Oct 20, 2022 Version 3

# Monkeypox virus multiplexed PCR amplicon sequencing (PrimalSeq) V.3

DOI
dx.doi.org/10.17504/protocols.io.5qpvob1nbl4o/v3

Nicholas F.G. Chen*[1], Luc Gagne*[2], Matthew Doucette[2], Sandra Smole[2], Erika Buzby[2], Joshua Hall[2], Stephanie Ash[2], Rachel Harrington[2], Seana Cofsky[2], Selina Clancy[2], Curtis J Kapsak[3], Joel Sevinsky[3], Kevin Libuit[3], Mallery I Breban[1], Chrispin Chaguza[1], Nathan D. Grubaugh[1], Daniel J. Park[4], Glen R. Gallagher#[2], Chantal B.F. Vogels#[1]

[1]Department of Epidemiology of Microbial Diseases, Yale School of Public Health; [2]Massachusetts Department of Public Health; [3]Theiagen Genomics; [4]Broad Institute, Cambridge, Massachusetts

Nathan D Grubaugh
Department of Epidemiology of Microbial Diseases, Yale Schoo...





**DOI: https://dx.doi.org/10.17504/protocols.io.5qpvob1nbl4o/v3**

**Protocol Citation:** Nicholas F.G. Chen*, Luc Gagne*, Matthew Doucette, Sandra Smole, Erika Buzby, Joshua Hall, Stephanie Ash, Rachel Harrington, Seana Cofsky, Selina Clancy, Curtis J Kapsak, Joel Sevinsky, Kevin Libuit, Mallery I Breban, Chrispin Chaguza, Nathan D. Grubaugh, Daniel J. Park, Glen R. Gallagher#, Chantal B.F. Vogels# 2022. Monkeypox virus multiplexed PCR amplicon sequencing (PrimalSeq) . **protocols.io** **https://dx.doi.org/10.17504/protocols.io.5qpvob1nbl4o/v3**Version created by **Nathan D Grubaugh**



**Protocol status:** In development
**We are still developing and optimizing this protocol**

**Created:** October 19, 2022

**Last Modified:** February 11, 2023

**Protocol Integer ID:** 71507

**Keywords:** Monkeypox, MPXV, Amplicon sequencing, PrimalSeq, hMPXV, current global outbreak of human monkeypox virus, human monkeypox virus, efforts of the current hmpxv outbreak, pathogen whole genome sequencing, hmpxv primer scheme with clinical specimen, current hmpxv outbreak, monkeypox virus, updated monkeypox virus, monkeypox outbreak, hmpxv reference sequence, zika virus, covidseq sar, many of the early hmpxv case, genome dna virus, metagenomic sequencing, early hmpxv case, primer scheme for hmpxv, high viral dna concentration, whole genome sequencing, clade iia dna from cultured virus, sequencing analysis, pathogen, high viral concentration, approach for hmpxv, hmpxv primer scheme, low viral concentration, current global outbreak, cultured virus, sequencing approach, need for genomic surveillance, current outbreak, genomic surveillance, hmpxv, sequencing effort, comparable genome coverage between amplicon, comparable genome coverage, metagenomics approach, genome coverage via amplicon, sequencing, sequencing platform

## Disclaimer

## Abstract

***Version 3 Updates:***

- *Updated Monkeypox virus (MPXV) nomenclature*
- *Resolved typo in step 7.3*
- *Included recommendations for sample selection and assigning library reads*
- *Included link to validation study pre-print*

