

METHODOLOGY

Open Access



Global mpox lineage discovery and rapid outbreak tracking with nanopore sequencing

Michael S. Bosmeny¹, Adam A. White^{1†}, Adrian A. Pater^{2†}, Jennifer Crew³, Joshua Geltz³ and Keith T. Gagnon^{1,2*}

Abstract

Insufficient tracking of virus introduction, spread, and new lineage emergence for the human monkeypox (mpox) virus 1 (hMPXV1) outbreak of 2022 hindered epidemiological studies and public health response. hMPXV1 mutations accumulated unexpectedly faster than predicted. Thus, new variants with altered pathogenicity could emerge and spread without early detection. Whole genome sequencing addresses this gap when implemented but requires widely accessible and standardized methodologies to be effective both regionally and globally. Here we developed a rapid nanopore whole genome sequencing method complete with working protocols, from DNA extraction to phylogenetic analysis tools. Using this method, we sequenced 84 complete hMPXV1 genomes from Illinois, a Midwestern region of the United States, spanning the first few months of the outbreak. The resulting five-fold increase in hMPXV1 genomes from this region established two previously unnamed global lineages, several mutational profiles not seen elsewhere, multiple separate introductions of the virus into the region, and the likely emergence and spread of new lineages from within this region. These results demonstrate that a dearth of genomic sequencing of hMPXV1 slowed our understanding and response to the mpox outbreak. This accessible nanopore sequencing approach makes near real-time mpox tracking and rapid lineage discovery straightforward and creates a blueprint for how to deploy nanopore sequencing for genomic surveillance of diverse viruses and future outbreaks.

Introduction

Since its first identification in May 2022, the current outbreak of mpox virus, formerly the monkeypox virus [1], has been the largest observed in non-endemic countries. While there have been over eighty-five thousand mpox cases worldwide to date during this outbreak, less than 6,000 viral genomes have been sequenced and shared publicly [2, 3]. Additionally, the rate of single-nucleotide polymorphisms (SNPs) is higher than would be expected

and novel deletions have been identified. General orthopoxviruses are estimated to acquire one or two substitutions per genome per year. However, the current mpox outbreak has over 50 SNP mutations more than the 2018/2019 reference strains [4]. In the months since the outbreak began, half a dozen more mutations have accumulated in some lineages. This equates to a clock rate of about 11–12 mutations per genome per year, several-fold more than expected. It is therefore conceivable that new variants with gain-of-function mutations or altered viral pathogenicity could arise.

Described in 1970 as an infection in humans distinct from other orthopoxviruses, mpox has similar symptoms to smallpox: fever, headaches, enlarged lymph nodes, and a rash leading to lesions [5]. These lesions, which cause the skin to crust, can also lead to secondary bacterial infections. Prior to the current outbreak, most human cases were observed in Africa, specifically West Africa, which were classified as Clade II, and the Congo Basin,

[†]Adam A. White and Adrian A. Pater authors contributed equally.

*Correspondence:

Keith T. Gagnon
ktgagnon@siu.edu

¹ Department of Biochemistry and Molecular Biology, Southern Illinois University School of Medicine, Carbondale, IL, USA

² Department of Chemistry and Biochemistry, Southern Illinois University, Carbondale, IL, USA

³ Illinois Department of Public Health, Springfield, IL, USA



which were classified as Clade I [6] (Fig. 1A). The current outbreak is an offshoot of the West African Clade II but has sufficient mutations to be categorized as Clade IIb, also known as human mpox virus 1 (hMPXV1). Within Clade IIb, there is a fine-scale classification based on a lineage naming scheme, for example A, A.1, A.1.1, and so forth. Nearly all the current outbreak sequences, by this categorization, fall into lineage B.1 (Fig. 1B). The current geographical and lineage makeup of mpox genome sequences in the world and United States are shown in Fig. 1C, D, respectively.

During the SARS-CoV-2 pandemic, there were publicly-available protocols to sequence viral genomes via large staggered amplicons using the Oxford Nanopore Technologies (ONT) MinION nanopore platform [7]. This allowed many smaller labs worldwide to sequence local samples quickly and without costly equipment. To facilitate a similar system for mpox, we developed an all-in-one MinION approach that retrofitted similar reagents and software that we and others previously used for SARS-CoV-2 genome sequencing [8]. This approach

provides an accessible methodology that allows a greater number of laboratories to begin sequencing and submitting mpox genomes to global databases. It also serves as a blueprint for how to rapidly design, deploy, and share standardized and accessible genome sequencing technology for virtually any viral pathogen during regional or global outbreaks.

Our rapid targeted nanopore sequencing method was used to sequence 103 mpox patient samples from the state of Illinois, a Midwest region of the United States, resulting in 84 full genomes. Sample processing and sequencing of 1–24 samples could be accomplished in as little as 48 h by a single researcher. This sequencing revealed new insight into the regional introduction and emergence of lineages, as well as their subsequent spread, which included discovery of new lineages that may have emerged in the Midwest region and spread internationally. Thus, regional efforts by even small laboratories equipped with practical sequencing tools can have a large impact on the global scientific consensus for a disease epidemic like mpox. While the current mpox outbreak

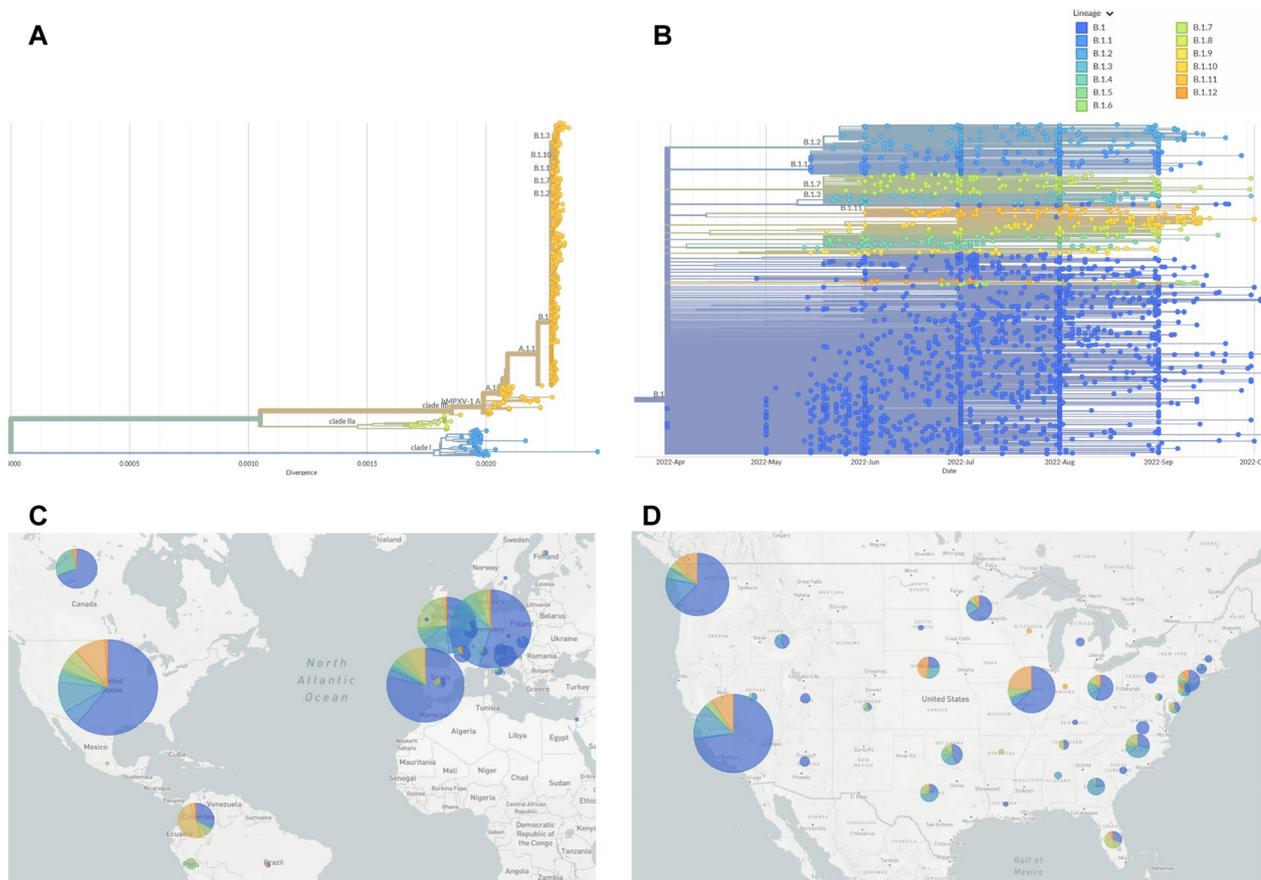


Fig. 1 **A** Phylogenetic tree of sequenced Mpox cases, covering multiple clades. The current outbreak is Clade IIb. **B** Phylogenetic tree of only Clade IIb, lineage B.1 Mpox cases. Color coded based on lineages within B.1. **C** World map showing distribution of sequences and lineages. The larger the circle, the more sequences. **D** United States map showing distribution of sequences and lineages

appears to be subdued, future outbreaks are possible. In addition, amplicon paneling with nanopore sequencing is a flexible platform that can be retrofitted to achieve rapid genomic surveillance for a diverse array of pathogenic viruses to help track and control outbreaks.

Results

Targeted amplicon paneling and nanopore sequencing

A set of primer pools, protocols, and computational commands were designed to sequence hMPXV1 genomes on a MinION nanopore flow cell (Fig. 2A). All necessary protocols are briefly described in the Methods section and available as Additional file 2: *Working Laboratory Protocols*. Initial amplicon primers were designed using PrimalScheme [9], which constructs a series of tiled amplicons of set size using a known genetic template,

against NCBI reference sequence NC_063383.1, a Clade IIb, Lineage A sequence from 2018. To ensure that the primers had the maximum compatibility with both major clades of mpox and modern outbreak samples, six different genomes encompassing mpox samples from 2003 to the modern day were aligned and the designed primers compared against them. Only primers that matched all six genomes, indicating they resided in areas of low genetic variability, were retained. Those that contained mismatches were redesigned. These ensure the highest likelihood of the primers working with new mutations of mpox in this and future outbreaks. The final design incorporated 73 tiled amplicons of approximately 3,000 base-pairs (bp) in length, which were divided up into two pools. Pool 1 contained primers for odd-numbered amplicons and Pool 2 contained primers for

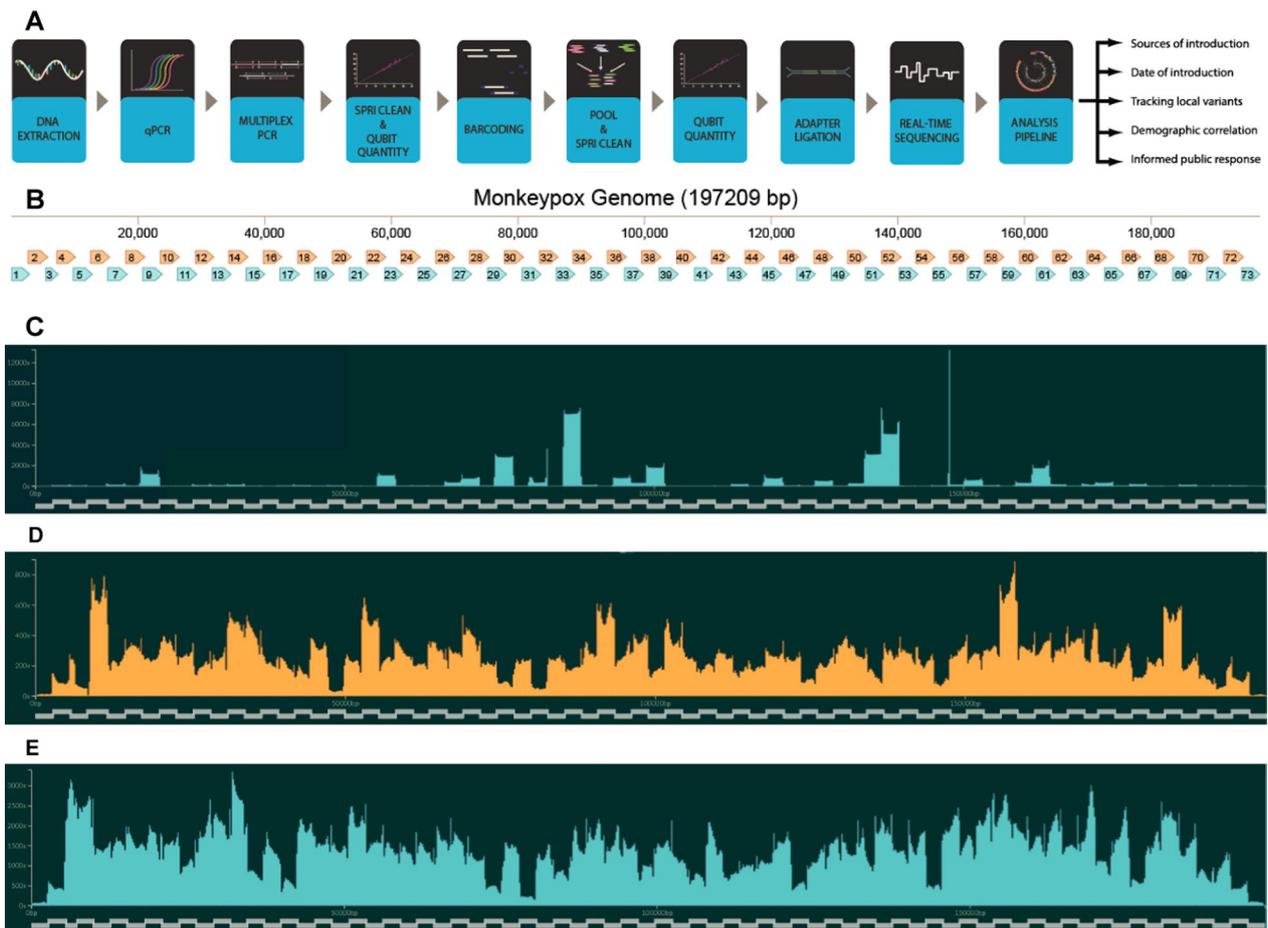


Fig. 2 **A** Diagram showing workflow for preparation and sequencing of mpxv sample. **B** Amplicon Map of Mpxv Genome. Pool 1 amplicons are shown in blue, while Pool 2 amplicons are shown in orange. Amplicons 1 and 73 are inverted matches of each other, as are amplicons 2 and 72. **C–E** Coverage of individual amplicons in the Mpxv genome during sequencing. Derived from FASTQ reads using RAMPART software [7]. The tiled amplicons are shown as the X-axis. The Y-axis indicates how many reads match that location in the genome. **C** First sequencing attempt of genome USA-2003, using USA-2003 as a reference genome. **D** Refined sequencing attempt of USA-2003. **E** An example sequencing attempt of a modern 2022 outbreak genome, using a lineage B.1 sequence as the reference genome

