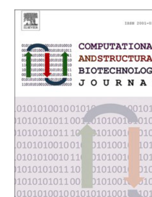
Computational and Structural Biotechnology Journal 21 (2023) 2197–2203

Contents lists available at ScienceDirect

Computational and Structural Biotechnology Journal

journal homepage: www.elsevier.com/locate/csbj



Research article

Bioinformatic approaches to draft the viral genome sequence of Canary Islands cases related to the multicountry mpox virus 2022-outbreak

Adrián Muñoz-Barrera [a,1], Laura Ciuffreda [b,1], Julia Alcoba-Florez [c,1], Luis A. Rubio-Rodríguez [a], Héctor Rodríguez-Pérez [b], Helena Gil-Campesino [c], Diego García-Martínez de Artola [c], Josmar Salas-Hernández [b], Julia Rodríguez-Núñez [b], Antonio Íñigo-Campos [a], Víctor García-Olivares [a], Oscar Díez-Gil [c], Rafaela González-Montelongo [a], Agustín Valenzuela-Fernández [d], José M. Lorenzo-Salazar [a,2], Carlos Flores [a,b,e,f,*,2]

[a] *Genomics Division, Instituto Tecnológico y de Energías Renovables, 38600 Santa Cruz de Tenerife, Spain*
[b] *Research Unit, Hospital Universitario Ntra. Sra. de Candelaria, 38010 Santa Cruz de Tenerife, Spain*
[c] *Servicio de Microbiología, Hospital Universitario Ntra. Sra. de Candelaria, 38010 Santa Cruz de Tenerife, Spain*
[d] *Laboratorio "Inmunología Celular y Viral", Unidad de Farmacología, Sección de Medicina, Facultad de Ciencias de la Salud, Universidad de La Laguna, 38200 San Cristóbal de La Laguna, Spain*
[e] *CIBER de Enfermedades Respiratorias (CIBERES), Instituto de Salud Carlos III, 28029 Madrid, Spain*
[f] *Facultad de Ciencias de la Salud, Universidad Fernando Pessoa Canarias, 35450 Las Palmas de Gran Canaria, Spain*

ARTICLE INFO

Article history:
Received 16 November 2022
Received in revised form 13 March 2023
Accepted 13 March 2023
Available online 15 March 2023

Keywords:
Bioinformatics
De novo genome assembly
Viral surveillance
Mpox
MPXV

ABSTRACT

On July 23, 2022, monkeypox disease (mpox) was declared a Public Emergency of International Concern (PHEIC) by the World Health Organization (WHO) due to a multicountry outbreak. In Europe, several cases of mpox virus (MPXV) infection related to this outbreak were detected in the Canary Islands (Spain). Here we describe the combination of viral DNA sequencing and bioinformatic approaches, including methods for *de novo* genome assembly and short- and long-read technologies, used to reconstruct the first MPXV genome isolated in the Canary Islands on the 31st of May 2022 from a male adult patient with mild symptoms. The same sequencing and bioinformatic approaches were then validated with three other positive cases of MPXV infection from the same mpox outbreak. We obtained the best results using a reference-based approach with short reads, evidencing 46–79 nucleotide variants against viral sequences from the 2018–2019 mpox outbreak and placing the viral sequences in the new B.1 sublineage of clade IIb of the MPXV classification. This study of MPXV demonstrates the potential of metagenomics sequencing for rapid and precise pathogen identification.

© 2023 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Monkeypox (mpox) virus (MPXV) is a zoonotic Orthopoxvirus (OPV) (family *Poxviridae*) [1,2] endemic to West and Central Africa [3,4] that causes mpox disease [5]. Mpox has been described in humans in Central and Western Africa (occurring mainly in tropical forest areas of Central Africa) as well as in other parts of the world [6–11]. On approximately 23 July 2022, because mpox cases were reported in several non-endemic countries, predominantly in Europe [12–17], the WHO declared community transmission of the virus [18,19]. This global outbreak requires epidemiological surveillance due to newly reported mutations of the virus [20,21]. Most cases reported by the 7th of November 2022 have been presented through sexual health or other health services and have involved mainly men who have sex with men (MSM) [22–25]. Several clinical studies have demonstrated the human transmission of the virus through sexual contact [26–28]. By early June 2022, 129 viral genomes had been deposited into GISAID (https://gisaid.org) [29] with 46 single

* Correspondence to: Unidad de Investigación, Hospital Universitario Nuestra Señora de Candelaria, Carretera del Rosario s/n, 38010 Santa Cruz de Tenerife, Spain.
E-mail address: cflores@ull.edu.es (C. Flores).
[1] Equal contribution as first authors
[2] Equal contribution

https://doi.org/10.1016/j.csbj.2023.03.020
2001-0370/© 2023 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

nucleotide variants (SNPs) shared by all these sequences and differing from the viral genome sequences from the 2018–2019 mpox outbreak [30]. Preliminary data from polymerase chain reaction (PCR) assays indicate that these MPXV strains detected in Europe and other non-endemic areas belong to clade two (II) [31–38].

In Europe, several cases of MPXV infection have been associated with 2022 outbreaks in the Canary Islands and Spain [31]. A few viral sequences from samples collected in Spain have been reported [39]. As of June 2022, a total of 15 positive cases were confirmed in the Canary Islands [40], and this number increased to a total of 176 by early November 2022 [41]. Here, we describe the combination of methodological approaches to obtain the draft sequences of the first MPXV genomes isolated in the Canary Islands on the 31st of May 2022 from a male adult patient with one week-onset mild symptoms (fever, odynophagia) who presented at the emergency room but did not necessitate hospital admission. An expedited description of the case and the resulting draft sequences was publicly posted in mid-June 2022 [12]. Validation of procedures and results was performed using samples from three other positive cases isolated in the Canary Islands from the same mpox outbreak.

2. Materials and methods

2.1. DNA extraction and PCR testing

For the first incident mpox case (MPXV01), viral DNA was extracted at the Hospital Universitario Ntra. Sra. de Candelaria (Santa Cruz de Tenerife, Spain) from five samples (nasopharyngeal swab, lesion crust, and vesicles) from the same patient using the eMAG system (Biomerieux) following the manufacturer's instructions. Virus inactivation was conducted under a biosafety class II cabinet (TELSTAR bio-II-A) following ECDC procedures [42]. The diagnosis of the MPXV infection was confirmed using the LightMix Modular Orthopox (Roche) and a real-time PCR assay described elsewhere [43]. This assay yielded threshold cycle values in the range of 17–33 in these samples. To further validate the sequencing data and bioinformatic analysis results, additional viral DNA samples were obtained from three other patients (MPXV05 to 07) from the same outbreak (Table 1).