**Complete validation study:** https://doi.org/10.1101/2022.10.14.22280783

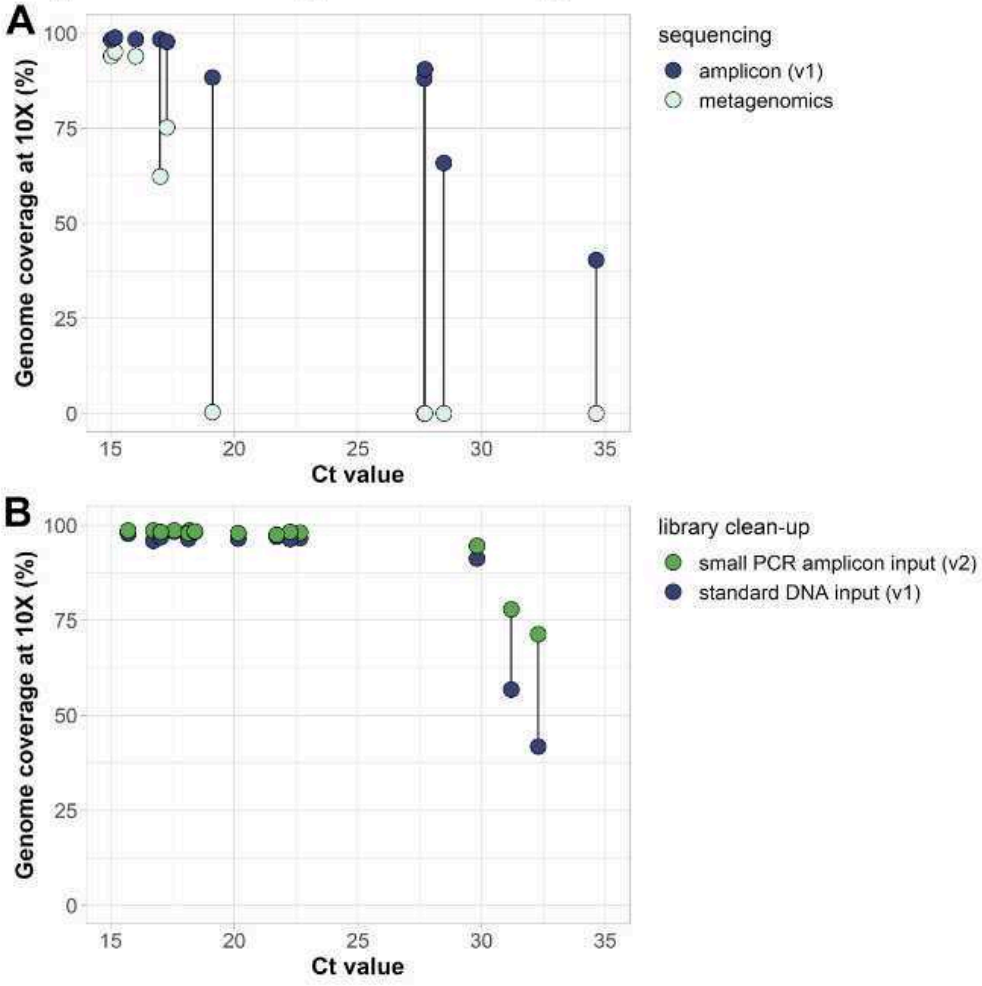
**Background:** The current global outbreak of human Monkeypox virus (hMPXV) concurrent with an ongoing SARS-CoV-2 pandemic has further highlighted the need for genomic surveillance and pathogen whole genome sequencing. While metagenomic and hybrid capture sequencing approaches were used to sequence many of the early hMPXV cases, the viability of these methods is dependent on samples with high viral DNA concentrations. Given the atypical clinical presentation of cases associated with the current outbreak and uncertainty regarding viral load across both the course of infection and anatomical body sites, there is a strong need for a more sensitive and broadly applicable sequencing approach. Amplicon-based sequencing (PrimalSeq) was initially developed for sequencing of Zika virus, and later adapted as the main sequencing approach for SARS-CoV-2. Here, we used PrimalScheme to design a primer scheme for hMPXV and we validated it with widely used SARS-CoV-2 sequencing protocols. Based on initial validation, our approach shows notably higher depth and breadth of coverage across the genome, particularly with higher PCR cycle threshold (Ct) samples, as compared to metagenomic sequencing. While further testing is needed, the early success of this approach has significant implications for sequencing efforts of the current hMPXV outbreak and serves as a proof of concept of amplicon-based sequencing for use with other large-genome DNA viruses and potentially bacterial genomes.