even-numbered amplicons (Fig. 2B). Due to the terminal repeat regions (~6,400 bp) at the ends of the genome, known as the inverted terminal repeats (ITRs), amplicons 1 and 73 have the same primers, as do amplicons 2 and 72. Amplicons 3 and 71 have the same nested outer primers, but different internal primers.

PCR amplification with primer pools was tested against a known mpox virus genome, USA-2003 from a 2003 outbreak in Wisconsin, which was cultured in green monkey cells (BEI Resources, NR-4928). Because commercial samples contained low virus copy numbers, an unbiased whole-genome amplification method was first used prior to PCR amplification. Resulting amplicons from each primer pool after PCR were combined, bar-coded, and sequenced on a MinION instrument. Reads were used to generate a consensus sequence. This initial PCR and sequencing attempt only yielded 20× coverage of 78.3%, which limits full genome acquisition since tools like Nanopolish require 20× coverage to call a consensus sequence. Amplicon distribution was skewed towards producing a high number of a handful of amplicons. In addition, only 52% of the total reads mapped to the mpox genome, indicating that non-specific amplification was occurring (Fig. 2C).

After careful analysis, primers prone to off-target amplification of green monkey or human genomic DNA were identified and replaced. We also found that amplicons 18, 20, 38, 40, 43, and 52 were slightly under-represented while amplicons 32 and 51 were over-represented. We subsequently doubled the concentration of primers for amplicons 18, 20, 38, 40, 43, and 52. Other changes included optimized thermocycler conditions and changing barcoding from a ligation-based method to a transposase-based rapid barcoding kit, which yielded better and faster results. A repeated sequencing attempt utilizing these changes had 20× coverage of 97.10% and 97.42% of reads matched to the USA-2003 genome, indicating a clean PCR amplification (Fig. 2D). The resulting consensus sequence was found to exactly match the expected USA-2003 genome in all regions with sufficient coverage.

We then deployed this method to sequence current 2022 mpox-infected patient samples from the state of Illinois across the first few months of the outbreak. Magnetic bead-based DNA extraction was performed, followed by viral load quantification by quantitative PCR (qPCR). Testing facilities where samples were obtained used a generic non-variola *orthopoxvirus* (NVO) qPCR probe designed by the Center for Disease Prevention and Control (CDC) for testing for the presence of mpox viral DNA. Other testing facilities may use mpox-specific CDC probe-based qPCR. This probe, while more precise in detecting mpox, has failed on a small subset of mpox

samples first observed in California. This is due to partial ITR deletions where the qPCR probe binds. These samples were only identified by testing with the NVO qPCR. To guard against similar problems, we designed our own probe-based qPCR reaction, OPG123, based on analysis of low-mutational areas of the mpox genome. To standardize reads, sample DNA was tested in the same qPCR conditions and using the same baseline threshold value. In general, our cycle threshold (C_t) values were approximately four C_t values higher than those generated by testing facilities (Additional file 1: Fig. S1). This could be due to different qPCR conditions, degradation during shipping, or freeze–thaw cycles.

To quantify the number of genomes needed for successful sequencing, a short double-stranded DNA fragment was synthesized and used for amplification with a known copy number (Additional file 1: Fig. S2). Based on these results, between 5,000 and 10,000 copies of the Mpox genome are needed for optimal amplicon production and sufficient genome sequence coverage. This correlated to a C_t value of 32 or lower with our qPCR probe OPG123. However, based on qPCR comparisons between our probe, the CDC mpox-specific probe, and an RNase P probe (human genomic DNA), this copy number limitation may be linked to the amount of mpox DNA compared to contaminating human DNA (Additional file 1: Fig. S3). Samples with low OPG123 C_t values also had a high ratio of Mpox DNA to human DNA (8–12 C_t difference) while samples with high OPG123 C_t values had lower ratios of Mpox DNA to human DNA (3–4 C_t difference).

After qPCR quantification, hMPXV1 genomic DNA was PCR amplified separately with primer Pool 1 and Pool 2. Resultant amplicons from the two reactions were then combined and proceeded through a series of clean up and library preparation steps (see Additional file 2: *Working Protocols*). After nanopore sequencing, which could be achieved by barcoding up to 24 samples on one flow cell, consensus sequences were generated for each sample. In total, 103 patient samples were processed, resulting in 84 full genomes based on consensus sequences. Additional file 1: Table S1 shows C_t values and 100×/20× coverage numbers for these 84 samples. The mean coverage of the consensus sequences was 98.79%, with a median coverage of 99.41%. An example of 99.95% sequence coverage as viewed in Rampart is shown in Fig. 2E. One challenge to genome coverage was the terminal ITR repeats. Minimap2, the program used to align reads to a template genome, attempts to match these reads to one end of the genome or the other, but not both. This is an issue if there is a low number of reads in that region. This can result in uncalled bases (Ns) in Amplicons 1 and 73 in the final consensus. If insufficient

reads are generated, one strategy is to remove the terminal repeat region from one end of the template genome so that minimap2 aligns all reads on the other terminal repeat. The removed repeat can then be manually re-added.

A diverse set of viral introductions

After sequencing, all consensus sequences were analyzed using NextClade software [10]. This compares the sequence of interest against a generic Clade IIb, lineage B.1 template genome, based on ON563414.3, MPXV_USA_2022_MA001, a sample from the current outbreak collected in May 2022 in Massachusetts, USA, and highlights any indels or mutations. The sequences can then be categorized into specific sub-lineages within B.1. As of October 2022, there are twelve sub-lineages, as defined by the NextClade group [11] (Fig. 1B). Each has one to four unique nucleotide mutations. There are also the remaining approximately 60% of the current outbreak sequences that fall within the B.1 lineage, but do not fit within a designated sub-lineage. Our sequencing suggests that Illinois has seen a significant number of different viral introductions. Table 1 shows some of the mutations encountered in our sequencing.

In addition, another eight sequences possessed unique mutations not in common with any of the other sequences in the NCBI database. These diverse mutational patterns in a small number of cases indicate that a substantial number of viral versions entered the region and spread.

Novel B.1 lineage discovery

Our genome sequencing also suggests that several B.1 lineages may have emerged in Illinois or the Midwest region.

Of the 84 samples sequenced, a quarter of them had four mutations in common: C18133T, G67611A, G130231A, and G159277A (Fig. 3). Three of these mutations are in non-coding regions, while the last, G159277A, produces mutation E121K in protein OPG185. At the time of this sequencing, this mutation categorization was still unnamed. It has subsequently been designated as lineage B.1.11 by Nextclade. Currently, B.1.11 is the fifth-most common lineage in the current Mpox outbreak, with eighty-four total sequences. Our sequencing contributed 25% of the genomes needed to help establish this recently named lineage. This lineage is primarily found in the U.S., with just a few samples from outside North America: five August and September infections from Columbia, an early August infection in Portugal, and a September infection in Germany. Although the dearth of sequencing makes it difficult to know with certainty the origin of this lineage, the earliest cases are all found in Illinois in late June. It is not until July that similar strains appear in California, Washington, and other regions of the U.S.

Another major mutational pattern seen in our sequencing has only just recently received a lineage designation, B.1.13. (Fig. 4). Nineteen of our eighty-four samples have G175093A, and of those, eighteen have a secondary mutation of C132520T. G175093A confers the mutation D188N in protein OPG204, while C132520T is in a noncoding region. B.1.13 is primarily found in the Midwest of the U.S. It appears to have originated there, with infections appearing with these mutations in mid-June. From there, it appears to have gradually spread to other Midwestern areas, like Oklahoma, Minnesota, Ohio, Tennessee, and Kentucky. So far, no sequences with this mutational pattern have appeared in California and only a small number in Washington state, despite their high

Table 1 List of most common mutations found in sequenced mpox samples, as compared to a Clade IIb, Lineage B.1 template genome

Nucleotide mutations	Lineage name	Mutational pattern sequenced	Mutational pattern found elsewhere	Total
G186165A (OPG210:D1604N)	B.1.2	2	192	194
G74360A (OPG094:R194H)	B.1.1	3	189	192
C25644T	B.1.7	5	148	153
C18133T, G67611A, G130231A, and G159277A (OPG185:E121K)	B.1.11	21	68	89
G34308A (OPG053:S128L)	B.1.4	1	49	50
C132520T, G175093A (OPG204:D188N)	B.1.13	18	19	37
C77462T (OPG098:S188F)	-	4	21	25
G22814A (OPG038:H116Y), G55162A (OPG074:S655L), G137128A, G167910A (OPG193:E193K)	-	2	1	3
C88308T	-	3	0	3
C156533T	-	3	0	3

Resulting protein changes are in parentheses after nucleotide changes

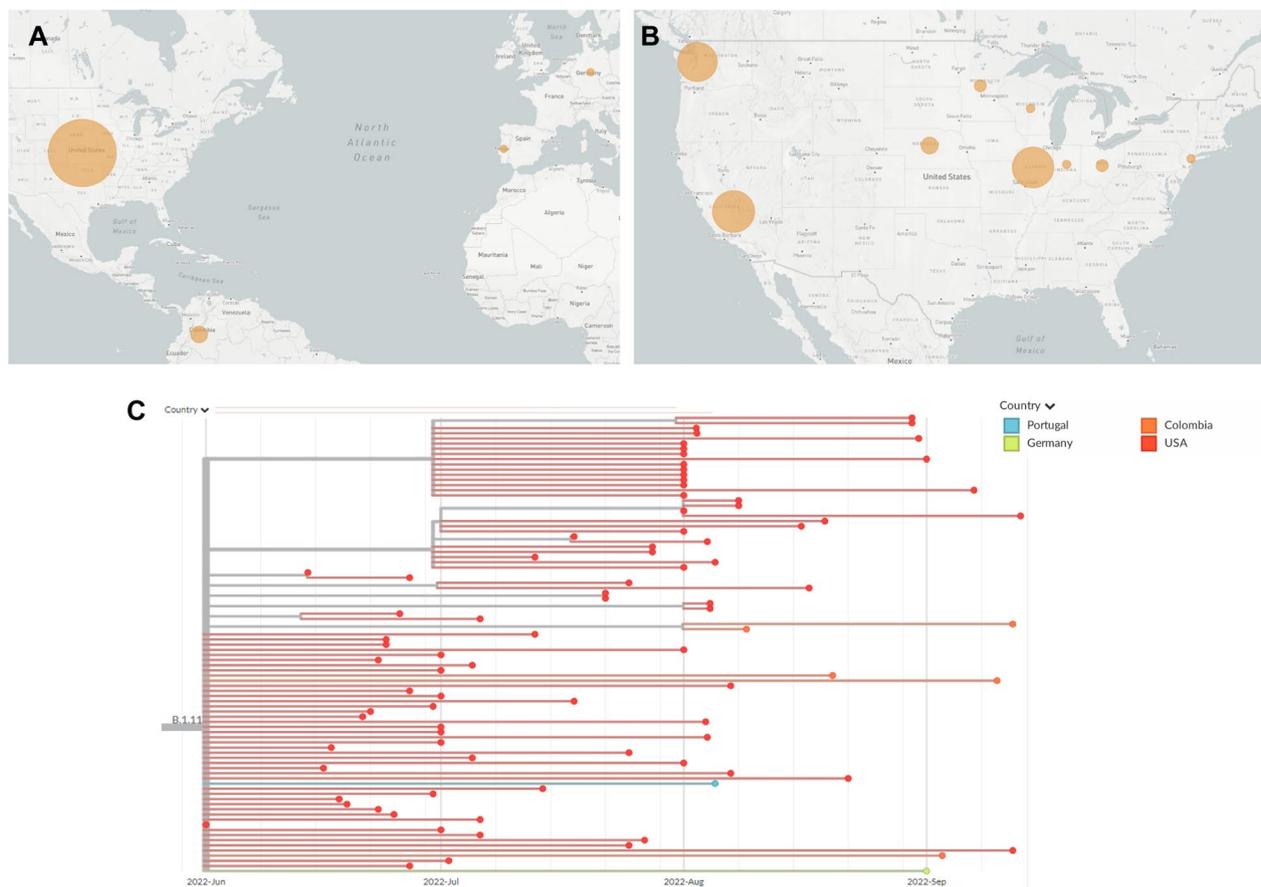


Fig. 3 Geographic and phylogenetic data for Mpox lineage B.1.11. **A** World map showing lineage distribution by country. The larger the circle, the higher the number of sequences located there. Compare with total sequences in Fig. 1C. **B** United States map showing lineage distribution by state. Compare with total sequences in Fig. 1D. **C** Phylogenetic tree showing collection dates and locations for each sequence of B.1.11

sequencing rate. Finally, in late July a similar sequence was seen in the United Kingdom and in late August another sequence in the group was seen in Germany.