2.2. Short- and long-read DNA sequencing

Five independent DNA dual index libraries (one for each sample) were processed from the first positive case at Instituto Tecnológico y de Energías Renovables (ITER) with a Nextera XT DNA Library Preparation Kit (Illumina Inc.), following the manufacturer's recommendations with manual library normalization, and pooled prior to sequencing. The quality of the libraries was assessed with a D1000 ScreenTape kit on the 4200 TapeStation System (Agilent). Library concentrations ranged from 7.4 to 10.4 nM and showed a fragmentation profile ranging from 721 to 808 bp. The mean fragment size for the sequencing pool was 677 bp, as measured with a D1000 High Sensitivity ScreenTape kit (Agilent). Paired-end sequences were obtained on a MiSeq Sequencing System (Illumina Inc.) using reagent kit v3 chemistry with 150 cycles and an expected throughput of 3.3–3.8 Gb. The pool concentration was 15 pM, and 5% of PhiX Control V3 was used as the internal control.

DNA libraries for nanopore sequencing were also prepared from the sample with the highest yield (taken from a skin lesion exudate) from this first mpox case using the Rapid Barcoding kit (SQK-RBK004) from Oxford Nanopore Technologies (ONT). To increase the quantity of the starting material, the protocol used 30–45 ng of the DNA extract in 7.5 μl of reaction to generate 12 independently barcoded libraries that were pooled to obtain the maximum yield from the run. The pooled libraries were loaded onto an R9.4.1 flow cell and run in a MinION (ONT) for 42 h. Basecalling of raw ONT signal data as well as demultiplexing and adapter trimming was carried out using Guppy v.6.0.7 with default parameters and the high-accuracy basecalling model.

The three additional samples from the same outbreak were sequenced with an Illumina Nextera XT DNA Library Preparation Kit (Illumina Inc.) following the same protocol as for the first positive sample. Library concentrations ranged from 8.3 to 10.7 nM and showed a fragmentation profile ranging from 603 to 646 bp. For ONT sequencing, the same Rapid Barcoding kit (SQK-RBK004) was used, and three barcoded libraries for each sample were pooled to maximize the run performance.

2.3. Bioinformatic analyses and assembly comparisons

As the first step, the pair individual demultiplexed FASTQ of Illumina files was interleaved with BBMap (Reformat tool) and then merged into a single interleaved FASTQ file. Then, two different bioinformatic tools were tested to identify and remove the human reads: Kraken2 [44] and NCBI SRA Human Scrubber v.1.0.2021_05_05 (only used for Illumina sequencing data). The remaining Illumina and ONT reads were subjected to different bioinformatic procedures to obtain draft sequences (Fig. 1).

A reference-based analysis was conducted with Illumina unclassified reads that were mapped to the MPXV genome MPXV-UK_P2 (GenBank MT903344.1) by means of three alternative aligners: minimap2 v.2.24-r1122 [45], BWA-MEM v.0.7.17 [46], and Bowtie2 v.2.4.5 [47]. At this stage, duplicate metrics from PICARD v.2.18.7 [48] and coverage metrics from SAMtools v.1.6 [49] and mosdepth v.0.3.3 [50] were obtained from the remaining interleaved paired-end reads. Variant calling was carried out with two alternative algorithms, iVar v1.3.1 [51] and LoFreq v.2.1.5 [52], using default parameters against the MPXV MT903344.1 genome. For downstream analyses, a consensus sequence was obtained by piping a SAMtools v.1.6 pileup with iVar v.1.3.1 consensus as described elsewhere [53].

Additionally, a hybrid *de novo* assembly was obtained by combining the filtered Illumina and ONT reads using custom script based on the Unicycler v.0.5.0 [54] assembler. Bandage v.0.9.0 [55] was used to visualize the resulting contigs in the assembly. A refined version of this hybrid *de novo* assembly was obtained after running Kraken2 v.2.1.2 with the PlusPF database to remove nonviral assembled contigs. The consensus sequence of this assembly was obtained by mapping the resulting contigs to the MT903344.1 genome and piping SAMtools v.1.6 pileup with BCFtools v.1.6 and seqtk v.1.3-r106.

Finally, the two selected consensus sequences (Illumina-only and hybrid *de novo* assembly) were compared against MT903344.1 as the reference genome with QUAST v.5.0.2 [56].

Table 1
Sample information from the four patients analyzed in the study.

Sample	Age	Sex	Collection date (mm-day-year)	Sample type	Ct
MPXV01	36	Male	05–31–2022	Skin lesion	17
				Nasopharyngeal	33
				Skin lesion	21
				Skin lesion	22
				Skin lesion	17
MPXV05	37	Male	07–01–2022	Skin lesion	17
MPXV06	34	Male	07–01–2022	Skin lesion	15
MPXV07	30	Male	07–05–2022	Skin lesion	21

Ct, Cycle threshold.