**Overview of Design:** We used PrimalScheme (**https://primalscheme.com/**) to generate an hMPXV primer scheme using a pre-outbreak A.1 lineage reference genome (GenBank accession: MT903345). The primer scheme comprises a total of 163 primer pairs with an amplicon length ranging between 1597 and 2497 bp (average length of 1977 bp). The primer scheme is compatible with current ARTIC and COVIDSeq SARS-CoV-2 sequencing protocols, and while validated only with Illumina library prep methods and sequencing platforms, it would likely see a high degree of success with other sequencing platforms such as the Oxford Nanopore Technologies MinION.

**Initial Validation:** We validated our hMPXV primer scheme with clinical specimens at the Massachusetts Department of Public Health, Massachusetts State Public Health Laboratory under the IRB Project Titled “Rash Illness: Alternate Specimen Types and Sequencing” (protocol number 1917413). A total of 25 clinical specimens were included in this initial validation comprised of both throat and swabs of fluid from lesions of 8 individuals, belonging to the current outbreak-associated hMPXV B.1 lineage as determined by prior characterization at the Centers of Disease Control and Prevention. Ct values were determined with the non-variola orthopox diagnostic qPCR assay developed for use by the Laboratory Response Network (**Rapid Diagnostic Testing for Response to the Monkeypox Outbreak — Laboratory Response Network, United States, May 17–June 30, 2022 | MMWR (cdc.gov)**). Clinical samples ranged in cycle threshold (Ct) values from 15.03 (high viral

concentration) to 34.63 (low viral concentration), and each sample was sequenced in parallel using a metagenomics approach and the provided amplicon-based approach. Extractions and sequencing analysis were performed in accordance with current biosafety guidance including extraction in a BSL-3 setting. An evaluation of site-specific biosafety practices should be developed in consultation with your organization's biosafety officer. Libraries were prepared with the Illumina DNA prep kit and sequenced on the MiSeq (v2 kit running 2×150 nt reads). Consensus genomes were generated at 10X coverage using the TheiaCoV_Illumina_PE Workflow Series on Terra.bio. An hMPXV "fork" was developed for us by Curtis Kapsak from Theiagen, which included the hMPXV reference sequence, primer scheme, and consensus genome length, and can currently be accessed at: https://github.com/theiagen/public_health_viral_genomics/tree/cjk-MPXV-theiacov. We found comparable genome coverage between amplicon and metagenomic sequencing with low Ct (<18) samples, yet a significant increase in genome coverage with amplicon sequencing in higher Ct samples (>18; **Figure 1A.**) The library clean up stage of the amplicon based Illumina DNA prep protocol was conducted using the 'standard DNA input' option for comparison to metagenomics. Following optimization with 15 additional paired samples unrelated to the initial 10 samples, we found improved genome coverage with higher Ct samples (>25) using the 'small PCR amplicon input' option of the Illumina DNA prep protocol (**Figure 1B.**) Our findings highlight how amplicon-based approaches can significantly expand hMPXV sequencing to a wider variety of samples.

**Figure 1: Percent genome coverage at 10X**




**A.** Comparison in genome coverage for 10 paired clinical samples sequenced with metagenomics and amplicon approaches using the v1 standard DNA input library clean up method. **B.** Comparison between the v2 Illumina DNA prep small PCR amplicon input and the v1 standard DNA input library clean up methods for 15 paired clinical samples

Amongst samples with a low Ct (<18), genome coverage via amplicon sequencing was consistently >97%, with minimal amplicon drop-outs (**Figure 2**). Amplicons 11, 75, and 118 showed consistent drop-out across the sequenced samples, while none of the primers had mismatches, except for a single nucleotide mismatch in the 11_RIGHT primer. We did obtain coverage for these amplicons when sequencing clade IIa DNA from cultured virus (strain USA-2003; NR-4928) obtained from BEI Resources (NIAID, NIH). This suggests that the dropouts of these primers may be a specific issue related to the current hMPXV B.1 lineage genomes. As this protocol is still in development, we will further investigate performance as we continue to sequence additional samples.

