Evol.* 2024; 10(1):veae043.
- Aksamentov I, Roemer C, Hodcroft EB, Neher RA. Nextclade: Clade assignment, mutation calling and quality control for viral genomes. *J Open Source Softw.* 2021;6(67):3773.

- Alarabi AB, Mohsen A, Mizuguchi K, Alshbool FZ, Khasawneh FT. Co-expression analysis to identify key modules and hub genes associated with COVID-19 in platelets. *BMC Med Genomics*. 2022;15(1):83.
- Bosmeny MS, White AA, Pater AA, Crew J, Geltz J, Gagnon KT. Global mpox lineage discovery and rapid outbreak tracking with nanopore sequencing. *Virology*. 2023;20(1):90.
- Choi CH, Lee M, Lee SE, Kim JW, Shin H, Choi MM, et al. Genomic analysis of monkeypox virus during the 2023 epidemic in Korea. *J Korean Med Sci*. 2024;39(18):e165.
- Giorgi FM, Pozzobon D, Di Meglio A, Mercatelli D. Genomic and transcriptomic analysis of the recent Mpox outbreak. *Vaccine*. 2024;42:1841–9.
- Imani S, Aminnezhad S, Alikarami M, Abedi Z, Mosleh IS, Maghsoudloo M et al. Exploration of drug repurposing for Mpox outbreaks targeting gene signatures and host-pathogen interactions. *Sci Rep*. 2024; 14(1):29436.
- Loganathan T, Fletcher J, Abraham P, Kannangai R, Chakraborty C, El Allali A, et al. Expression analysis and mapping of viral-host protein interactions of Poxviridae suggests a lead candidate molecule targeting Mpox. *BMC Infect Dis*. 2024;24(1):483.
- Martinez-Fernandez DE, Fernandez-Quezada D, Casillas-Munoz FAG, Carrillo-Ballesteros FJ, Ortega-Prieto AM, Jimenez-Guardeno JM et al. Human monkeypox: A comprehensive overview of epidemiology, pathogenesis, diagnosis, treatment, and prevention strategies. *Pathogens*. 2023;12(7):947.
- MathWorks. MATLAB. Available from: <https://www.mathworks.com/products/matlab.html>. Accessed August 16, 2024.
- Monzon S, Varona S, Negredo A, Vidal-Freire S, Patino-Galindo JA, Ferressini-Gerpe N, et al. Monkeypox virus genomic accordion strategies. *Nat Commun*. 2024;15(1):3059.
- Munoz-Barrera A, Ciuffreda L, Alcoba-Florez J, Rubio-Rodriguez LA, Rodriguez-Perez H, Gil-Campesino H, et al. Bioinformatic approaches to draft the viral genome sequence of Canary Islands cases related to the multicountry mpox virus 2022-outbreak. *Comput Struct Biotechnol J*. 2023;21:2197–203.
- Nasri F, Kongkitimanon K, Wittig A, Cortés JS, Brinkmann A, Nitsche A, et al. MpoxRadar: a worldwide MPXV genomic surveillance dashboard. *Nucleic Acids Res*. 2023;51(W1):W331–7.
- Nnaemeka J, Oladokun P, Abubakar I, Afolabi S. Leveraging AI and deep learning in predictive genomics for MPOX virus research using MATLAB. *Int J Computer Applications Technol Res*. 2024;13(9):1-13.
- RCSB PDB. 4QWO: 1.52 Angstrom Crystal Structure of A42R Profilin-like Protein from Monkeypox Virus Zaire-96-I-16. 2014. Available online: <https://www.rcsb.org/structure/4QWO> [accessed on 31 May 2022].
- Tang Z, Han Y, Meng Y, Li J, Qiu X, Bajinka O, et al. A bioinformatics approach to systematically analyze the molecular patterns of monkeypox virus-host cell interactions. *Heliyon*. 2024;10(9):e30483.
- Zhu J, Yu J, Qin H, Chen X, Wu C, Hong X, et al. Exploring the key genomic variation in monkeypox virus during the 2022 outbreak. *BMC Genom Data*. 2023; 24(1):67.
-