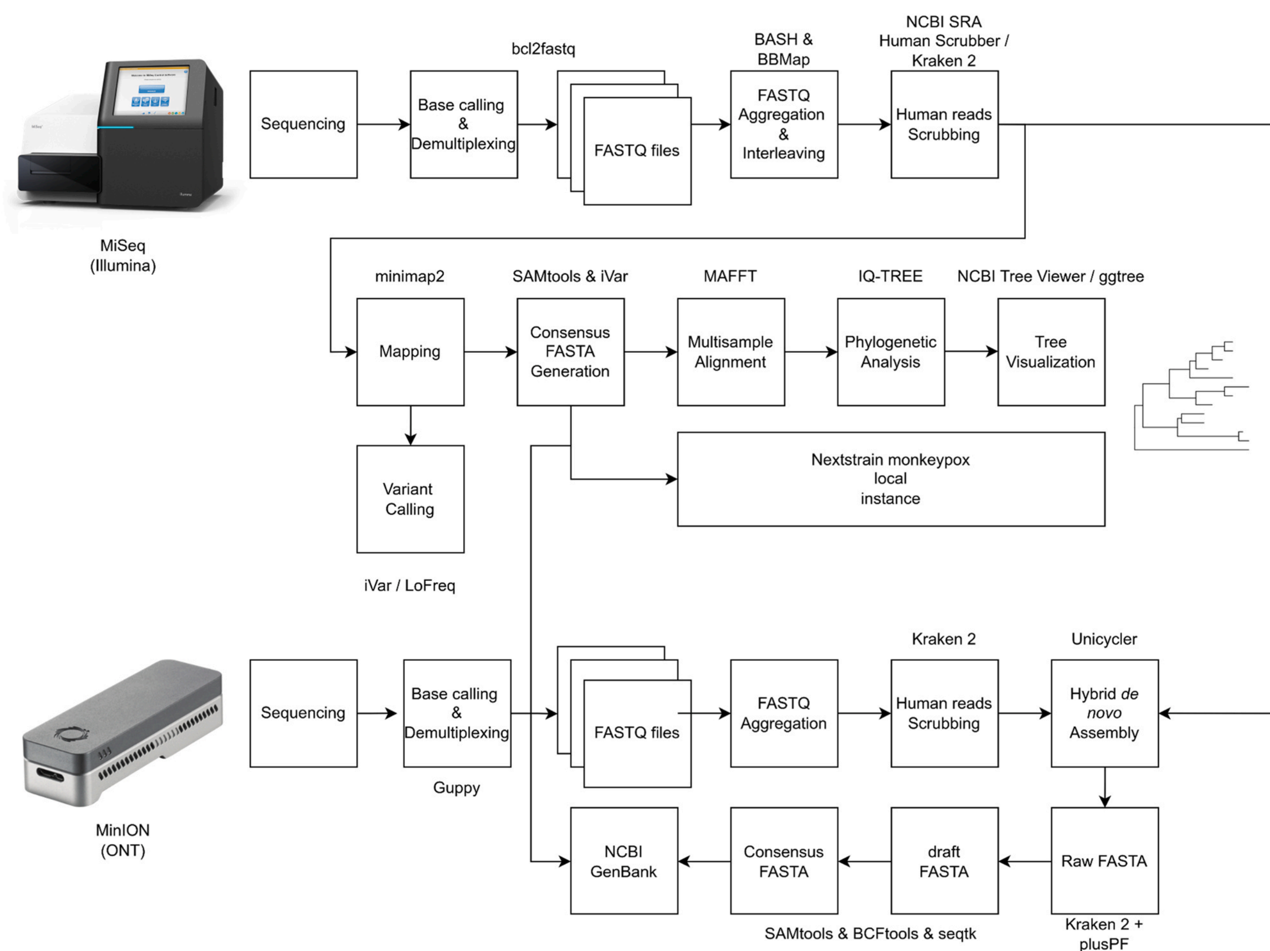


Fig. 1. Full bioinformatic pipeline to obtain the MPXV sequences from Illumina-read only and the hybrid *de novo* assembly to infer phylogenetic relationships with other MPXV genomes available from public repositories.

2.4. Phylogenetic analysis

The most complete consensus sequence resulting from the previous stage was aligned with 126 MPXV sequences downloaded from NCBI GenBank (Table S1) using MAFFT v.7.505 [57]. A phylogenetic analysis was performed using both IQ-TREE v.2.2.0.3 [58] with the K3Pu+F+I model as the best-predicted model and default parameters and using a local instance of Nextstrain [59].

3. Results

The Illumina sequencing run produced 3.88 Gb and 25.5 M reads in total. A mapping of 101,814 and 100,897 Illumina reads was obtained using minimap2 on the NCBI SRA Human Scrubber (mean depth: 38.3×) and Kraken2 (mean depth: 38.1×), respectively, thus providing equivalent results (Table 2). We estimated as few as 2.81% of the reads as duplicates and that 99% of the MPXV genome was covered ≥ 1×, with a fraction of 85% of the viral genome covered at ≥ 10×. The combinations of mapper and variant caller that yielded the smallest and largest number of nucleotide variants against the reference were BWA+LoFreq (46 nucleotide variants) and minimap2 +iVar (67 nucleotide variants), respectively (Table 2). To maintain the maximum sequence variability for downstream analyses, the consensus sequence for Illumina-only reads was obtained for the minimap2 +iVar combination (total size of 197,221 bp), providing a near-fully complete viral genome (99.91%) against the reference.

The ONT run provided 1.98 Gb and a total of 1.38 M reads, ranging from 499 to 101,895 bp in length, with a mean length of 1,432 bases. ONT sequencing provided 2,246 nonhuman mapping reads after filtering with Kraken2, resulting in a theoretical viral genome depth of 14.9×. A hybrid *de novo* assembly based on Illumina and ONT Kraken2-filtered reads was performed and resulted in four contigs (Fig. 2). Contigs 1 and 2 accounted for 186,315 bp and 4,703 bp (191,018 bp total), respectively, and mapped to the MPXV genome Zaire-96-I-16. Contigs 3 and 4 spanned 10,530 bp in total but did not map to the MPXV reference. Thus, a consensus sequence was built from this hybrid *de novo* assembly (including only contigs 1 and 2)

Table 2
Comparative of mapped Illumina reads and coverage using different aligners (Bowtie2, BWA, Minimap2) and called variants using iVar and LoFreq callers.

MPXV01	Kraken2						NCBI SRA Human Scrubber					
Total reads	Non-human Reads	Aligners	Mapped reads	Coverage	iVar	LoFreq	Non-human Reads	Aligners	Mapped reads	Coverage	iVar	LoFreq
51,042,414	4,009,480	Bowtie2	101,092 (2.52%)	38.15	65	47	10	Bowtie2	100,851 (0.96%)	38.34	54	30
		BWA	101,456 (2.53%)	38.25	56	46		BWA	105,484 (1.00%)	38.84	58	46
		Minimap2	100,907 (2.52%)	38.08	67	48		Minimap2	101,838 (0.96%)	38.28	67	48

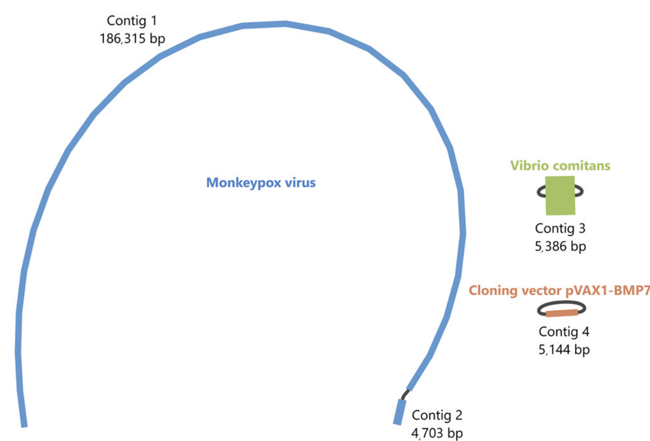


Fig. 2. Bandage representation of the hybrid *de novo* assembly based on long-read sequencing technology. Contigs 1 and 2 (solid blue line) contained MPXV sequences and accounted for a total of 191,018 bp. Contigs 3 (solid green line) and 4 (solid brown line) contained sequences that corresponded to non-viral assembled contigs and that were discarded from further analyses. Whenever necessary, continuity between or within contigs is represented as a thin black line to indicate that there is a gap between them.