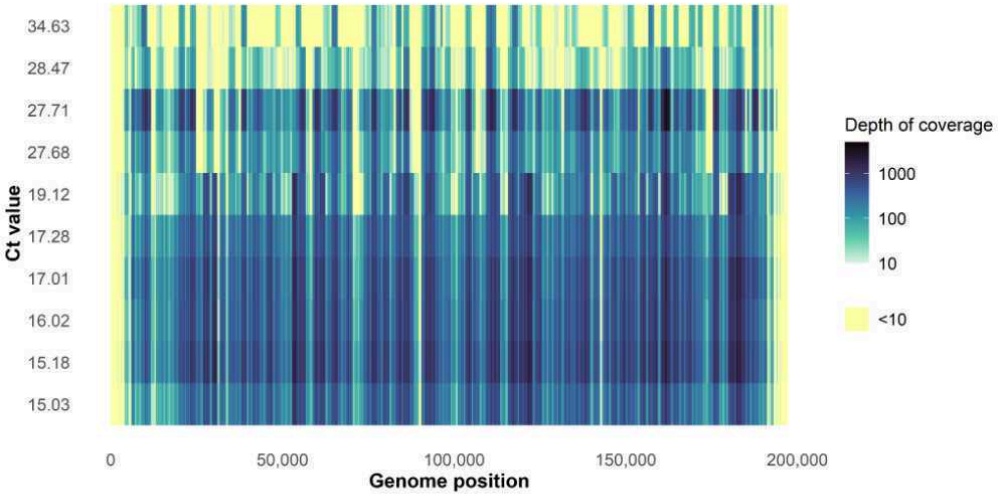



**Figure 2.** Depth of coverage by genome position and Ct value for 10 clinical specimens sequenced using the amplicon-based sequencing approach.

**Conclusion:** We developed an amplicon-based sequencing (PrimalSeq) approach for hMPXV that improved the depth and breadth of genome coverage with low viral concentration specimens as compared to metagenomic sequencing.

*This protocol represents the third iteration of development. Further versions will be uploaded to protocols.io with an accompanying description of changes as appropriate.*

## Materials

**MPXV Primer Scheme**

MPXV-primer_genome-positions.tsv

Primers may be ordered from any oligonucleotide company using the above file as a reference. For instructions on how to prepare hMPXV Primer Pool 1 and 2 (10 uM) see step 1 of the protocol.

**Library Preparation Method: Illumina DNA Prep**

**Reagents:** https://www.illumina.com/products/by-type/sequencing-kits/library-prep-kits/nextera-dna-flex.html https://www.neb.com/products/m0494-q5-hot-start-high-fidelity-2x-master-mix#Product%20Information

**Library Preparation Method: CovidSeq**

**Reagents**: https://www.illumina.com/products/by-type/ivd-products/covidseq.html

| | A | B | C |
|---|---|---|---|
| | **Illumina COVIDSeq Test Box 1 – 3072 Samples, Part # 20044408** | | |
| | **Reagent** | **Description** | **Storage** |
| | ITB | Illumina Tune Beads | Room Temperature |
| | ST2 HT | Stop Tagment Buffer 2 HT | Room temperature, post-amp environment |
| | **Illumina COVIDSeqTest Box 2 – 3072 Samples, Part # 20044409** | | |
| | EBLTS HT | Enrichment BLT HT | 2°C to 8°C post-amp environment |
| | TWB HT | Tagmentation Wash Buffer HT | 2°C to 8°C post-amp environment |
| | RSB HT | Resuspension Buffer HT | 2°C to 8°C, post-amp environment |
| | **Illumina COVIDSeq Test Box 3 – 3072 Samples, Part # 20044410** | | |
| | IPM HT | Enhanced PCR Mix HT | -25°C to -15°C, pre-amp environment |
| | TB1 HT | Tagmentation Buffer 1 HT | -25°C to -15°C, post-amp environment |
| | EPM HT | Enhanced PCR Mix HT | -25°C to -15°C, pre-amp environment |

| | A | B | C |
|---|---|---|---|
| | **Index Adapater Part Numbers : 20043132, 20043133, 20043134, 20043135** | | |
| | Index Adapters | IDT for Illumina- PCR Indexes Set 1-4 | Room Temperature |

**Additional Materials**

| | A | B | C |
|---|---|---|---|
| | **Reagent** | **Description** | **Storage** |
| | 80% EtOH | 80% Ethanol | Room Temperature |
| | Nuclease-free water | | Room Temperature |

## Troubleshooting

## Safety warnings

Processing of any sample type which could potentially be positive for hMPXV should be conducted in BSL2+ settings. Before starting work with these samples, please contact your local EHS (environment, health and safety) or biosafety office for proper guidance on how to work with these samples in your laboratory.