Currently, B.1.13 is the eleventh most common lineage in the 2022 global outbreak, with a total of thirty-seven sequences. Half of the sequences in B.1.13 were generated from this study. This demonstrates the need for deeper regional sequencing and the ability of a small number of regional genomes to impact global epidemiology.

Discussion

It seems likely that B.1.11 and B.1.13 lineages originated in the U.S. Midwest, possibly even Illinois given that Chicago is a travel hub. However, it is impossible to know with any certainty due to the lack of systematic mpox genome sequencing in the U.S. and internationally. Many U.S. states have less than a dozen sequences submitted. Washington, Connecticut, and California are standouts with high numbers of genome sequences, but they are exceptions. Our recent sequencing increased the number of genomes from Illinois by

five-fold. There could certainly be new mutational variants appearing that have remained uncharacterized because sequencing is absent in many U.S. states, such as New York, Florida, and Texas, and very low in most other countries (Fig. 5). In Europe, both Portugal and Germany are submitting an impressive number of sequences. In contrast, Spain and France have uploaded samples for less than 1% of their total cases.

The 2022 mpox outbreak is noted to have 6–12 fold more mutations from previous 2018–2019 strains than expected, based on the calculated substitution rates for *orthopoxviruses* [4]. As a great number of these mutations are of the same pattern, GA → AA or TC → TT, it has been speculated that this new mutation rate may be driven by the human APOBEC3 enzyme. APOBEC is a family of cytidine deaminases that can attempt to disrupt viral replication by nucleotide alteration based on specific motifs within the nucleotides [12, 13]. APOBEC3 specifically targets TC positions and modifies the C to U, which, upon replication will cause TC to



Fig. 4 Geographic and phylogenetic data for Mpxv lineage B.1.13. **A** World map showing lineage distribution by country. The larger the circle, the higher the number of sequences located there. Compare with total sequences in Fig. 1C. **B** United States map showing lineage distribution by state. Compare with total sequences in Fig. 1D. **C** Phylogenetic tree showing collection dates and locations for each sequence in B.1.13

become TT. GA → AA mutations are the result of a TC → TT mutation on the other strand.

With limited genome sequencing and faster than expected mutation rates, it is likely that the full picture of pathogen evolution and host adaptation is not being captured. Although the number of mpx cases may wax and wane, the hMPXV1 virus may become endemic to many new international regions, including the U.S. While the likelihood seems low, individual mutations or groups of mutations could nevertheless arise that significantly alter virus pathogenicity, severity, or infectivity. With SARS-CoV-2, thanks to global sequencing efforts, public health officials are now quickly and regularly alerted to new, more contagious variants, like Omicron. A similar approach should be preemptively applied to mpx and all new viral outbreaks.

The hMPXV1 genome is six times larger than SARS-CoV-2 and possesses inverted terminal repeats, a number of microsatellites, and homopolymeric mononucleotide repeats, which all introduce sequencing challenges [14]. Nonetheless, genome sequencing shortfalls are not the result of these technical challenges but instead a lack of accessible resources. Most mpx virus genome

sequencing has been performed with metagenomics shotgun approaches, which sequence everything and can be time consuming and expensive compared to targeted sequencing [15]. Near the end of this study, a commercially available kit was released for targeted sequencing on Illumina platforms. However, it is designed primarily for sequencing core facilities and private contract laboratories that may process large numbers of samples. In contrast, many small research groups and developing regions purchased and used nanopore sequencing supplies to sequence SARS-CoV-2 genomes during the COVID-19 pandemic. Thus, our approach should be nearly “plug-and-play” with existing resources for many research groups.

Amplicon paneling is a method introduced previously and developed by the ARTIC group for targeted SARS-CoV-2 sequencing [16]. It is the method used for Illumina and nanopore sequencing platforms. For nanopore sequencing, amplicon size is very flexible and can range from only a few hundred base-pairs up to tens of thousands of base-pairs. In contrast, Illumina technology has traditionally been severely limited to short reads, which requires substantially more amplicons to be generated.

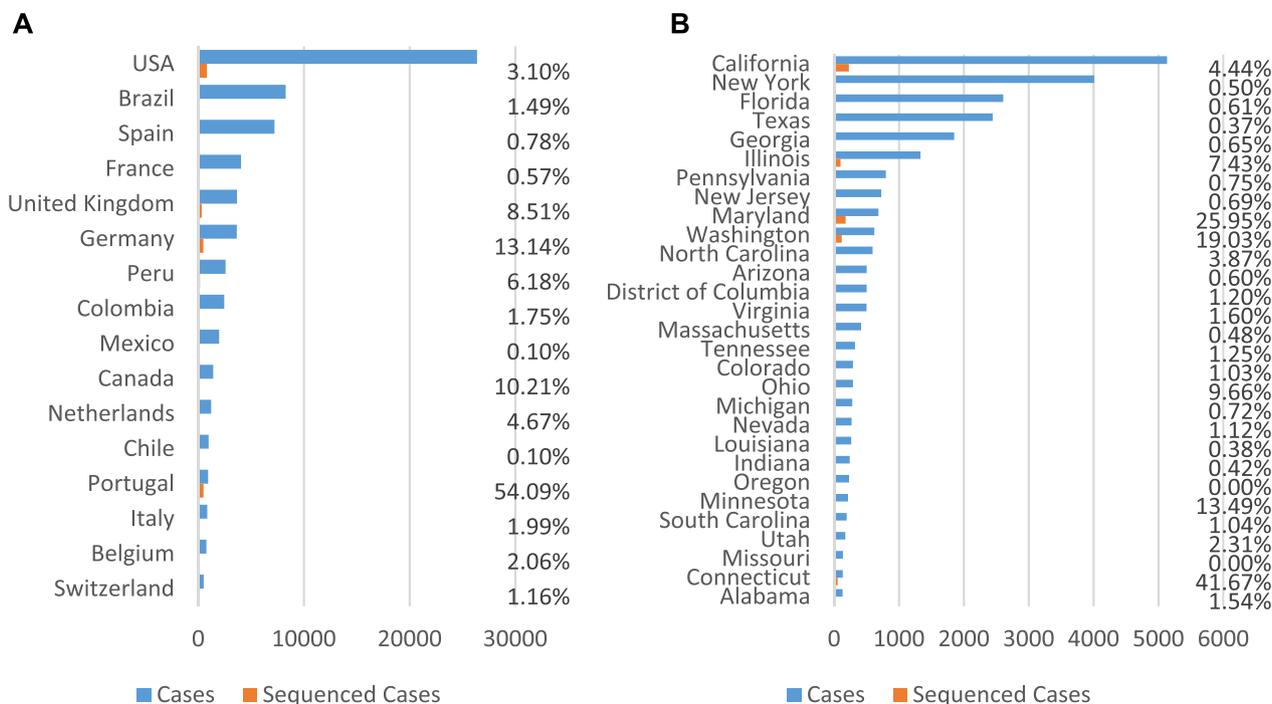


Fig. 5 Comparisons of mpox cases versus mpox sequences for the current 2022 outbreak. **A** Countries with more than 500 cases. **B** US States with more than 100 cases

Larger amplicon sizes are particularly valuable for mpox genome sequencing due to repeat regions. For example, there is an ATATACATT motif that repeats dozens of times, starting at base 197,077. Small, 100–200 bp fragments that an Illumina sequencing run will generate may not fully span this repeat motif, and it is therefore extremely difficult to accurately count how many times the motif occurs. A look at current sequences (Additional file 1: Fig. S4) serves to illustrate this as over half of all sequences list a deletion in this area. Interestingly, there are no patterns to these deletions on the phylogenetic tree, leading to the conclusion that these deletions are sequencing artifacts. Nanopore sequencing does not have this problem because the amplicon can be larger than the repeat region, making it possible to know with certainty whether deletions actually exist. One disadvantage of nanopore sequencing is base calling for mononucleotide, or homopolymeric, repeats. Nanopore technology reads bases as grouped k-mers using a voltage-measurement as the nucleic acid strand passes through the pore. Long stretches of mononucleotides can sometimes be miscounted. Consensus-finding programs can help correct for this and newer flow cell chemistries can boast a higher base-calling accuracy.

In summary, we have developed a rapid nanopore sequencing method that should prove highly accessible to small laboratories and international groups participating

in local mpox genome sequencing efforts. It borrows from the success of nanopore and amplicon tiling methods, can be high-throughput with 96-well plates and barcoding, and we provide full laboratory *Working Protocols* to support immediate implementation. The amplicon tiling primer sets created in this study are compatible with all known mpox genomes and were designed to extend to likely future variants. By employing this method on 103 regional samples, we sequenced 84 complete hMPXV1 genomes, which significantly increased the number of genome sequences regionally and nationally. It also offered unexpected insight into both regional and global dynamics of the current mpox outbreak. Multiple introductions of mpox occurred in the U.S. state of Illinois early in the outbreak, some of which then spread regionally. Well-placed travel restrictions or other public health policies might have limited the impact of the outbreak. In addition, multiple mutational profiles were observed in Illinois that have not been found elsewhere. It is possible that these emerged in Illinois or neighboring Midwest regions. New lineages of hMPXV1 were identified, one of which has been designated as B.1.11 on the global phylogenetic tree and one of which remains undesignated. Both of these lineages have a high likelihood of emergence within Illinois or the U.S. Midwest region. Our sequencing was able to fill gaps in the global phylogenetic tree that provide a better picture of the evolution of the

virus and its adaptation to the human host. Finally, our approach creates a straightforward blueprint for adapting amplicon paneling and nanopore sequencing for accessible and affordable genomic surveillance to track nearly any viral outbreak where the virus' genome sequence, or a close relative, is known. This blueprint should also be applicable to future mpox outbreaks.

Methods

Extraction of viral genomes

Initial testing and optimization was done with a genome acquired from BEI Resources (<https://www.beiresources.org/Catalog/BEINucleicAcids/NR-4928.aspx>). Low copy number in this sample was augmented with whole-genome amplification via Repli-G treatment (Qiagen, # 150023). All sequenced samples were obtained from the Illinois Department of Public Health.

A detailed protocol is available in Additional file 2: Working Protocol 1. Briefly, samples are stored at -80°C until ready for processing. Virus sample should be delivered in an inactivation (lysis) solution, referred to here as the viral transport media (VTM). Extractions should be done under sterile conditions in a BSL-2-class biosafety cabinet. For consistency, extraction and all future steps should also be performed on a negative control sample containing only VTM or phosphate-buffered saline (PBS). Any genomic sequences that show up in the negative control at the end of the workflow should be considered to be possibly present as a contamination in any other samples processed at the same time. DNA is extracted from virus samples using a magnetic bead-based process. First, DNA is bound to the magnetic bead in an isopropanol solution. After binding, a magnet is used to separate the beads and bound DNA from the supernatant liquid. The beads and DNA is then washed twice with an 80% ethanol solution. Finally, the DNA is eluted from the beads by the addition of a DNA elution buffer.

qPCR to estimate viral load

To quantify the relative amount of viral genomic DNA in each extraction, qPCR is performed on each sample. In initial testing, we quantified DNA using two different probes: One which uses the same probe sequences as the CDC's Mpox qPCR tests (Sequences: Forward: 5'-GGA AATGTAAGACAACGAATACAG-3'; Reverse: 5'-GCT ATCACATAATCTGGAAGCGTA-3'; Probe: 5'-AAGCCGTAA TCTATGTTGTCTATCGTGTCC-3') and one of our own design, called OPG123 (Sequences: Forward: 5'-GGG ATGCTAATATTCCCAAACATAC-3'; Reverse: 5'-TCTAAA TCCGCATTAGACACCTT-3'; Probe: 5'-AGAGATATCAGC CGCAATAGCATCCC-3'). Later, we tested exclusively with OPG123. Both probes were manufactured by IDT.

Typically, the standard manufacturer's qPCR protocol for the PrimeTime gene expression assay was followed (Additional file 2: Working Protocol 2). Supplementary Fig. 5 shows negative and positive bands for these qPCR tests.