Table 3
Assessment of the two MPXV genome assemblies from the first patient (MPXV01) against the MPXV genome reference.

Metrics	Hybrid *de novo* assembly (ON782054)	Illumina-only (ON782055)
# contigs	1	1
Largest contig (bp)	197,222	197,221
Total length (bp)	197,222	197,221
Reference length (bp)	197,233	197,233
GC (%)	32.93	33.02
Reference GC (%)	33.02	33.02
N50	197,222	197,221
L50	1	1
# misassemblies	0	0
Genome fraction (%)	96.75	100.00
Duplication ratio	1.034	1.000
# N's per 100 kbp	3,281.07	92.79
# mismatches per 100 kbp	22.01	25.86
# indels per 100 kbp	0.00	1.52
Largest alignment (bp)	190,825	197,221
Total aligned length (bp)	190,825	197,221
NA50	190,825	197,221
LA50	1	1

and the MT903344.1 MPXV reference genome, spanning a total size of 197,222 bp. It should be noted, however, that because this sequence includes 6,471 undetermined bases, the consensus sequence only covered 96.75% of the reference genome (Table 3).

In addition to the number of undetermined bases, which were much higher in the hybrid *de novo* assembly sequence than in the Illumina-only consensus sequence, we observed small differences overall between the two MPXV genome assemblies obtained (Table 3). In brief, the hybrid *de novo* assembly was able to retrieve a slightly lower proportion of the reference genome, although it had a similar GC content and fewer mismatches than the Illumina-only assembly.

Based on the above findings supporting more completeness for the Illumina-only assembly, we opted to place this consensus sequence in a phylogenetic context (Fig. 3). This analysis indicated that the draft MPXV genome sequence belongs to the new B.1 sublineage of clade II [60–65]. In addition, the closest sequences were related to the Slovenian-Mpox GenBank-released genomes, contributing further evidence of community spread in the present worldwide mpox outbreak.

To validate the utility of these approaches, samples from three other positive patients from the same mpox outbreak detected in the Canary Islands were sequenced with both sequencing technologies following the same laboratory and bioinformatic procedures. All samples offered similar sequencing results to those described for the first positive case in terms of the number of nonhuman sequenced reads, ranging from 48.2 to 307.3 M short reads and 2.4–24.3 M long reads. For the analysis of these data, as part of the Illumina-only strategy, minimap2 and iVar were used for mapping and variant calling, following the selected combination as for the first positive case to maximize sequence variability. For the hybrid *de novo* assembly, variants were calculated from the mismatches in the sequences reported by QUAST. The results also showed the best performance for the Illumina-only assembly approach, supporting the previous results and placing all three viral sequences in the B.1 sublineage of clade IIb (Table 4).

4. Discussion

Here, we provide the draft sequences of the first MPXV viral genomes isolated in the Canary Islands on 31 May 2022, corresponding to the B.1 sublineage of clade IIb. We used diverse sequencing and bioinformatic approaches, including methods for *de novo* genome assembly and combining short and long sequencing reads. The best results (higher sequence similarity and higher genome coverage compared to the reference) were obtained using a reference-based approach with Illumina-only reads. With this approximation and using combinations of mappers and variant callers, between 46 and 67 nucleotide variants were observed for the viral sequence isolated from the first patient when compared to the reference viral genome, which is compatible with the divergence of this new mpox outbreak from that of the 2018–2019 outbreak [21,30,66–68]. Both the procedures and findings were validated with samples isolated from three other positive patients infected with MPXV during the same Canary Islands outbreak.

For this first characterization, we relied on a metagenomic sequencing approach, providing a low proportion of sequence reads corresponding to the viral genome. In the study of the first patient, only 0.39% and 0.16% of the reads obtained from the Illumina and ONT runs, respectively, mapped to the MPXV reference. These data were further validated in independent viral samples from three other patients from the same mpox outbreak. As we demonstrate and other studies also report [69–71], this approach is straightforward and is not as dependent on viral sequence rearrangements in outbreaks, as is common for OPVs, demonstrating its value for detecting RNA and DNA viral pathogens in a few hours [72]. However, the costs per sample as well as the viral DNA concentration requirements in the patient samples hinder extended use. The use of alternative DNA extraction methods that enrich for viral DNA or remove host DNA could help to further improve the cost-efficiency of the overall procedure. After this work was completed [12], a more sensitive approach based on tiling amplicon sequencing for both Illumina and ONT workflows was developed and validated across a number of laboratories for MPXV [73]. Despite this method being more cost-efficient, periodic long-read metagenomics sequencing was recommended by the authors to monitor the emergence of viral variants with genomic rearrangements.