## Before start

This protocol is currently in the developmental phase. It is for research purposes only and should not be used in a diagnostic capacity.

If using clinical samples, DNA will need to be extracted and purified prior to beginning library preparation.

Primer diluting and pooling requires several hours and may be conducted well in advance so long as the pooled primers are stored at -20°C.

Following primer dilution and pooling, this workflow can be completed in one day, however, it is recommended to be conducted across two, with the amplicon generation step on the first day and all subsequent wet-lab steps on the second day.

**Protocol Recommendations:**
To maximize sequencing coverage, we recommend the following best practices based on our **multi-site validation study**:

1. Prioritizing samples with cycle threshold (Ct) values <29 if resources are limited
2. Generating at least 1 million sequencing reads per sample

## Dilute and Pool Primers

1 **Reagents:**

**Primer Preparation**

| Reagent | Storage | Instructions |
|---|---|---|
| 82 Odd Numbered Primer Pairs (100 µM) | -20°C | Thaw at RT |
| 81 Even Numbered Primer Pairs (100 µM) | -20°C | Thaw at RT |
| Nuclease-free water | RT | |

Primers should be ordered lyophilized or resuspended (100 uM; recommended).
Lyophilized primers should be resuspended to 100 uM in nuclease-free water.

Note

**Primer Scheme:** MPXV-primer_genome-positions.tsv

1.1 If not already done, separate odd and even numbered primer pairs into two separate boxes. These will constitute the two pools

Note

*For example:*
***Primer Pool 1:*** *1 left, 1 right, 3 left, 3 right, etc.*
***Primer Pool 2:*** *2 left, 2 right, 4 left, 4 right, etc.*

1.2 Label 164, 8-strip tubes with the corresponding odd-numbered primer name (e.g. 3 left)

1.3 To each tube add 90 µL of nuclease-free water

Note

This will comprise the 10X dilution to arrive at a final primer concentration of 10uM

1.4 For each odd-numbered primer tube:
- Spin down
- Pipette 10 times to mix
- Add 10 µL to the corresponding labeled tube
- Pipette 10 times to mix

1.5 After all 164 primers have been aliquoted, combine 10uL from each tube into a 2mL tube. This will be the odd-numbered primer pool

Note

To most efficiently pool all 164 primers, use a multi-channel pipette to remove 10uL from each 8-strip and pool into a new 8-strip.
Then combine each of the pooled 8 strip tubes into a single 2mL tube

1.6 Repeat steps 1.2-1.5 with the even -numbered primers (**NOTE:** There will be two fewer primers in the even set than the odd)

1.7 ***Safe Stopping Point: Pooled Primers Can Be Stored at -20°C***

2 **Library Preparation Method**

STEP CASE

Illumina CovidSeq Test (RUO) 53 steps

## Amplicon Generation

3 **Reagents:**

| Reagent | Storage | Instructions |
|---|---|---|
| MPXV Primer Pool 1 (10 uM) | -20°C | Thaw at RT |
| MPXV Primer Pool 2 (10 uM) | -20°C | Thaw at RT |
| IPM HT | -20°C | Thaw at RT |

3.1 In two separate tubes, prepare the following master mixes:

**Pool 1**

| Reagent | Volume in 20 µL master mix |
|---|---|
| IPM HT | 12.5 µL |
| MPXV Primer Pool 1 (10 uM) | 3.6 µL |
| Nuclease-free water | 3.9 µL |

**Pool 2**

| Reagent | Volume in 20 µL master mix |
|---|---|
| IPM HT | 12.5 µL |
| MPXV Primer Pool 2 (10 uM) | 3.6 µL |
| Nuclease-free water | 3.9 µL |

Note

**Master mix volumes are for one reaction and do not account for lost volume due to pipetting. Multiply volumes by reaction number accordingly.**

3.2 Label two sets of PCR tubes/plates for Pool 1 and Pool 2

3.3 Add the following:

- 20 µL **Pool 1 master mix** to each Pool 1 tube/well
- 20 µL **Pool 2 master mix** to each Pool 2 tube/well