Multiplex amplicon tiling PCR to amplify viral genomes

For all samples that have sufficient material to be sequenced, based on the qPCR results, the next step is to amplify the DNA using a series of 73 overlapping, 3000 base pair amplicons. To prevent interference between adjacent amplicons, odd-numbered amplicons are used in one pool, and even-numbered amplicons are used in a second pool. This is modeled after the ARTIC Network's SARS-CoV-2 V3 sequencing primers, which produce a similar pattern of amplicons. NEB Q5 Hot-Start Polymerase was used for our PCR reactions (Additional file 2: Working Protocol 3). Several adjustments were made to the concentrations of the oligos to produce the best sequencing results:

It is recommended, at least for initial sequencing runs, to verify PCR products for each sample and pool on a 1% agarose gel via electrophoresis. Correct band sizes are approximately 3000 bp. Supplementary Fig. 6 shows an example of several samples checked on a gel, along with a negative control.

Pooling, purification, and quantification of PCR amplicons

After PCR amplification, Pool 1 and Pool 2 for each virus sample are pooled together and cleaned of PCR reactants using SPRI AMPure XP beads (Additional file 2: Working Protocol 4). The ratio of bead solution to DNA solution can be varied to select out smaller DNA fragments, if desired. Initial testing, using 1:1 bead solution to DNA solution ratios, showed that there were many small, <200 bp DNA fragments that occupied bandwidth in the later MinION sequencing reaction. By reducing the ratio down to 0.75:1 bead solution to DNA solution, an abundant amount of 3000 bp fragments were preserved while excluding many of these smaller products, giving a cleaner sequencing run.

After purification, the quantity of DNA in each sample needs to be determined. DNA was quantified using a Qubit 2.0 fluorometer and the Qubit double-stranded DNA assay kit (Additional file 2: Working Protocol 5). This quantification allows for an equal amount of each sample to be carried forward to future steps and, ideally, an equal number of reads per sample when the finished process is loaded into the flow cell for sequencing.

DNA library barcoding and adapter ligation

To prepare the amplicons for sequencing, our lab utilized the Oxford Nanopore RAPID Barcoding kit. This kit

uses a transposase which cleaves template and attaches barcoded tags to the cleaved ends (Additional file 2: Working Protocol 6). Afterwards, the barcoded samples were pooled, and then cleaned via bead extraction. Rapid adapters were added to pooled barcoded samples and allowed to attach by a room temperature incubation (Additional file 2: Working Protocol 7).

Nanopore sequencing with a MinION

Sequencing was done using the Oxford Nanopore Technologies MinION Mk1B (MIN-101B) sequencer and Spot-ON Flow Cell, R9 version (FLO-MIN106D). Oxford Nanopore's MinKNOW software (v21.11.9) was used to read from the flow cell and base-call the reads (Additional file 2: Working Protocol 8).

During sequencing, ARTIC's RAMPART software (v1.0.6) was used to monitor the reads on each barcode separately and the respective coverage of the Mpox genome (Additional file 2: Working Protocol 9).

Phylogenetic analysis, variant calling, and database depositing

The ARTIC bioinformatics pipeline (v1.2.1) was used to consolidate FASTQ reads from MinKNOW and align them against the NextClade B.1 Mpox template genome (Additional file 2: Working Protocol 10).

NextClade software was used to identify mutations in each sequence against the template genome (Additional file 2: Working Protocol 11).

All sequences were uploaded into the NCBI GenBank system using their BankIt upload process.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12985-023-02059-2>.

Additional file 1. Supplementary Figures and Table.

Additional file 2. WorkingProtocols.

Acknowledgements

We thank the Illinois Department of Public Health for access to patient samples.

Author contributions

MSB developed and validated all protocols with assistance of other authors, supervised the project, extracted samples, ran multiplex PCR, prepared sequencing libraries, sequenced samples, analyzed data, interpreted results, prepared figures, and wrote the manuscript. AAW assisted in protocol development, especially custom qPCR probe design, reaction optimization and validation. AAP assisted in protocol development, especially in development of multiplex PCR primers and Rampart visualization. JC and JG provided validated samples and assisted in editing of the manuscript. KTG conceived of and directed the project, as well as assisted in figure preparation and writing of the manuscript. All authors read and approved the final manuscript.

Funding

There is no funding to report for this project.

Availability of data and materials

All sequencing data is available. Complete consensus mpox genome sequences have been deposited at the National Center for Biotechnology Information (NCBI). Accession numbers correlating with each sample are listed in Additional file 1: Table S1. Compressed fastq files for all sequencing runs are available at <https://zenodo.org/record/7659572>. Fast5 files of raw nanopore data for all sequencing runs are available upon request due to the large file size associated with the fast5 file type.

Declarations

Ethics approval and consent to participate

Patient consent was waived for this study by the Southern Illinois University School of Medicine Institutional Review Board (IRB) since no significant patient-identifiable information was shared or generated in this study.

Consent for publication

Publication consent was not required for this study.

Competing interests

The authors declare no competing interests.

Received: 26 April 2023 Accepted: 1 May 2023

Published online: 06 May 2023

References

1. WHO (<https://www.who.int/news/item/28-11-2022-who-recommends-new-name-for-monkeypox-disease>; 2022).
2. Sayers EW, et al. Database resources of the national center for biotechnology information. *Nucleic Acids Res.* 2022;50:D20-d26.
3. Prevention, C.f.D.C.a. (2022).
4. Isidro J, et al. Phylogenomic characterization and signs of microevolution in the 2022 multi-country outbreak of monkeypox virus. *Nat Med.* 2022;28:1569–72.
5. McCollum AM, Damon IK. Human Monkeypox. *Clin Infect Dis.* 2013;58:260–7.
6. Happi C, et al. Urgent need for a non-discriminatory and non-stigmatizing nomenclature for monkeypox virus. *PLoS Biol.* 2022;20:e3001769.
7. Tyson, JR et al. Improvements to the ARTIC multiplex PCR method for SARS-CoV-2 genome sequencing using nanopore. *bioRxiv* (2020).
8. Pater AA, et al. High throughput nanopore sequencing of SARS-CoV-2 viral genomes from patient samples. *J Biol Methods.* 2021;8:e155.
9. Quick J, et al. Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. *Nat Protoc.* 2017;12:1261–76.
10. Aksamentov I, Roemer C, Hodcroft E, Neher R. Nextclade: clade assignment, mutation calling and quality control for viral genomes. *J Open Source Softw.* 2021;6:3773.
11. mpvx-lineages/lineage-designation (<https://github.com/mpvx-lineages/lineage-designation>; 2022).
12. Xu WK, Byun H, Dudley JP. The role of APOBECs in viral replication. *Microorganisms.* 2020;8:1899.
13. Cervantes-Gracia K, Gramalla-Schmitz A, Weischel J, Chahwan R. APOBECs orchestrate genomic and epigenomic editing across health and disease. *Trends Genet TIG.* 2021;37:1028–43.
14. Hatcher EL, Wang C, Lefkowitz EJ. Genome variability and gene content in chordopoxviruses: dependence on microsatellites. *Viruses.* 2015;7:2126–46.
15. Quince C, Walker AW, Simpson JT, Loman NJ, Segata N. Shotgun metagenomics, from sampling to analysis. *Nat Biotechnol.* 2017;35:833–44.
16. Davis JJ, et al. Analysis of the ARTIC version 3 and version 4 SARS-CoV-2 primers and their impact on the detection of the G142D amino acid substitution in the spike protein. *Microbiol Spectr.* 2021;9:e0180321.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary Data

Global Mpox Lineage Discovery and Rapid Outbreak Tracking with Nanopore Sequencing

Michael S. Bosmeny¹, Adam A. White^{1,†}, Adrian A. Pater^{2,†}, Jennifer Crew³, Joshua Geltz³, and Keith T. Gagnon^{1,2,*}

¹ Department of Biochemistry and Molecular Biology, Southern Illinois University School of Medicine, Carbondale, IL, USA.

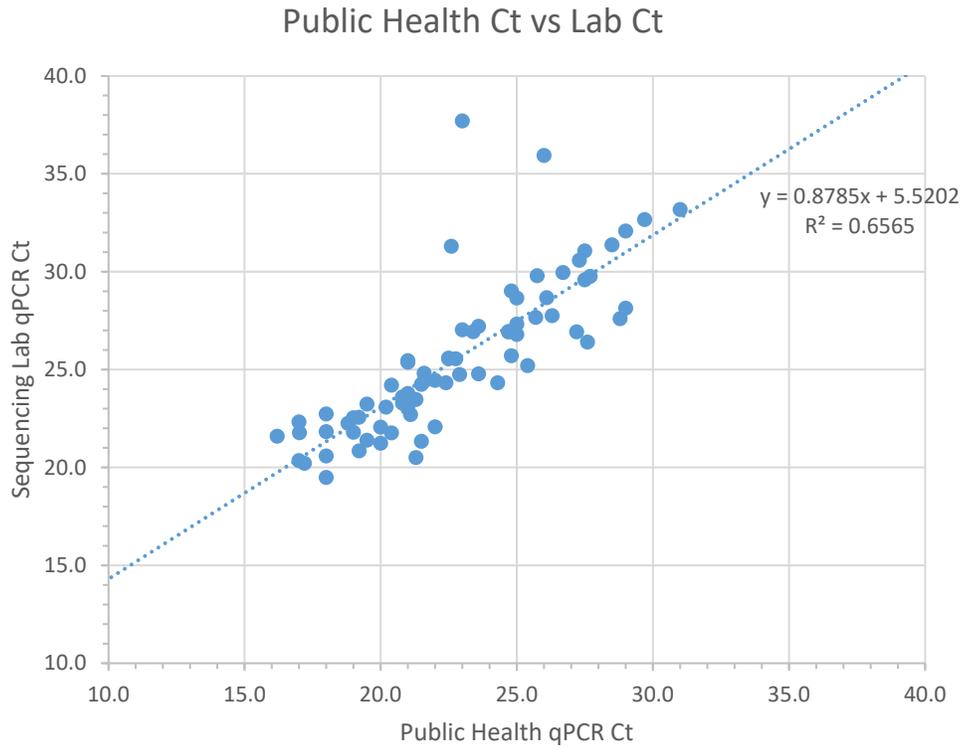
² Department of Chemistry and Biochemistry, Southern Illinois University, Carbondale, IL, USA.

³ Illinois Department of Public Health, Springfield, IL, USA.

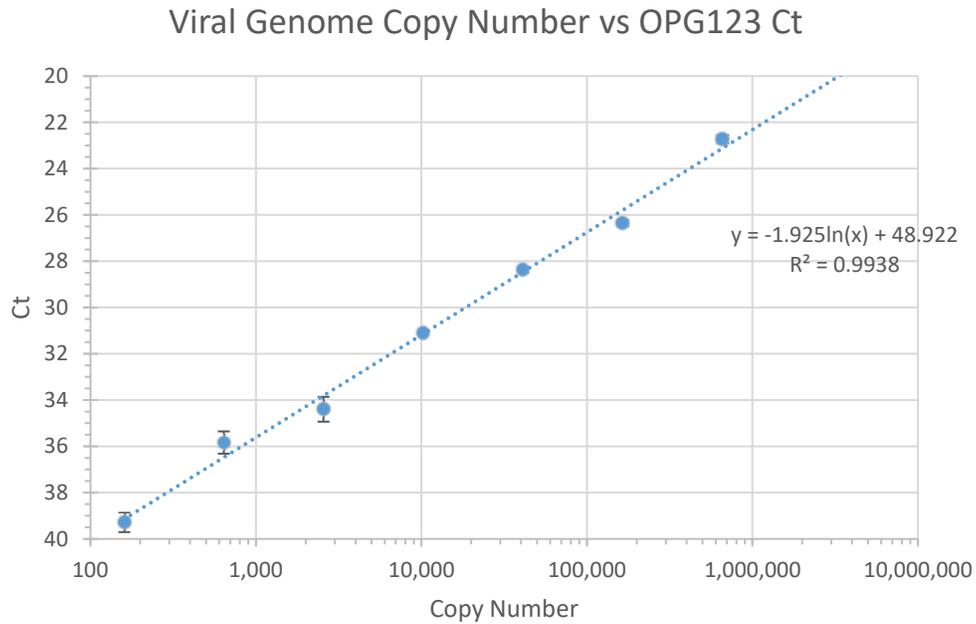
[†] These authors contributed equally.