The main limitation of this study was the low viral genomic diversity of the samples that were circulating in the island throughout the study period, which hindered the possibility of testing our method for the detection of different lineages. However, we predict that because we used a metagenomics-based approach, our methods will be able to easily identify the presence of diverse viral variants and can be used to study genomic diversity in the future. Notably, all

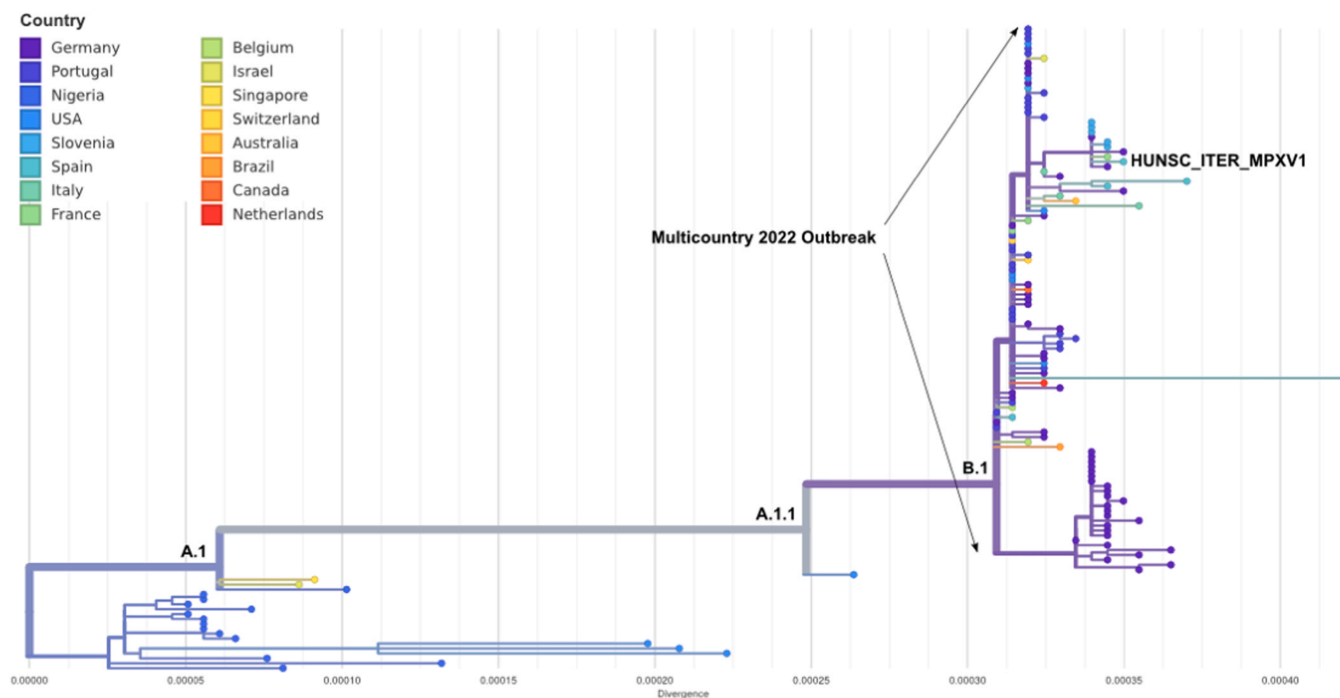


Fig. 3. A phylogenetic tree depicting the draft MPXV sequence isolated on May 31, 2022 from a patient from the Canary Islands mpox outbreak along with publicly available viral sequences at the NCBI GenBank.

Table 4
MPXV genome sequencing and classification from samples of three other positive patients from the same mpox outbreak to validate the findings of the previous patient.

Samples	Sequence type	Total length (bp)	Total Ns	Variants	Clade	Lineage
MPXV05	Illumina-only	197,189	1,387	68	IIb	B.1.10
	Hybrid *de novo* assembly	190,835	19,848	78	IIb	B.1.10
MPXV06	Illumina-only	197,210	145	79	IIb	B.1
	Hybrid *de novo* assembly	192,429	1,796	37	IIb	B.1
MPXV07	Illumina-only	197,203	131	73	IIb	B.1
	Hybrid *de novo* assembly	190,837	323	44	IIb	B.1

MPXV genome sequences analyzed here correspond to the same viral B.1 lineage, which was observed across Europe and other non-endemic areas during the 2022 outbreak, confirming the reliability of our results [60–65]. Another major limitation of this study is the poor quality of the read alignments at both ends of the MPXV reference genome due to the presence of tandem repeat regions (TRs) [74]. TRs produce high variability and structural complexity resulting in assembly and alignment limitations, increasing the likelihood of false-positive SNPs in the variant calling in these regions [69,75,76].

Overall, our results provide a proof of concept of the potential of introducing sequencing technologies, and metagenomics in particular, for rapid and precise pathogen diagnosis and surveillance, as is the case for the ongoing mpox outbreak caused by the emerging MPXV.

Ethics statement

The study was conducted at the University Hospital Nuestra Señora de Candelaria (Santa Cruz de Tenerife, Spain) during May 2022. The institutional review board approved the study (ethics approval number: CHUNSC_2022_83).

Funding

This study has been funded by Cabildo Insular de Tenerife (CGIEU0000219140 and "Apuestas científicas del ITER para colaborar en la lucha contra la COVID-19"); Instituto de Salud Carlos III (FI18/00230 and CD22/00138) cofunded by European Union (ERDF) "A way of making Europe"; by Fundación Canaria Instituto de Investigación Sanitaria de Canarias (PIFIISC 21/37); by the agreement with Instituto Tecnológico y de Energías Renovables (ITER) to strengthen scientific and technological education, training, research, development and innovation in Genomics, Personalized Medicine and Biotechnology (OA17/008); by the European Center for Disease Prevention and Control (ECDC/HERA/2021/024 ECD.12241); and by Convenio Marco de Cooperación Consejería de Educación-Cabildo Insular de Tenerife 2021–2025 (CGIAC0000014697).

CRediT authorship contribution statement

Adrián Muñoz-Barrera: Data acquisition, Software, Formal analysis, Writing – Original draft preparation, Writing - review & editing. **Laura Ciuffreda:** Data acquisition, Formal analysis, Writing - review & editing. **Julia Alcoba-Florez**: Conceptualization, Data acquisition, Formal analysis, Funding acquisition, Writing - review & editing. **Luis A. Rubio-Rodríguez:** Data acquisition, Formal analysis, Writing - review & editing. **Héctor Rodríguez-Pérez:** Data acquisition, Formal analysis, Writing - review & editing. **Helena Gil-Campesino:** Data acquisition, Writing - review & editing. **Óscar Díez-Gil:** Data acquisition, Writing - review & editing. **Josmar Salas-Hernández:** Data acquisition, Writing - review & editing. **Julia Rodríguez-Núñez:** Data acquisition, Writing - review & editing. **Antonio Íñigo-Campos:** Data acquisition, Writing - review & editing. **Víctor García-Olivares:** Data acquisition, Formal analysis,

Writing - review & editing. **Diego García-Martínez de Artola:** Data acquisition, Writing - review & editing. **Rafaela González-Montelongo:** Data acquisition, Funding acquisition, Writing - review & editing. **Agustín Valenzuela-Fernández:** Writing – Original draft preparation, Writing - review & editing. **José M. Lorenzo-Salazar:** Conceptualization, Software, Formal analysis, Writing – Original draft preparation, Writing - review & editing. **Carlos Flores:** Conceptualization, Data acquisition, Formal analysis, Writing – Original draft preparation, Funding acquisition, Writing - review & editing.