3.4 Add 5 µL **DNA** to each tube in **both** sets

- Mix by pipetting up and down 10 times
- Briefly centrifuge tubes/plates

Note

Be sure to include a negative PCR control (NTC; nuclease-free water) for **each pool**

3.5 Place on thermocycler and run the following program (choose preheat lid option):

| PCR Program | | | |
|---|---|---|---|
| **Step(s)** | **Temperature** | **Time** | **Cycles** |
| Initial denaturation | 98°C | 3min | 1 |
| Denaturation | 98°C | 15sec | 35 times |
| Anneal and extension | 63°C | 5min | |
| Hold | 4°C | Hold | Hold |
| *Volume Amount: 25 µL* | | *Lid: 105°C* | |

3.6 ***Safe Stopping Point: Amplified DNA may remain in the thermocycler at 4°C or stored at -20°C until ready to use***

## Amplicon Tagmentation and Clean Up

4 **Reagents:**

| Reagent | Storage | Instructions |
|---|---|---|
| EBLTS HT | 4°C | Thaw at RT; vortex to mix |
| TB1 HT | -20°C | Thaw at RT; vortex to mix |
| ST2 HT | RT | Vortex before use |
| TWB HT | 4°C | Vortex before use |

4.1 Spin down PCR tubes/plates

4.2 Prepare the following master mix:

| Reagent | Volume in 30 µL master mix |
|---|---|
| TB1 | 10 µL |
| EBLTS | 3.3 µL |
| Nuclease-free water | 16.7 µL |

4.3 In **each** PCR tube add:

- 30 µL **Master mix**
- 10 µL **Pool 1 amplicons**
- 10 µL **Pool 2 amplicons**

For a total reaction volume of 50 µL per PCR tube/well

Note

Pooling of amplicons should be conducted on a dedicated post-PCR bench to prevent contamination

4.4 Mix by pipetting up and down and briefly centrifuge

4.5 Place on thermocycler and run the following program (choose preheat lid option):

| Temp | Time |
|---|---|
| 55°C | 3 minutes |
| 10°C | hold |

4.6 Once the thermocycler reaches 10 °C , remove tubes/plates and spin down

4.7 Add 10 µL **ST2** to each tube/well, mix by pipetting up and down, and briefly centrifuge

4.8 Incubate at Room temperature for 00:05:00 5m

4.9 Place on magnetic stand and wait until liquid is clear (a few minutes)

4.10 Remove and discard all supernatant

4.11 Remove tubes/plates from magnetic stand and add 100 µL **TWB** to each tube/well

- Mix by pipetting up and down 10 times and spin down. **Be careful to not introduce bubbles**.

4.12 Repeat steps 4.9 - 4.11. **Leave the supernatant after the second wash so that the beads don't dry out**

## Amplify Tagmented Amplicons

5 **Reagents:**

| Reagent | Storage | Instructions |
|---|---|---|
| EPM HT | -20°C | Invert to mix |
| Index adapters | -20°C | Thaw at RT; vortex to mix |

5.1 Prepare the following master mix:

| Reagent | Volume in 40 µL master mix |
|---|---|
| EPM | 20 µL |
| Nuclease-free water | 20 µL |

5.2 Place the tubes/plates with tagmented amplicons on the magnetic stand and remove the supernatant once the liquid is clear

5.3 Use a **20µL pipette** to remove any residual TWB from tubes/wells

5.4 Remove the tubes/plates from the magnetic stand

5.5 Add 40 µL **master mix** to each tube/well

5.6 Add 10 µL **dual-barcoded index adapters** to each tube/well

Note

**Note which set of indexes are used for each tube for bioinformatic processing**

5.7 Pipette up and down to mix and spin down

5.8 Place on thermocycler and run the following program (choose preheat lid option):

| Temp | Time | Repeat |
|---|---|---|
| 72°C | 3 minutes | |
| 98°C | 3 minutes | |
| 98°C | 20 seconds | 7 times |
| 60°C | 30 seconds | |
| 72°C | 1 minute | |
| 72°C | 3 minutes | |
| 10°C | hold | |