* Corresponding Author: ktgagnon@siu.edu

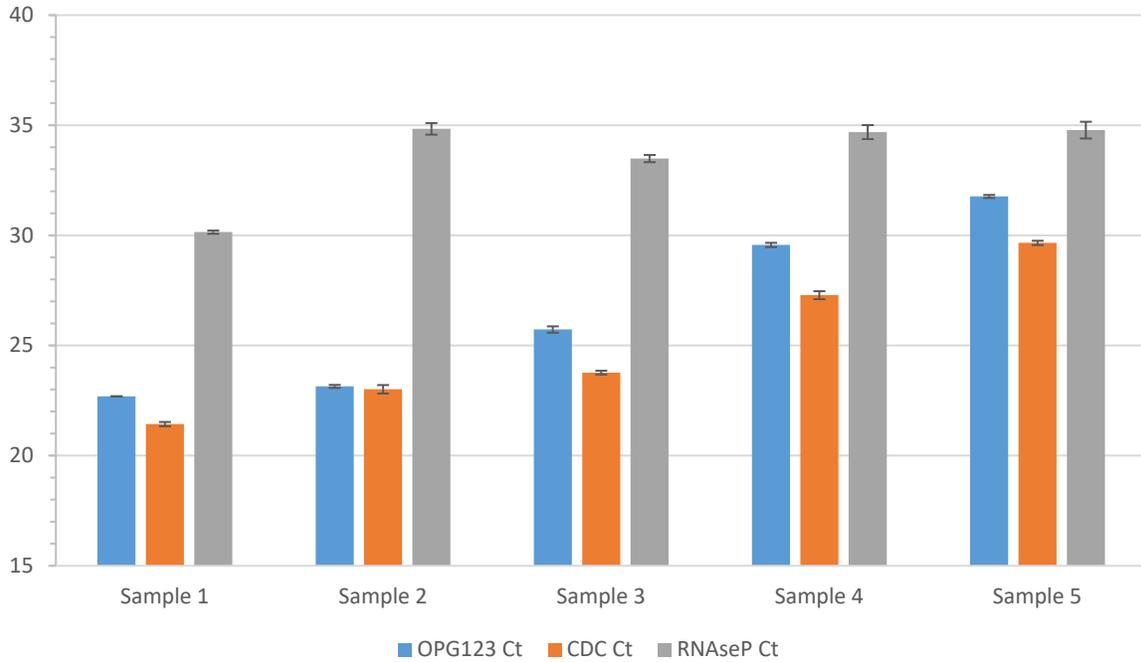
Supplementary Figures



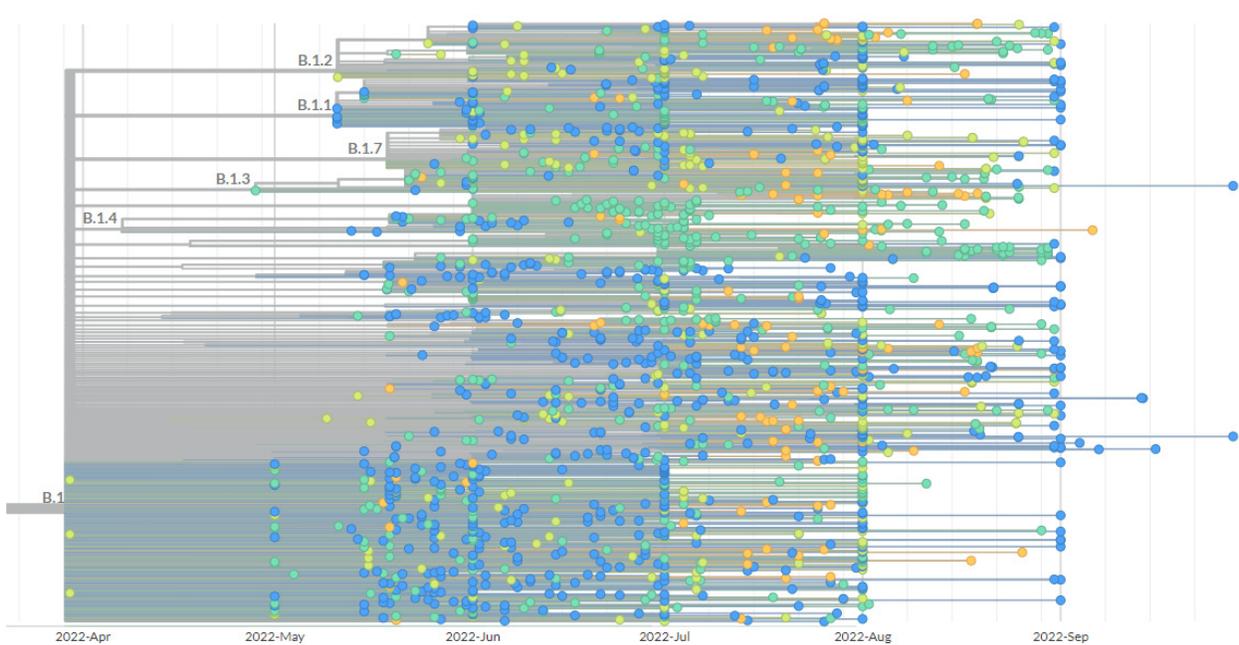
Supplementary Figure S1: Comparison of qPCR Ct values from Illinois Department of Public Health and those generated using OPG123 qPCR probes.



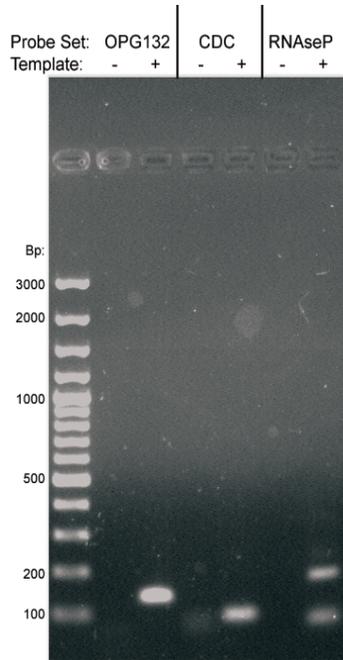
Supplementary Figure S2: Comparison of number of copies of a DNA gBlock and the qPCR Ct values generated from them using the OPG123 qPCR probe set.



Supplementary Figure S3: Ct values for five different mpox samples, using three different qPCR probes. Each sample/probe combination was replicated three times.

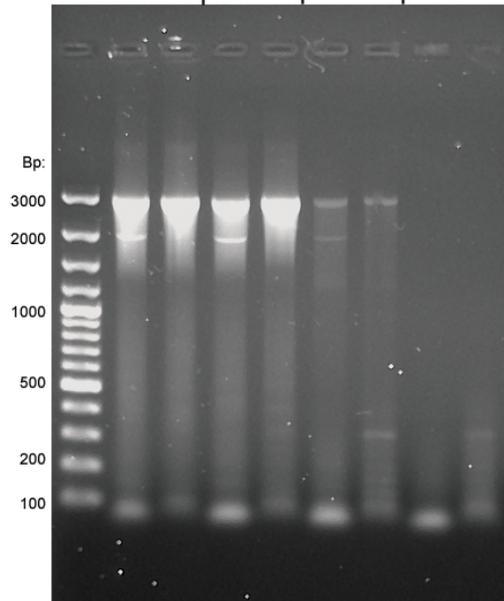


Supplementary Figure S4: Phylogenetic tree showing Clade IIb, Lineage B.1 mpox sequences. Color of each sequence indicates status of indicated deletion starting at base 179077. Teal: no deletion. Lime-green: Deletion to base 179094 (17 bases removed). Blue: Deletion to base 179148 (71 bases removed). Orange: Deletion to base 179175 (98 bases removed). Each lineage follows no clear pattern in deletions, indicating this could be an artifact of sequencing method, rather than actual deletion.



Supplementary Figure S5: 1% agarose gel showing the positive and negative control results for three qPCR probe sets. Ladder size markers are shown on the left.

Template:	IDPH-22		IDPH-42		IDPH-52		No template	
Pool:	1	2	1	2	1	2	1	2
Coverage:	99.74%		99.42%		99.95%		-	



Supplementary Figure S6: 1% agarose gel showing three randomly-selected multiplex PCR samples, along with a negative control (no template). Final sequence coverage of each sample is also shown. Ladder size markers are shown on the left.

Working Protocols

Working Protocol 1 – DNA Extraction

Introduction

The first step in the sequencing of Mpox virus genomes is extracting viral DNA from positive Mpox samples. DNA is extracted from inactivated samples in lysis solution using magnetic beads.

Materials & Reagents

Note: All reagents were estimated for 96 samples. Omega Mag-Bind® Viral DNA/RNA 96 Kit is used for DNA extraction.

Reagents

- Viral Transport Media (VTM, 17.28 mL)
- Mag-Bind Bead Solution CNR (144 µL)
- 100 % Isopropanol (20.16 mL)
- Nuclease-free water (4.8 mL)

Materials

- 96-Well Deep-Well PCR Plate
- Bio-Rad Microseal B
- 70% Ethanol

Equipment

- BSL-2 Hood
- Full PPE
- Thermomixer
- 96-well magnetic platform

Procedure

Caution: This is the step in the protocol when infection is possible. All steps are to be performed in the BSL-2 biosafety cabinet. Full PPE should be worn during the entire extraction process. Samples should only be handled by authorized personnel.

Preparing Working Stock Lysis Buffer

1. If samples are not already inactivated by addition of Lysis Buffer, create a working stock
2. to supply 180 µL/sample.
3. Remove Carrier RNA from the -20°C freezer and allow to thaw.
4. Working stock lysis buffer is composed of 180 µL TNA Lysis Buffer and 0.75 µL carrier RNA.

NOTE: We developed a homemade lysis buffer as a substitute for the TNA lysis buffer, which allows us to send pre-filled vials to testing facilities and send us back inactivated sample. The homemade lysis buffer performed similarly to TNA lysis buffer (data not shown). HM lysis buffer recipe: 100 mM sodium citrate (pH 6), 5 M guanidine isothiocyanate, 4% Triton X-100, 2 mM EDTA, 40 µg/mL purified (*Torula*) yeast tRNA.

Inactivating Samples

1. Aliquot 150 μL of sample in VTM into each well of a 96-well deep-well plate.
2. Add 180 μL of Working Stock Lysis Buffer to each well with sample.
 - a. If receiving frozen viral samples pre-mixed with homemade lysis buffer from testing facilities, allow samples to come to room temperature in biosafety cabinet and transfer to deep-well plate.
3. Seal the plate with a Bio-Rad Microseal B clear plate sealer.
4. Place the sealed plate on a thermomixer. Allow to mix at room temperature for 1 minute at 1100 RPM.

Binding DNA to Mag-Beads

1. Remove the Mag-Bind bead solution from 4°C storage and thoroughly vortex to resuspend beads.
2. Aliquot 211.5 μL of Mag-bind solution into each well of the plate, composed of 210 μL of 100 % isopropanol and 1.5 μL Mag-Bind beads. Mix solution by inversion and shaking of the tube before aliquoting. Note: Failing to mix can cause the beads to settle and unequally distribute, impacting purification yields.
3. Reseal the plate with Bio-Rad seal.
4. Place sealed-plate on thermomixer, and shake at RT for 10 minutes at 1100 RPM.

Separating and Washing Beads

1. After removing the plate from the thermomixer, place the plate in a 96-well magnetic platform designed to accommodate deep-well plates. Wait for the beads to be attracted to the magnet and cleared from the solution, approximately 10 minutes.
2. After solution is clear, carefully discard the supernatant, leaving beads untouched.
3. Wash the beads with VHB Buffer:
 - a. Remove plate from magnetic platform and added 262.5 μL of VHB Buffer per well.
 - b. Re-seal plate and place on thermomixer. Mix at RT for 1 minute at 1100 RPM.
 - c. Place on magnetic platform for approximately 7 minutes to clear beads from solution.
 - d. Carefully discard supernatant.
4. Wash the beads with SPR Wash Buffer:
 - a. Remove plate from magnetic platform and added 262.5 μL of SPR Wash Buffer per well.
 - b. Re-seal plate and place on thermomixer. Mix at RT for 1 minute at 1100 RPM.
 - c. Place on magnetic platform for approximately 7 minutes to clear beads from solution.
 - d. Carefully discard supernatant.
5. Wash the beads a second time with SPR Wash Buffer:
 - a. Remove plate from magnetic platform and added 262.5 μL of SPR Wash Buffer per well.
 - b. Re-seal plate and place on thermomixer. Mix at RT for 1 minute at 1100 RPM.
 - c. Place on magnetic platform for approximately 7 minutes to clear beads from solution.
 - d. Carefully discard supernatant.

Elution of DNA from Mag-Beads

1. After second wash, be sure to remove all remaining liquid from wells. A brief spin of the deep-well plate in a centrifuge to pool all liquids, followed by several more minutes on the magnetic platform to re-clear the magnetic beads is suggested.

2. Leave plate open inside biosafety cabinet for approximately 5 minutes to allow Mag-Bind particles to air dry. Do not over-dry Mag-Bind particles.
3. Remove deep-well plate from magnetic platform and add 50 μ l of nuclease-free water. Alternately, Elution buffer (20 mM Tris, pH 7.2, 1 mM EDTA) can be used instead for more stable storage of extracted DNA samples.
4. Reseal plate and place on the thermomixer. Mix at room temperature for 10 minutes at 1100 RPM.
5. Place on magnetic platform for 10 minutes to clear beads from solution.
6. Transfer elution to new 0.3 mL qPCR plate.
7. Seal plate with a fresh seal and label appropriately.
8. Store at 4°C until ready for next step for same-day workflow. For long-term storage, store at -80°C.

Working Protocol 2 – Viral Load estimation by qPCR

Introduction

qPCR is used to confirm positive samples and determine which samples can result in full or nearly full genomes for the workflow. Any sample with a cycle threshold (C_t) value exceeding a threshold C_t will not proceed through the workflow. This reduces cost and hands-on time. In addition to chances of genome retrieval, the C_t value provides insight on the viral load.

Materials & Reagents

Note: All reagents & materials are estimated for 96 samples, scale as needed.