Data availability statement

The code used for this study is available at https://github.com/genomicsITER/monkeypox.

MPXV sequences of MPXV01, MPXV05, MPXV06 and MPXV07 samples obtained from the Illumina-only consensus approach have been released in the NCBI GenBank with accessions ON782054, OQ581847, OQ581848, and OQ581849, respectively. Hybrid *de novo* assemblies of MPXV01, MPXV05, MPXV06 and MPXV07 samples have also been released with accessions ON782055, OQ581850, OQ581851, and OQ581852, respectively.

Conflicts of interest

The authors declare no conflict of interest.

Acknowledgments

We would like to deeply thank the patients participating in the study. A.M-B., L.A.R-R., and J.M.L-S. acknowledge the University of La Laguna for the training support during the PhD studies.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.csbj.2023.03.020.

References

[1] Esposito J, Fenner F. Others. Poxviruses. Fields Virology. fourth ed. Lippincott Williams & Wilkins; 2001.
[2] Ryan KJ, George Ray C, Sherris JC. Sherris Medical Microbiology: An Introduction to Infectious Diseases. McGraw-Hill; 2004.
[3] Ligon BL. Monkeypox: a review of the history and emergence in the Western hemisphere. Semin Pediatr Infect Dis 2004;15:280–7. https://doi.org/10.1053/j.spid.2004.09.001
[4] Jezek Z, Fenner F. Human Monkeypox. Basel: Karger; 1988.
[5] WHO recommends new name for monkeypox disease n.d. https://www.who.int/news/item/28-11-2022-who-recommends-new-name-for-monkeypox-disease (accessed March 2, 2023).
[6] Parker S, Nuara A, Buller RML, Schultz DA. Human monkeypox: an emerging zoonotic disease. Future Microbiol 2007;2:17–34. https://doi.org/10.2217/17460913.2.1.17
[7] Gispen R, Brand-Saathof BB, Hekker AC. Monkeypox-specific antibodies in human and simian sera from the Ivory Coast and Nigeria. Bull World Health Organ 1976;53:355–60.
[8] Hutin YJF, Williams RJ, Malfait P, Pebody R, Loparev V, Ropp S, et al. Outbreak of human monkeypox, Democratic Republic of Congo, 1996 to 1997. Emerg Infect Dis J 2001;7:434. https://doi.org/10.3201/eid0703.017311
[9] Di Giulio DB, Eckburg PB. Human monkeypox: an emerging zoonosis. Lancet Infect Dis 2004;4:15–25. https://doi.org/10.1016/s1473-3099(03)00856-9
[10] Parker S, Buller RM. A review of experimental and natural infections of animals with monkeypox virus between 1958 and 2012. Future Virol 2013;8:129–57. https://doi.org/10.2217/fvl.12.130
[11] Reed KD, Melski JW, Graham MB, Regnery RL, Sotir MJ, Wegner MV, et al. The detection of monkeypox in humans in the western hemisphere. N Engl J Med 2004;350:342–50. https://doi.org/10.1056/NEJMoa032299
[12] A draft of the first genome sequence of Monkeypox virus associated with the multi-country outbreak in May 2022 from the Canary Islands, Spain. Virological 2022. https://virological.org/t/a-draft-of-the-first-genome-sequence-of-monkeypox-virus-associated-with-the-multi-country-outbreak-in-may-2022-from-the-canary-islands-spain/864 (accessed November 7, 2022).
[13] Peiró-Mestres A, Fuertes I, Camprubí-Ferrer D, Marcos MÁ, Vilella A, Navarro M, et al. Frequent detection of monkeypox virus DNA in saliva, semen, and other clinical samples from 12 patients, Barcelona, Spain, May to June 2022. Eur Surveill 2022;27. https://doi.org/10.2807/1560-7917.ES.2022.27.28.2200503.
[14] Orviz E, Negredo A, Ayerdi O, Vázquez A, Muñoz-Gomez A, Monzón S, et al. Monkeypox outbreak in Madrid (Spain): clinical and virological aspects. J Infect 2022;85:412–7. https://doi.org/10.1016/j.jinf.2022.07.005
[15] Miura F, van Ewijk CE, Backer JA, Xiridou M, Franz E, Op de Coul E, et al. Estimated incubation period for monkeypox cases confirmed in the Netherlands, May 2022. Eur Surveill 2022;27. https://doi.org/10.2807/1560-7917.ES.2022.27.24.2200448
[16] Vivancos R, Anderson C, Blomquist P, Balasegaram S, Bell A, Bishop L, et al. Community transmission of monkeypox in the United Kingdom, April to May 2022. Eur Surveill 2022;27. https://doi.org/10.2807/1560-7917.ES.2022.27.22.2200422
[17] Thornhill JP, Barkati S, Walmsley S, Rockstroh J, Antinori A, Harrison LB, et al. Monkeypox Virus Infection in Humans across 16 Countries - April-June 2022. N Engl J Med 2022;387:679–91. https://doi.org/10.1056/NEJMoa2207323
[18] Multi-country monkeypox outbreak in non-endemic countries n.d. https://www.who.int/emergencies/disease-outbreak-news/item/2022-DON385 (accessed November 7, 2022).