## Pool and Clean Up

6 **Reagents:**

| Reagent | Storage | Instructions |
|---|---|---|
| ITB | RT | Vortex thoroughly to mix |
| RSB HT | 4°C | Bring to RT; vortex to mix |
| 80% EtOH | RT | Prepare immediately before use |

6.1 Briefly centrifuge tubes/plates

6.2 Place on magnetic stand and wait until liquid is clear (a few minutes)

6.3 Pool libraries by equal volume:

| Number of individual samples | Volume to pool per sample |
|---|---|
| 1-24 | 40 µL |
| 25-48 | 20 µL |
| 49-72 | 10 µL |
| 72-96 | 5 µL |

6.4 Pipette up and down to mix pooled libraries and briefly spin down

6.5 Calculate the volume of **ITB** to reach a **0.6X** beads:total pool volume ratio

Note

*For example: Total pooled volume is 400 µL (10 samples x 40 µL/sample) then add 0.6 x 400 µL = 240 µL ITB*

6.6 Mix by pipetting up and down, briefly centrifuge, and incubate at Room temperature for 00:05:00 5m

6.7 Place on magnetic stand and wait until liquid is clear (a few minutes)

6.8 Transfer supernatant to a new tube **(do not discard, this is your final library)**. This volume will be slightly lower than the total combined volume

Note

*For example: 400µL+ 240µL = total 640µL > transfer 630µL*

6.9 Calculate the volume of beads to add to the supernatant to attain a second clean-up beads:sample ratio of **0.9X:**

$$\frac{\text{beads from 1st step} + \text{beads added in 2nd step} \times \frac{Vo}{Vt}}{\text{Volume of DNA sample from right-side clean-up}}$$

**Vo = total volume of sample + beads from step 5.5**
**Vt = transferred volume of supernatant**

Note

***For example**: 630µL of supernatant is transferred to a new tube*

$$240 + \mathbf{118} * \frac{640}{630} = 360$$

$$\frac{360}{400} = \mathbf{0.9}\boldsymbol{X}$$

*Add **118µL** of ITB to **630µL** of transferred supernatant*

***Calculation source**: https://support.illumina.com/bulletins/2020/07/library-size-selection-using-sample-purification-beads.html*

6.10 Add **beads** to supernatant, mix by pipetting up and down, briefly centrifuge, and incubate at Room temperature for 00:05:00 5m

6.11 Place on magnetic stand and wait until liquid is clear (a few minutes)

6.12 Carefully discard supernatant

6.13 Add 1000 µL **80% EtOH**

6.14 Wait 00:00:30 30s

6.15 Repeat steps 6.12 - 6.14

6.16 Remove supernatant using a **20µL pipette** to remove all residual EtOH

6.17 Add 55 µL of **RSB** to the tube, mix by pipetting up and down, and briefly centrifuge

6.18 Incubate at Room temperature for 00:02:00 2m

6.19 Place on magnetic stand and wait until liquid is clear (a few minutes)

6.20 Transfer 50 µL of final pooled library to a new 1.5mL tube

6.21 Quantify library on a Qubit and obtain fragment distribution using a Bioanalyzer/Tape Station

Note

**Qubit SOP:** Agilent High Sensitivity DNA Kit Gui...

**Bioanalyzer SOP:** Agilent High Sensitivity DNA Kit Gui...

## Sequencing

7 Protocol validated on the MiSeq (v2 kit running 2×150 nt reads)

Note

**Note:** For sequencing we recommend generating at least 1 million reads per sample for optimal sequencing coverage

Sequencing may be performed on Illumina and Oxford Nanopore Technologies sequencing platforms following standard protocols

## Bioinformatics/Analysis

8 Sequencing results may be analyzed utilizing a standard amplicon sequencing bioinformatics pipeline, including those employed for SARS-CoV-2 sequencing.

Note

If utilizing a SARS-CoV-2 based bioinformatics pipeline, adjustments may be required to account for the consensus genome length

**Optional bioinformatics pipeline: Terra.bio_TheiaCov_Illumina_PE_workflow**

**Reference Sequence:** MPXV.reference.fasta

**Bed file:** MPXV.primer.bed