Reagents

- Nuclease-Free water (318 μ L)
- 2X IDT PrimeTime (1060 μ L)
- 10 μ M N2 Forward Primer (106 μ L)
- 10 μ M N2 Reverse Primer (106 μ L)
- 2.5 μ M N2 Probe (106 μ L)

Materials

- Bio-Rad 96 Well PCR Plate
- 96-Well Plate Seals

Equipment

- qPCR Instrument (Bio-Rad CFX-96)

Procedure

Master Mix Assembly

Note: qPCR should be done on a clean bench to avoid any chance of contamination. Therefore, decontaminate your bench with 70% ethanol before beginning.

1. Thaw 2X PrimeTime, 10 μ M Forward Primer, 10 μ M Reverse Primer, and 2.5 μ M Probe
2. Gently flick, and spin down briefly before placing on ice.
Note: Avoid prolonged exposure of reference dye to light. Aluminum foil can be used to reduce exposure to light.
3. Prepare the master mix in a 2 mL microcentrifuge tube by combining the reagents in the following order, stopping after the 2.5 μ M Probe.

Reagents	1x	96x
Nuclease-Free Water	3 μ L	318 μ L
2X PrimeTime	10 μ L	1060 μ L
10 μ M Forward Primer	1 μ L	106 μ L
10 μ M Reverse Primer	1 μ L	106 μ L
2.5 μ M Probe	1 μ L	106 μ L
Sample DNA	4 μ L	-
Final Volume	20 μ L	1696 μ L

4. Retrieve the DNA plate from 4°C storage. Vortex and spin briefly. Place the plate directly on ice.
5. From the DNA plate, add 4 µL of cDNA to each well of the 96-Well qPCR plate, corresponding with their position on the plate layout.
6. Add 16 µL of the master mix to each well. Mix by pipetting up and down, gently.
7. Seal the PCR plate. Vortex then spin down the plate briefly.
8. Reseal the DNA extract plate store at 4°C or on ice until ready to use again.

qPCR Program

Run the following thermocycler conditions:

Step	Temp	Time
1 - Polymerase activation	95°C	3 min
2 - Denaturation	95°C	10 sec
3 - Annealing	55°C	30 sec
4 - Plate Read	Read plate (Cy5)	
Repeat steps 2-4 for 39 cycles.		

This program will run for ~ 1 hour.

Data Retrieval

Set the baseline threshold at 200 and export the C_t values. Proceed with future steps with samples that have adequately low C_t values. (Typically C_t 32 or lower.)

Working Protocol 3 – ARTIC-Style PCR Amplification

Introduction

To have enough material to obtain sufficient reads during sequencing, DNA is exponentially amplified using PCR. This multiplex amplification protocol utilizes two pools of ARTIC-style primers which contains approximately 70 primers, each, to produce ~3000 bp amplicons.

Materials & Reagents

Note: All reagents & materials are estimated for 96 samples, scale as needed. Primer scheme: https://github.com/gagnonlab/artic-mpxv/blob/main/Mpox_Sequencing_Primers.xlsx

Reagents

- NEB 5X Q5 Reaction Buffer (Cat #: M0493L) (960 µL)
- NEB Q5 Hot-Start High-Fidelity Polymerase (Cat #: M0493L) (48 µL)
- 10 µM ARTIC-MPXV Primer Pool 1 Mix (IDT) (240 µL)
- 10 µM ARTIC-MPXV Primer Pool 2 Mix (IDT) (240 µL)
- NEB 10 mM dNTP Mix (Cat #: N0447L) (96 µL)

Materials

- 2x 96-Well PCR Plate

Equipment

- Thermocycler

Procedure

Plate Setup

Create a new PCR layout that removes any sample with a C_t value greater than the chosen threshold. This layout will be utilized in future PCR steps.

Master Mix Assembly

PCR is extremely susceptible to contamination, work on a clean bench and wipe working areas with 70% ethanol. It is also highly advised to include a negative control (water only) beside the DNA samples and carry it forward to sequencing to determine the extent of contamination.

1. Thaw 5X Reaction Buffer, 10 mM dNTP mix, and Primer Pools 1 & 2. Mix at room temperature by vortexing, pulse centrifuge using a benchtop centrifuge and place on ice.
2. Dilute the 10 mM dNTP mix to 2.5 mM.
3. Gently flick the Q5 High-Fidelity Polymerase multiple times and immediately place on ice.
4. Dilute 100 µM primer pools in nuclease free water, to generate 10 µM primer stocks and place on ice.

Note: The MPXV primers are used at a final concentration of 12.5 nM per primer.

Our initial sequencing observed a low yield in several amplicons. To overcome this, we spike in primers for these amplicons at 2x in the 10 µM primer set. These primers are for amplicon 43 in pool 1 and amplicons 18, 20, 38, 40, and 52 in pool 2.

5. Prepare the pool 1 and pool 2 master mixes:

Reagent	1x	96x
Nuclease-Free Water	11.25 µL	1,080 µL

5X Q5 Reaction Buffer	5 μ L	480 μ L
Pool 1 or 2 10 μ M Primer Mix	2.5 μ L	240 μ L
2.5 mM dNTP Mix	2 μ L	192 μ L
Q5 High-Fidelity Polymerase	<u>0.25 μL</u>	<u>24 μL</u>
Sub Total	21 μL	2016 μL

- Retrieve sealed PCR plate of extracted DNA samples. Spin down the plate briefly. Place on ice.
- Transfer 4 μ L of DNA for each sample that meets the Ct criteria to wells corresponding to the sample's new plate position, for both Pool 1 and Pool 2 plates. The rest of the extracted DNA pcr plate can be stored at -20°C in case another reaction is desired.
- Add 21 μ L of Pool 1 Master Mix and Pool 2 Master Mix to the wells containing samples in plate 1 and plate 2, respectively.
- Seal the PCR plate, making sure that each well is tightly sealed to prevent evaporation.
- Vortex gently and spin the plate down.
- Place plates in thermocyclers and run

Thermocycle Program

Step	Cycle Step	Temp	Time	Number of Cycles
1	Heat activation	98°C	0:30	1
2	Denaturation	94°C	0:16	21
3	Annealing / Extension (Touchdown: -0.1 °C each cycle)	65-63°C	8:00	21
4	Denaturation	94°C	0:16	16
5	Annealing / Extension	63°C	8:00	16
6	Extension	72°C	4:00	1
7	Hold	4°C	∞	

This program will run for ~ 5.5 hours.

- After PCR is complete, it is recommended to confirm a successful reaction by running a selection of each of Pool 1 and Pool 2 samples, as well as Pools 1 and 2 of the negative control, on a 1% agarose gel. A lack of 3 kb bands in all lanes indicates something wrong with the PCR reaction, while bands in every well, including the negative control, could indicate major contamination that must be remedied before proceeding. Samples from DNA extracts with higher C_ts (typically in the C_t 30 – 32 range) may produce weaker 3 kb bands. These can still be sequenced, but may require longer sequencing times.

Working Protocol 4 - Bead Clean-up of PCR Products

Introduction

The clean-up is performed by using AMPure XP SPRI beads to remove leftover PCR contaminants (dNTPs, salts, primers, etc.) from the two previously performed PCR reactions. Leftover contaminants may lead to poor DNA library end-prep and barcoding efficiency. The advantage of using SPRI beads is their ability to size select DNA depending on the ratio of bead to sample volume used. Expected recovery is 60-80%.

Materials & Reagents

Note: All reagents & materials were estimated for 96 samples.

Reagents

- AMPure XP beads (4 mL)
- 100% Ethanol (40 mL)
- Omega Elution Buffer (EB) (2.9 mL)

Materials

- 96-well PCR plate
- 96-well magnetic separator
- Multichannel pipettes
- Plate Seal

Procedure

You will begin this procedure with two 96-well plates containing the PCR product from the previously performed PCR reaction. The plates will be referred to as PCR pool 1 plate and PCR pool 2 plate.

Note: Amplicon clean up should be done on a clean bench to avoid any chance of contamination. Therefore, decontaminate your bench with 70% ethanol.

Pooling of PCR Products

CAUTION: Be very careful when pooling PCR products. Avoid cross-contaminating neighboring wells with small droplets.

1. Spin down each PCR product plate of pool 1 and pool 2 at 2,000 RPM for 30 seconds. Remove seal on both plates.
2. Using an 8-channel P200 pipette, transfer all 25 μ L from column 1 of PCR pool 2 plate to the corresponding wells of column 1 of the PCR pool 1 plate (Fig 1) and mix by pipetting up and down 3 times.

AMPure Bead Addition

3. Resuspend the stock AMPure XP Beads by vortexing for 20 seconds. The solution should be a homogenous brown color.
4. Using an electronic pipettor, remove 4 mL of AMPure XP beads and place it in a clean trough on ice. Place the stock beads back at 4°C.
5. Using an 8-channel P200 multichannel pipette set to 200 μ L, mix the beads in the trough by pipetting up and down 5 times.

6. Immediately, use the multichannel pipette to transfer 37.5 μL of beads from the trough to each well of a column of the 96-well PCR plate. Repeat this with a fresh set of tips and each following column.
7. Between each transfer, mix the beads by pipetting up and down 2 times to ensure the beads do not settle.
8. Seal plate.
9. Incubate the plate at room temperature for 15 minutes to allow beads to bind to genomic DNA.

Purify the Pooled Reactions

CAUTION: Be careful not to touch the pelleted beads when removing the supernatant. Removal of beads will negatively impact yield.

10. After incubation, place 96-well plate on the magnet for 5 minutes. Check that all bead material has moved to the side of the well.
11. Carefully discard the supernatant without disturbing the beads.
12. Wash the beads with 200 μL of 80% ethanol.
13. Keep the plate on the magnet for 3 minutes.
14. Carefully discard the ethanol without disturbing the beads.
15. Repeat steps 3-5 (80% ethanol wash) for a second ethanol wash.
16. Centrifuge the 96-well plate at 2,000 RPM for 30 seconds to force any remaining ethanol wash to the bottom of the well.
17. Place the 96-well plate on the magnetic platform.
18. Use a P10 single channel pipette to aspirate any residual ethanol.
19. Let plate air-dry for 1-2 minutes to remove last traces of ethanol. Do not over-dry the beads.
20. Remove the plate from the magnetic platform and place onto the benchtop.
21. Elute DNA by adding 30 μL of Omega elution buffer to each of the well. Mix by pipetting up and down five times.
22. Cover the plate and incubate for 5 minutes at room temperature.
23. Place the plate on the magnetic platform for 5 minutes.
24. Carefully remove and retain 30 μL of the eluate containing the DNA library per well into a new 96-well PCR Plate. Retain the same plate layout.
25. Seal and label the new plate containing 30 μL of the eluate. Store at 4°C until ready for next step.

Working Protocol 5 – Qubit Quantification

Introduction

The Qubit dsDNA HS assay uses fluorescent dyes that bind specifically to dsDNA providing a more accurate quantification. This assay relies on a 2-point curve by reading the two standards every time measurements are taken. The assay is designed to quantitate 0.2–200 ng/ μL of DNA.

Materials & Reagents

Reagents

- Qubit dsDNA HS Buffer
- Qubit dsDNA HS Reagent (Dye)
- Qubit dsDNA HS Standard #1
- Qubit dsDNA HS Standard #2

Materials

- Qubit assay tubes
- Qubit 2.0 Fluorometer

Procedure

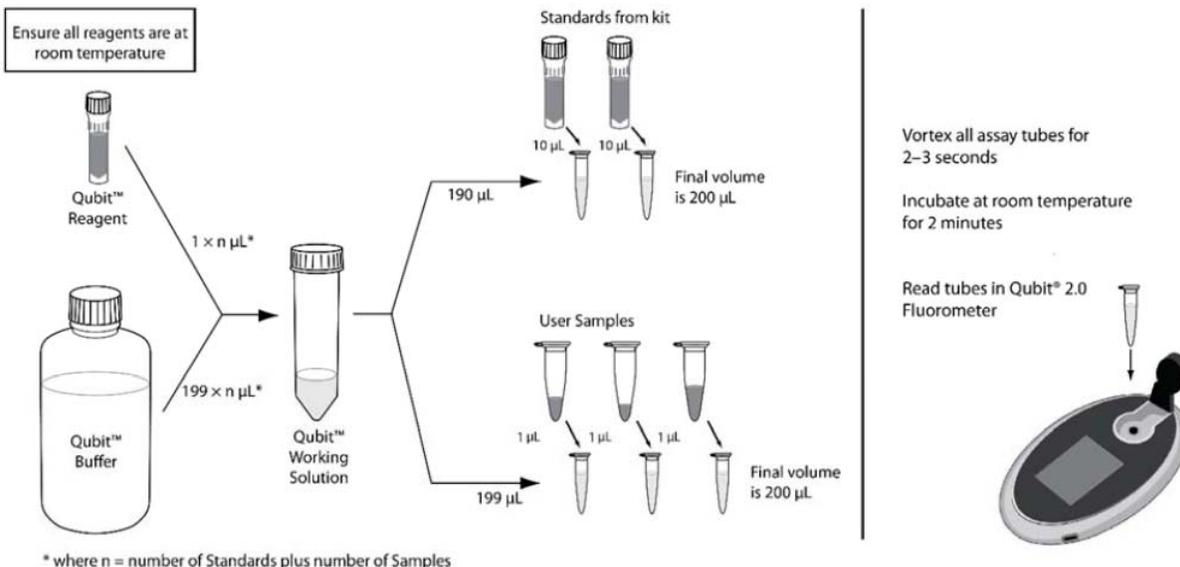


Figure WP1. Workflow of Qubit quantification

Qubit Reaction Preparation

1. Remove the Qubit standards and reagent (dye) from 4°C and let them equilibrate to room temperature.
2. Determine the total number of samples + two standards.
3. Set up two Qubit Assay tubes for the two standards and one tube for each sample to be quantified. Label the tube lids.
4. Make sufficient Qubit working solution (WS) in a single microcentrifuge tube for the total number of reactions (standards and samples), by combining 1 μL Qubit dsDNA reagent (dye) to 199 μL Qubit dsDNA buffer for each reaction. For example, for 96 samples, prepare enough working solution (WS) for a total of 100 reactions:

	<u>1X</u>	<u>100X</u>
dsDNA Buffer	199 μ L	1,990 μ L
dsDNA Reagent	<u>1 μL</u>	<u>100 μL</u>
Total working solution	200 μ L	2,000 μ L

5. Mix the Qubit working solution (WS) by vortexing for ~4 seconds and briefly centrifuge.
6. Add 190 μ L of Qubit working solution (WS) to each of the Qubit assay tubes used for standard 1 and standard 2.
9. Add 199 μ L of Qubit working solution (WS) to each of the Qubit assay tubes used for samples to be quantified.
10. Add 10 μ L of each Qubit standard to the appropriate tube, mix by vortexing for ~4 seconds and briefly centrifuge.
11. Add 1 μ L of each DNA sample to the appropriate tube, mix by vortexing for ~4 seconds and briefly centrifuge.
12. Incubate tubes at room temperature for 2 minutes.