[19] Wenham C, Eccleston-Turner M. Monkeypox as a PHEIC: implications for global health governance. Lancet 2022;400:2169–71. https://doi.org/10.1016/S0140-6736(22)01437-4
[20] Kumar N, Acharya A, Gendelman HE, Byrareddy SN. The 2022 outbreak and the pathobiology of the monkeypox virus. J Autoimmun 2022;131:102855. https://doi.org/10.1016/j.jaut.2022.102855
[21] Isidro J, Borges V, Pinto M, Sobral D, Santos JD, Nunes A, et al. Phylogenomic characterization and signs of microevolution in the 2022 multi-country outbreak of monkeypox virus. Nat Med 2022;28:1569–72. https://doi.org/10.1038/s41591-022-01907-y
[22] Multi-country monkeypox outbreak in non-endemic countries: Update n.d. https://www.who.int/emergencies/disease-outbreak-news/item/2022-DON388 (accessed November 7, 2022).
[23] Multi-country monkeypox outbreak: situation update n.d. https://www.who.int/emergencies/disease-outbreak-news/item/2022-DON390 (accessed November 7, 2022).
[24] Heskin J, Belfield A, Milne C, Brown N, Walters Y, Scott C, et al. Transmission of monkeypox virus through sexual contact – a novel route of infection. J Infect 2022;85:334–63. https://doi.org/10.1016/j.jinf.2022.05.028
[25] Ogoina D. Sexual behaviours and clinical course of human monkeypox in Spain. Lancet 2022;400:636–7. https://doi.org/10.1016/S0140-6736(22)01497-0
[26] Girometti N, Byrne R, Bracchi M, Heskin J, McOwan A, Tittle V, et al. Demographic and clinical characteristics of confirmed human monkeypox virus cases in individuals attending a sexual health centre in London, UK: an observational analysis. Lancet Infect Dis 2022;22:1321–8. https://doi.org/10.1016/S1473-3099(22)00411-X
[27] Vusirikala A, Charles H, Balasegaram S, Macdonald N, Kumar D, Barker-Burnside C, et al. Epidemiology of early monkeypox virus transmission in sexual networks of gay and bisexual men, England, 2022. Emerg Infect Dis 2022;28:2082–6. https://doi.org/10.3201/eid2810.220960
[28] Antinori A, Mazzotta V, Vita S, Carletti F, Tacconi D, Lapini LE, et al. Epidemiological, clinical and virological characteristics of four cases of monkeypox support transmission through sexual contact, Italy, May 2022. Eur Surveill 2022;27. https://doi.org/10.2807/1560-7917.ES.2022.27.22.2200421
[29] Shu Y, McCauley J. GISAID: Global initiative on sharing all influenza data – from vision to reality. Eur Surveill 2017:22. https://doi.org/10.2807/1560-7917.ES.2017.22.13.30494
[30] Multi-country outbreak of Monkeypox virus: genetic divergence and first signs of microevolution. Virological 2022. https://virological.org/t/multi-country-outbreak-of-monkeypox-virus-genetic-divergence-and-first-signs-of-microevolution/806 (accessed November 7, 2022).
[31] Illumina whole-genome sequence of Monkeypox virus in a patient travelling from the Canary Islands to France. Virological 2022. https://virological.org/t/illumina-whole-genome-sequence-of-monkeypox-virus-in-a-patient-travelling-from-the-canary-islands-to-france/829 (accessed November 7, 2022).
[32] Monkeypox: experts give virus variants new names n.d. https://www.who.int/news/item/12-08-2022-monkeypox-experts-give-virus-variants-new-names (accessed March 2, 2023).
[33] Paniz-Mondolfi A, Guerra S, Muñoz M, Luna N, Hernandez MM, Patino LH, et al. Evaluation and validation of an RT-PCR assay for specific detection of monkeypox virus (MPXV). J Med Virol 2023;95:e28247. https://doi.org/10.1002/jmv.28247
[34] Mills MG, Juergens KB, Gov JP, McCormick CJ, Sampoleo R, Kachikis A, et al. Evaluation and clinical validation of monkeypox (mpox) virus real-time PCR assays. J Clin Virol 2023;159:105373. https://doi.org/10.1016/j.jcv.2022.105373
[35] Gul I, Liu C, Yuan X, Du Z, Zhai S, Lei Z, et al. Current and perspective sensing methods for monkeypox virus. Bioengineering 2022;9. https://doi.org/10.3390/bioengineering9100571
[36] Elbaz M, Halutz O, Ali Y, Adler A. Diagnosis of monkeypox infection: validation of two diagnostic kits for viral detection using RT-PCR. J Virol Methods 2023;312:114653. https://doi.org/10.1016/j.jviromet.2022.114653
[37] Jiang Z, Sun J, Zhang L, Yan S, Li D, Zhang C, et al. Laboratory diagnostics for monkeypox: an overview of sensitivities from various published tests. Travel Med Infect Dis 2022;49:102425. https://doi.org/10.1016/j.tmaid.2022.102425