Qubit Quantification

13. On the Home screen of the Qubit Fluorometer, press **DNA**, then select **dsDNA High Sensitivity** as the assay type.
14. The "Read New Standards?" screen is displayed. Press **Yes**.
15. Insert Standard #1 into the Sample Chamber, close the lid and press **Read**.
16. Insert Standard #2, close the lid and press **Read**. Calibration of Qubit is now complete.
17. Choose **Sample** to go to the sample screen.
18. Insert a sample into the sample chamber, close the lid and press **Read**.
19. A result will display on the screen. Press **Calculate Stock Conc**.
20. Using the volume roller wheel, select **1 μ L**.
21. Change the units in which the original sample concentration is displayed, select **ng/ μ L**.
22. Record the concentration value for the sample.
23. Repeat for all DNA samples by pressing **Read** once a new sample has been inserted.

Working Protocol 6 – RAPID Barcoding

Introduction

The barcoding reaction is done using the Nanopore Rapid Barcoding Kit 96 (SQK-RBK110.96). The barcodes are added via a transposase that also fragments the dsDNA. Due to the size of the Mpx genome, it is recommended to only sequence 20 genomes per flow cell. After barcoding, samples are pooled and cleaned up using beads.

Materials & Reagents

Reagents

- Rapid Barcoding Plate

Materials

- 96-well PCR plate
- Plate seal
- Thermocycler

Procedure

Note: Barcoding is done in a 10 μ L reaction, utilizing 2.5 μ L of Rapid Barcoding kit reagent per well. This allows for up to 7.5 μ L of sample per reaction. Depending on number of samples, concentration of samples, and available time, it may be worth using different volumes of each sample in its individual reaction, making up remaining volume with nuclease-free water, in order to have each reaction have a similar nanogram amount of DNA per reaction. This will result in a more consistent sequencing experience.

1. Thaw the rapid barcode plate at room temperature.
2. Once thawed, centrifuge the rapid barcode plate briefly to ensure all reagent is pooled at the bottom of the well. Ensure that there is sufficient of each barcode being used per sample.
3. Transfer 7.5 μ L of each sample (see above note) from Bead-Cleaned PCR Product plate (S4 protocol) to a new 96-well PCR plate. Save remainder of 30 μ L in Bead-Cleaned PCR Plate by resealing and storing at -20°C.
4. Add 2.5 μ L of the appropriate barcode reagent, 1-96, to each well of new PCR plate with sample in it. Mix by pipetting up and down 3 times.
5. Seal plate.
6. Gently vortex the plate to further mix samples. Centrifuge the plate at 2000 RPM for 1 minute to collect all liquid at the bottom of the well after vortexing.
7. Place the PCR plate in a thermocycler and run the following cycles:

Thermocycle Program

Step	Cycle Step	Temp	Time
1	Barcoding	30°C	3:00
2	Inactivation	80°C	2:30
3	Hold	4°C	∞

This program will run for ~ 6 minutes.

Working Protocol 7 – RAPID Sample Pool, Clean, & Adapter

Introduction

After barcoding, all samples are pooled, cleaned, and then eluted into a single volume. This volume is quantified to determine the correct amount to load into a MinION flow cell. The RAPID adapter is then allowed to attach.

Materials & Reagents

Reagents

- AMPure XP beads (4 mL)
- Elution Buffer
- 80% Ethanol

Materials

- Magnetic rack for Eppendorf tubes
- Qubit dsDNA HS Assay Kit
- Rotation mixer

Procedure

Pooling and Bead Cleanup

1. Remove Qubit dsDNA HS reagents from 4°C and let them warm to room temperature.
2. Retrieve barcoding 96-well plate from S6 protocol.
3. Briefly centrifuge plate to pool all liquids at bottom of wells.
4. Unseal plate and pool all samples that are being loaded into a flow cell into an Eppendorf tube.

Note: Generally, approximately 20 samples can be easily loaded onto a single flow cell. If a full 96-well plate of samples has been prepared, divide them into multiple Eppendorf tubes of 19-20 samples each.
5. Resuspend AMPure XP beads by vortexing for 30 seconds.
6. Add 1x the volume of AMPure XP beads to the tube. Mix by pipetting up and down 5 times.
7. Incubate tube at room temperature for 10 minutes in a rotator mixer (Hula mixer).
8. Pulse-centrifuge the tube and place in magnetic rack for 5 minutes, or until solution is clear and beads have all collected on the side of the tube.
9. Carefully aspirate and discard the supernatant without disturbing the beads.
10. Add 500 µL of 80% ethanol while keeping the tube on the magnet.
11. Wait 3 minutes for the solution to clear and beads to collect on the side of the tube.
12. Remove ethanol and discard.
13. Repeat steps 10-12 for a second ethanol wash.
14. Briefly centrifuge tube to pool any remaining ethanol.
15. Place tube back on magnetic rack for an additional 3 minutes.
16. Use a P10 to aspirate any residual ethanol.
17. Let tube air-dry for 1-2 minutes to remove last traces of ethanol. Do not over-dry the beads.
18. Remove the tube from the magnetic rack and place onto the benchtop.
19. Elute DNA by adding 30 µL of Elution Buffer to each of the well. Mix by pipetting up and down five times.
20. Incubate at room temperature for 10 minutes.
21. Place the tube back on the magnetic rack for 5 minutes.

22. Transfer the 30 μL of DNA in elution buffer to a new Eppendorf tube. Keep on ice until ready to sequence.

Rapid Adapter Addition

23. As before, use the Qubit quantification protocol (S5) to quantify the $\text{ng}/\mu\text{L}$ in the resulting elution tube. As the amount will most likely be above 200 $\text{ng}/\mu\text{L}$, it is recommended to first make a 1:10 dilution of the sample in water and use this for the quantification.
24. After quantification, calculate how many μL of sample needed to load 1500-2000 ng of dsDNA into the flow cell. Make up the remaining volume, up to 11 μL , in elution buffer.
25. Add 1 μL of Rapid Adapter F (RAP-F) to the 11 μL of barcoded DNA. Mix well by pipetting up and down 5 times.
26. Incubate at room temperature for at least 10 minutes to allow adapter to attach.

Working Protocol 8 – MinION Loading and Running

Introduction

The MinION flow cell is kept filled with a storage buffer to maintain product integrity. This must be flushed out and replaced. The pooled, barcoded DNA is then loaded and a sequencing run is started. Sequencing progress can be monitored in real-time using MinION software and RAMPART.

Materials & Reagents

Reagents

- Loading Beads II – 25.5 μ L
- Sequencing Buffer II – 37.5 μ L
- Flush Buffer – 1 mL
- Flush Tether – 30 μ L

Materials

- MinION
- Flow Cell

Procedure

Prepare Priming Mix

1. Thaw the Sequencing Buffer II (SBII), Loading Beads II (LBII), Flush Tether (FLT) and Flush Buffer (FB). Place on ice once thawed.
2. Mix the SBII, FLT, and FB tubes by vortexing and then return to ice.
3. Open the MinION Sequencer lid (Figure 7) and remove the configuration cell.
4. Slide the flow cell into the sequence, under the clip. Press down gently on the flow cell to ensure correct thermal and electrical contacts.
5. Create Priming Mix by adding 30 μ L of Flush Tether (FLT) into the new tube of Flush Buffer (FB). Label the FB tube to indicate FTL has been added. Mix the contents of the tube.

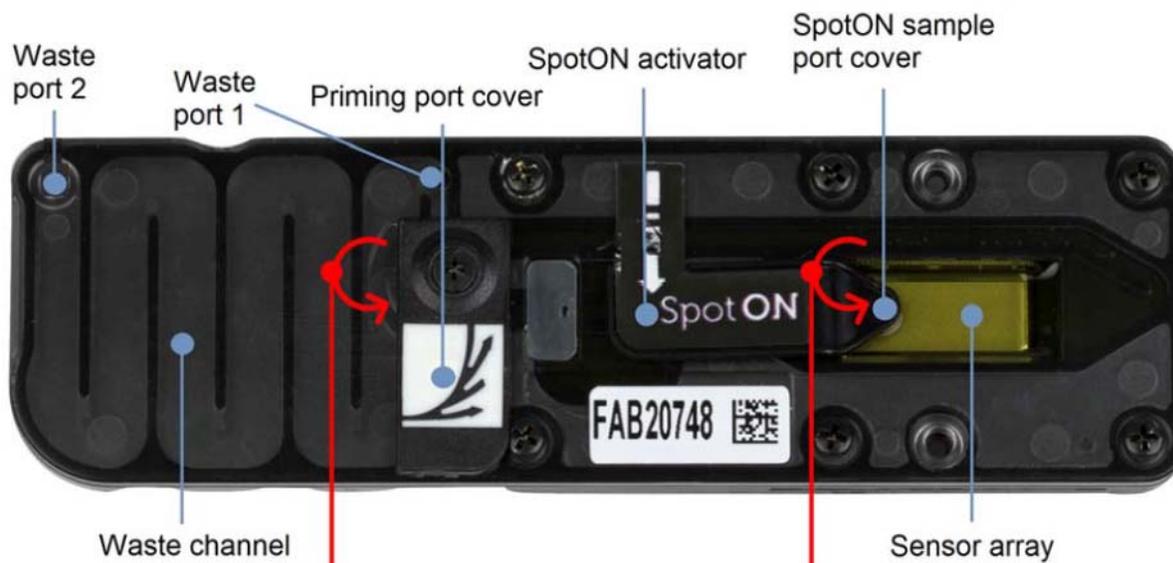


Figure WP2. MinION flow cell

Priming the Flow Cell

Caution: Throughout the following process it is essential that the sensor array remains submerged in buffer at all times. If an air bubble passes over any channels, those pores will be permanently damaged.

6. Open the priming port by sliding the cover clockwise so that the port is visible as shown in Figure S8.1.
7. Draw back a small volume of liquid from the priming port to remove any bubbles already present in the system:
 - a. Insert a P1000 pipette tip into the priming port
 - b. Adjust the P1000 volume upwards slowly until a small volume of liquid enters the pipette tip.
8. Slowly add 800 μL of Priming Mix to the priming port.
9. Close the priming port and allow flow cell to prime for five minutes. The Sequencing Mix can be prepared during this.

Prepare Sequencing Mix

Note: The Loading Beads II (LBII) tube contains a suspension of beads. These beads settle quickly. It is vital that they are mixed immediately before use by pipetting up and down.

10. In an Eppendorf tube, prepare the Sequencing Mix:

<u>Reagent</u>	<u>Volume</u>
Sequencing Beads II	37.5 μL
Loading Beads II	25.5 μL
Barcoded DNA library	<u>12 μL</u>
Total	75 μL

11. Mix contents of tube and pulse-centrifuge to pool.
12. Store on ice until priming of flow cell is complete.

Loading Flow Cell

13. Open the flow cell priming port by sliding the cover clockwise.
14. Gently lift the Spot-ON sample port cover to make the Spot-ON port accessible.
15. Slowly load another 200 μL of Priming Mix into the priming port. A droplet of buffer may rise from the Spot-ON port. Stop and wait for the droplet to recede before continuing to add the rest of the Priming Mix.
16. Resuspend the Sequencing Mix by gently pipetting up and down.
17. Add all 75 μL of Sequencing Mix to the flow cell via the Spot-ON sample port in a dropwise fashion. Each drop should flow into the port before adding another.
18. Close the Spot-ON port. Ensure the bung on the cover goes into the port hole.
19. Close the priming port.
20. Close the MinION lid.

Starting a Sequencing Run

21. Plug the MinION into the computer's USB port.
22. Open MinKNOW, the sequencing software for the MinION.
23. Ensure the software recognizes the flow cell.
24. Go to the Connection Manager window and click Start.
25. Name the experiment.