[38] Huo S, Chen Y, Lu R, Zhang Z, Zhang G, Zhao L, et al. Development of two multiplex real-time PCR assays for simultaneous detection and differentiation of monkeypox virus IIa, IIb, and I clades and the B.1 lineage. Biosaf Health 2022;4:392–8. https://doi.org/10.1016/j.bsheal.2022.10.005
[39] UPDATE: Two draft genomes from Madrid, Spain, of the Monkeypox virus 2022 outbreak. Virological 2022. https://virological.org/t/update-two-draft-genomes-from-madrid-spain-of-the-monkeypox-virus-2022-outbreak/848 (accessed November 7, 2022).
[40] Health related news from the Government of the Canary Islands n.d. https://www3.gobiernodecanarias.org/noticias/sanidad-contabiliza-tres-casos-confirmados-y-tres-en-estudio-de-viruela-del-mono-desde-el-viernes/ (accessed November 7, 2022).
[41] Informe de situación noviembre2022 Alerta sobre infección de viruela del mono en España y otros países no endémicos n.d. https://www.sanidad.gob.es/profesionales/saludPublica/ccayes/alertasActual/alertaMonkeypox/docs/Informe_de_situacion_MPX_20221102.pdf (accessed November 16, 2022).
[42] ECDC, Factsheet for health professionals on monkeypox. European Centre for Disease Prevention and Control n.d. https://www.ecdc.europa.eu/en/all-topics-z/monkeypox/factsheet-health-professionals (accessed November 7, 2022).
[43] Li Y, Zhao H, Wilkins K, Hughes C, Damon IK. Real-time PCR assays for the specific detection of monkeypox virus West African and Congo Basin strain DNA. J Virol Methods 2010;169:223–7. https://doi.org/10.1016/j.jviromet.2010.07.012
[44] Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. Genome Biol 2019;20:257. https://doi.org/10.1186/s13059-019-1891-0
[45] Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 2018;34:3094–100. https://doi.org/10.1093/bioinformatics/bty191
[46] Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. ArXiv [q-BioGN] 2013.
[47] Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods 2012;9:357–9. https://doi.org/10.1038/nmeth.1923
[48] Broad Institute. Picard Toolkit 2018. http://broadinstitute.github.io/picard/ (accessed November 7, 2022).
[49] Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. Gigascience 2021;10. https://doi.org/10.1093/gigascience/giab008
[50] Pedersen BS, Quinlan AR. Mosdepth: quick coverage calculation for genomes and exomes. Bioinformatics 2018;34:867–8. https://doi.org/10.1093/bioinformatics/btx699
[51] Grubaugh ND, Gangavarapu K, Quick J, Matteson NL, De Jesus JG, Main BJ, et al. An amplicon-based sequencing framework for accurately measuring intrahost virus diversity using PrimalSeq and iVar. Genome Biol 2019;20:8. https://doi.org/10.1186/s13059-018-1618-7
[52] Wilm A, Aw PPK, Bertrand D, Yeo GHT, Ong SH, Wong CH, et al. LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. Nucleic Acids Res 2012;40:11189–201. https://doi.org/10.1093/nar/gks918
[53] First French draft genome sequence of Monkeypox virus, may 2022. Virological 2022. https://virological.org/t/first-french-draft-genome-sequence-of-monkeypox-virus-may-2022/819 (accessed November 7, 2022).
[54] Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLOS Comput Biol 2017;13:e1005595. https://doi.org/10.1371/journal.pcbi.1005595
[55] Wick RR, Schultz MB, Zobel J, Holt KE. Bandage: interactive visualization of de novo genome assemblies. Bioinformatics 2015;31:3350–2. https://doi.org/10.1093/bioinformatics/btv383
[56] Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics 2013;29:1072–5. https://doi.org/10.1093/bioinformatics/btt086
[57] Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 2013;30:772–80. https://doi.org/10.1093/molbev/mst010
[58] Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 2015;32:268–74. https://doi.org/10.1093/molbev/msu300
[59] Hadfield J, Megill C, Bell SM, Huddleston J, Potter B, Callender C, et al. Nextstrain: real-time tracking of pathogen evolution. Bioinformatics 2018;34:4121–3. https://doi.org/10.1093/bioinformatics/bty407
[60] Urgent need for a non-discriminatory and non-stigmatizing nomenclature for monkeypox virus. Virological 2022. https://virological.org/t/urgent-need-for-a-non-discriminatory-and-non-stigmatizing-nomenclature-for-monkeypox-virus/853 (accessed November 7, 2022).
[61] Cohen J. Rename monkeypox strains to remove geographic stigma, researchers say 2022. https://doi.org/10.1126/science.add4325.
[62] Khosravi E, Keikha M. B.1 as a new human monkeypox sublineage that linked with the monkeypox virus (MPXV) 2022 outbreak - Correspondence. Int J Surg 2022;105:106872. https://doi.org/10.1016/j.ijsu.2022.106872
[63] Luna N, Ramírez AL, Muñoz M, Ballesteros N, Patiño LH, Castañeda SA, et al. Phylogenomic analysis of the monkeypox virus (MPXV) 2022 outbreak: emergence of a novel viral lineage? Travel Med Infect Dis 2022;49:102402. https://doi.org/10.1016/j.tmaid.2022.102402
[64] Luna N, Muñoz M, Bonilla-Aldana DK, Patiño LH, Kasminskaya Y, Paniz-Mondolfi A, et al. Monkeypox virus (MPXV) genomics: a mutational and phylogenomic analyses of B.1 lineages. Travel Med Infect Dis 2023;52:102551. https://doi.org/10.1016/j.tmaid.2023.102551
[65] Scarpa F, Sanna D, Azzena I, Cossu P, Locci C, Angeletti S, et al. Genetic variability of the monkeypox virus clade IIb B.1. J Clin Med Res 2022;11. https://doi.org/10.3390/jcm11216388
[66] Chen Y, Li M, Fan H. The monkeypox outbreak in 2022: adaptive evolution associated with APOBEC3 may account for. Signal Transduct Target Ther 2022;7:323. https://doi.org/10.1038/s41392-022-01181-x
[67] Dumonteil E, Herrera C, Sabino-Santos G. Monkeypox virus evolution before 2022 outbreak. Emerg Infect Dis 2023;29:451–3. https://doi.org/10.3201/eid2902.220962
[68] Chakraborty C, Bhattacharya M, Sharma AR, Dhama K. Evolution, epidemiology, geographical distribution, and mutational landscape of newly emerging monkeypox virus. Geroscience 2022;44:2895–911. https://doi.org/10.1007/s11357-022-00659-4
[69] Vandenbogaert M, Kwasiborski A, Gonofio E, Descorps-Declère S, Selekon B, Nkili Meyong AA, et al. Nanopore sequencing of a monkeypox virus strain isolated from a pustular lesion in the Central African Republic. Sci Rep 2022;12:10768. https://doi.org/10.1038/s41598-022-15073-1
[70] Israeli O, Guedj-Dana Y, Shifman O, Lazar S, Cohen-Gihon I, Amit S, et al. Rapid amplicon nanopore sequencing (RANS) for the differential diagnosis of monkeypox virus and other vesicle-forming pathogens. Viruses 2022;14. https://doi.org/10.3390/v14081817
[71] Cohen-Gihon I, Israeli O, Shifman O, Erez N, Melamed S, Paran N, et al. Identification and whole-genome sequencing of a monkeypox virus strain isolated in Israel. Microbiol Resour Announc 2020;9. https://doi.org/10.1128/MRA.01524-19
[72] Alcolea-Medina A., Charalampous T., Snell L.B., Aydin A., Alder C., Maloney G., et al. Novel, Rapid Metagenomic Method to Detect Emerging Viral Pathogens Applied to Human Monkeypox Infections 2022. https://doi.org/10.2139/ssrn.4132526.
[73] Chen NFG, Chaguza C, Gagne L, Doucette M, Smole S, Buzby E, et al. Multi-site validation of an amplicon-based sequencing approach for human monkeypox virus. MedRxiv 2022. https://doi.org/10.1101/2022.10.14.22280783
[74] Zhao K, Wohlhueter RM, Li Y. Finishing monkeypox genomes from short reads: assembly analysis and a neural network method. BMC Genom 2016;17(Suppl 5):497. https://doi.org/10.1186/s12864-016-2826-8
[75] Yeh T-Y, Hsieh Z-Y, Feehley MC, Feehley PJ, Contreras GP, Su Y-C, et al. Recombination shapes the 2022 monkeypox (mpox) outbreak. Med 2022;3:824–6. https://doi.org/10.1016/j.medj.2022.11.003
[76] Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res 1999;27:573–80. https://doi.org/10.1093/nar/27.2.573