26. Click **Continue to Kit Selection**.
27. Select **SQK-RBK110.96** as the kit.
28. In the Run Options menu, leave at default settings.
29. In the Basecalling Menu, enable **Basecalling** and use the Options menu to configure the software for **Super-High Accuracy Basecalling**.
 - a. Also enable **Barcoding**. Use the Options menu to set **Mid-read barcodes** to On. Set **Override minimum mid barcoding score** to 50.
30. In the Output menu, pick an output location and seq FAST5 and FASTQ Reads per File to 1000.
31. Click Start Run.

Working Protocol 9 – Rampart Monitoring

Introduction

Rampart (<https://artic.network/rampart>) is a useful program to supervise the progression of a MinION sequencing run. It uses the demultiplexed FASTQ read files that result from MinION base-calling and aligns them to a genomic template. This allows an estimation of coverage for each barcode being sequenced and is useful for determining how long to run the MinION sequencer.

Materials

- Rampart can be installed using the directions found here: <https://artic-network.github.io/rampart/docs/installation.html>.
- Mpxv genomic template and amplicon location files are needed for Rampart to function properly. These are available here: <https://github.com/gagnonlab/artic-mpxv>. After downloading, put the rampart directory in a known location.
- Location of MinION sequencing FASTQ files must be known.

Procedure

1. Launch a console window from the Linux environment.
2. Activate the conda artic environment with the command:

```
conda activate artic-ncov2019
```

3. Launch rampart using the following syntax:

```
rampart --clearAnnotated --protocol [location of template file] --basecalledPath [location of FASTQ files]
```

For example:

```
rampart --clearAnnotated --protocol /home/gagnonlab/artic-mpxv/rampart/B1 --basecalledPath /var/lib/minknow/data/220815_mpxv/no_sample/20220815_1200_MN24287_FAP93592_f886cd67/fastq_pass/
```

4. Running this command will cause rampart to begin running in the console window. This produces a status window (Figure 8). Clicking on the **http://localhost:3000** link will open a browser window with detailed breakdowns of each barcode (Figure 9).
 - a. Each barcode has its own panel which shows what percentage of the genome has 1x, 20x, 100x, and 200x coverage. It also shows how the reads match up along the length of the genomic, allowing analysis of low-amplifying amplicons.
 - b. The **Report** button in the upper-right corner allows for numerical analysis of all barcodes at once: Average read length, coverage percentages, and percentage matching to the template genome.
 - c. For best consensus-finding results, it is recommended to get at least 100x coverage of as much of the genome, for all sequences before terminating the sequencing run. If most genomes have good coverage but one or two barcodes are severely lagging, it is possible to load more of just those barcodes on the flow cell without the other barcodes to allow targeted read generation.

```
gagnonlab@gagnonlab: ~  
  
-----  
RAMPART daemon running  
socket open on port 3001 for all data communication  
-----  
  
-----  
Serving built bundle at http://localhost:3000  
-----  
  
Found 0 annotated CSV files in ../gagnonlab/annotations/. FASTQs with the same  
filename as these will be ignored.  
Scanning folder /home/gagnonlab/sc2_seq_data/201118_sc2_Plate9/201118_sc2_demult  
lplex/ for FASTQs  
(basecalled files which exist here (or are created here by Minknow) will be anno  
tated and loaded)  
  
Initial scan for FASTQs complete. Found 262 files, 262 of which are unannotated  
(unprocessed).  
Watching for new files
```

Figure WP3. Rampart software.



Figure WP4. Rampart software output.

Working Protocol 10 – Consensus Sequence Calling

Introduction

Once the MinION sequencing run is complete, the individual called reads need to be aligned against a template genome and compiled into a consensus sequence. The ARTIC network's pipeline is useful for streamlining the consensus calling process.

Materials

- The ARTIC pipeline installation instruction can be found here: <https://artic.readthedocs.io/en/latest/installation/>.
- This pipeline requires a Mpxv-specific primer scheme and template genome in order to create a consensus sequence. A directory containing these files can be found at: https://github.com/gagnonlab/artic-mpxv/tree/main/primer_schemes/B1/V1.
- Two useful BASH files can be found at: <https://github.com/gagnonlab/artic-mpxv/blob/main/concatBarcodes>
<https://github.com/gagnonlab/artic-mpxv/blob/main/consensusCall>

Procedure

1. First, the FASTQ reads for each barcode need aggregated into one file that can be processed.
 - a. Launch the ARTIC environment:

```
conda activate artic-ncov2019
```

- b. An individual barcode can be aggregated using the following syntax:

```
artic guppyplex --directory [location of barcode FASTQ files] --skip-quality-check --output [name of file to aggregate FASTQ files into]
```

For example:

```
Artic guppyplex --directory /var/lib/minknow/data/220815_mpxv/no_sample/20220815_1200_MN24287_FAP93592_f886cd67/fastq_pass/barcode10/ --skip-quality-check --output barcode10.fastq
```

- c. To aggregate a series of barcodes in a batch fashion, it is possible to use a BASH script, such as the **concatBarcodes** script linked in the Materials section. The format is:

```
./concatBarcodes FASTQ_DIR SAVE_DIR BARCODE_NUM_START BARCODE_NUM_END
```

For example:

```
./concatBarcodes /var/lib/minknow/data/220815_mpxv/no_sample/20220815_1200_MN24287_FAP93592_f886cd67/fastq_pass/ ./ 10 40
```

The above commands would aggregate the barcode FASTQ directories for barcodes 10 through 40 into files `barcode10.fastq` through `barcode40.fastq`.

2. After aggregation, the artic pipeline can be utilized to apply Nanopolish to the reads and deliver a consensus sequence file. Nanopolish requires FAST5 files and a sequencing summary in addition to the aggregated FASTQ reads in order to produce the consensus sequence.

- a. If not already in it, launch the ARTIC environment:

```
conda activate artic-ncov2019
```

- b. An individual barcode can be consensus called with the artic pipeline command:

```
artic minion --normalise [number of reads for a given position on the genome to consider] --threads [number of CPU threads] --scheme-directory [location of mpox primer scheme] [name of Mpox primer scheme] --read-file [location of aggregated fastq file] [name of finished consensus fasta] --fast5-directory [location of FAST5 minION reads] --sequencing-summary [location of MinION sequencing run summary]
```

For example:

```
artic minion --normalise 200 --threads 12 --scheme-directory /home/gagnonlab/artic-mpxv/primer_schemes B1 barcode10 --read-file ./barcode10.fastq --fast5-directory /var/lib/minknow/data/220911_MPXV/no_sample/20220911_1348_MN31708_FAR89910_172c327d/fast5_pass/barcode10/ --sequencing-summary /var/lib/minknow/data/220911_MPXV/no_sample/20220911_1348_MN31708_FAR89910_172c327d/sequencing_summary_FAR89910_8b733801.txt
```

This will produce a file called `barcode10.consensus.fasta` in the directory the artic minion command was run from.

- c. To derive a series of consensus sequences in a batch fashion, it is possible to use a BASH script, such as the **consensusCall** script linked in the Materials section. The format is:

```
./consensusCall BARCODE_DIR FAST5_DIR SUMMARY_PATH  
BARCODE_NUM_START BARCODE_NUM_END
```

For example:

```
./concatBarcodes ./  
/var/lib/minknow/data/220911_MPXV/no_sample/20220911_1348_MN31708_FAR89910_172c327d/fast5_pass/  
/var/lib/minknow/data/220911_MPXV/no_sample/20220911_1348_M
```

```
N31708_FAR89910_172c327d/sequencing_summary_FAR89910_8b7338  
01.txt 10 40
```

The above commands would use the aggregated FASTQ files named barcode10.fastq through barcode40.fastq, along with matching FAST5 files to derive consensus sequences, one after another, named barcode10.consensus.fasta, etc.

Working Protocol 11 – Visualizing Consensus Sequences

Introduction

When generating consensus sequences, it is important to check the result to ensure that nothing has gone wrong in the sequencing or consensus calling. NextClade, a website run by the NextStrain community, was originally designed for analyzing SARS-CoV-2 genomic sequences, but has since expanded to include Mpox and Influenza strains as well.

If needed, there is also an offline command-line interface version of NextClade, for processing a large number of samples.

Procedure

1. Navigate to <https://clades.nextstrain.org/> to launch the NextClade web interface.
2. Select pathogen of interest. For current 2022 outbreak mpox samples, select **Human Monkeypox Lineage B.1**.
3. Supply Nextclade with FASTA-format sequences. This can either be done by dragging-and-dropping the file into the website, or by copying and pasting the FASTA as text.
4. Click the **Run** button to have the website analyze the samples.

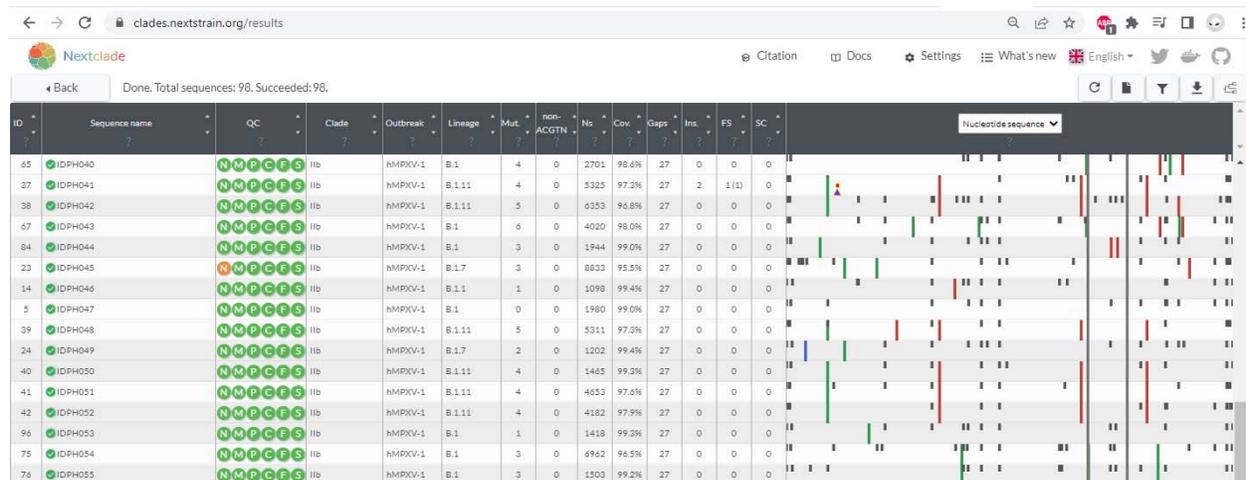


Figure WPS. NextClade software output.

Each column (Figure 10) provides different information about your sequence:

- Sequence name – name provided for sequence in FASTA file
- QC – General indications of quality of sequence. In general, green and yellow are acceptable, red is unacceptable. Hovering the mouse over each circle gives more details.
 - N circle – measurement of missing data. Based on the number of Ns in the sequence.
 - M circle – measurement of mixed sites. Will be flagged if more than 10 nucleotides have mixed states (such as R or Y), as this may be indicative of contamination.
 - P circle – measurement of private mutations. Based on the number of mutations present compared to the template sequence.
 - C circle – measurement of mutation clusters. Based on the clustering of mutations. If more than 6 mutations are within a 100 base range, the sequence will be flagged.

- F circle – measurement of frame shifts.
 - S circle – measurement of unusual stop codons.
- Lineage – What Nextstrain lineage the software determines the sequence to be. This label is reliable with good quality reads, but may be unreliable with too many missing nucleotides.
- Mut. – Number of mutations in the sequence compared to template genome. Hovering the mouse over a specific sequence's mutation column will generate a popup window with a list of nucleotide and amino-acid mutations.
- Non-ACGTN – Number of ambiguous, non-N nucleotides in the sequence. Hovering the mouse over a specific sequence's Non-ACGTN column will generate a list.
- Ns – Number of missing nucleotides in the sequence. Hovering the mouse over a specific sequence's Ns column will generate a list.
- Gaps – Number of intentional gaps in the sequence (-). Produced by deletions in the nucleotide code. Hovering the mouse over a specific sequence's Gaps column will generate a list.
- Ins. – Number of insertions in the sequence. Hovering the mouse of a specific sequence's Ins. Column will generate a list.
- FS – Number of frame shifts in the sequence. Hovering the mouse of a specific sequence's FS Column will generate a list.
- SC – Number of premature stop codons in the sequence. Hovering the mouse of a specific sequence's SC Column will generate a list.
- Sequence View – Color-coded visual representation of the mutations, gaps, and Ns in the sequence, as compared to template genome. Red, Mutation to A. Blue, Mutation to C. Yellow, mutation to G. Green, mutation to T. Dark grey, N. Light grey, gap. A visual representation of Mpox genes is provided below the sequences for comparison. Hovering the mouse over a specific color-coded bar will provide additional information about it.

In addition, there are several useful tool buttons at the top of the page:

- Filter – Funnel icon. Filter your sequence list based on various criteria, including a mutation at a specific nucleotide or amino-acid, the quality of the sequence, or the lineage it is categorized as.
- Download – Down pointing arrow icon. Allows the generation of a spreadsheet with all details of the analysis for later browsing.
- Show Tree – Phylogenetic tree icon. Places sequences on a phylogenetic tree for comparison with other sequences from the public sources.
- Settings – Gear icon. Adjust settings for each of the four quality control determinations.

Using these criteria, poor-quality sequences can be identified and removed from your FASTA file.

