

## Next-generation Sequencing of the DNA Virome from Fecal Samples

Cynthia L. Monaco<sup>1,\*</sup> and Douglas S. Kwon<sup>2,3</sup>

<sup>1</sup>Department of Medicine, Washington University School of Medicine, St. Louis, USA; <sup>2</sup>Ragon Institute of MGH, MIT, and Harvard, Cambridge, USA; <sup>3</sup>Division of Infectious Diseases, Massachusetts General Hospital and Harvard Medical School, Boston, USA

\*For correspondence: [cmonaco@wustl.edu](mailto:cmonaco@wustl.edu)

**[Abstract]** Herein we describe a detailed protocol for DNA virome analysis of low input human stool samples (Monaco *et al.*, 2016). This protocol is divided into four main steps: 1) stool samples are pulverized to evenly distribute microbial matter; 2) stool is enriched for virus-like particles and DNA is extracted by phenol-chloroform; 3) purified DNA is multiple-strand displacement amplified (MDA) and fragmented; and 4) libraries are constructed and sequenced using Illumina Miseq. Subsequent sequence analysis for viral sequence identification should be sensitive but stringent.

**Keywords:** Virome, Viral microbiome, Virus-like particles, MDA, Illumina Miseq

**[Background]** The virome, a dynamic community of eukaryotic viruses, bacteriophages and endogenous retroviruses, represents a minimally characterized component of the human microbiome (Virgin, 2014). In fact, it is estimated that only 1% of the virome has been sequenced and annotated (Mokili *et al.*, 2012). Next generation sequencing (NGS) enables examination of the entire virome, including unculturable viruses. Stool is a readily obtainable specimen type for study of the virome, and alterations in the fecal virome have been associated with a number of disease states (Handley *et al.*, 2012; Norman *et al.*, 2015; Monaco *et al.*, 2016). The fecal virome is largely comprised of bacteriophages, which affect the gastrointestinal tract through alterations in bacterial functions and populations (Duerkop and Hooper, 2013; Reyes *et al.*, 2013; Virgin, 2014). Enteric eukaryotic viruses, while less ubiquitous than bacteriophages, play a more direct role in gastrointestinal tract dysfunction by inducing gastroenteritis, enteritis and colitis. Despite the abundance of bacteriophages in fecal samples, only a few studies thus far have examined the contributions of fecal bacteriophages in human diseases. Inflammatory bowel disease has been associated with increased enteric bacteriophage richness (Norman *et al.*, 2015). In contrast, profound immunosuppression from AIDS in a sub-Saharan cohort resulted in an expanded eukaryotic virome, but had minimal impact on bacteriophage populations (Monaco *et al.*, 2016). More studies are needed to elucidate the role the fecal virome plays in disease states. A key roadblock to studying the stool virome is viral nucleic acid extraction and enrichment from fecal material. Several factors can contribute to difficulty in isolating viral sequences from fecal samples, chief among them the fact that viruses constitute a minority of fecal sample material. Additionally, dilution of feces in collection media (such as RNAlater RNA stabilization reagent) can further hamper the ability to find viral sequences. While many nucleic acid extraction protocols can be used for high input nucleic acid samples to enrich for viral nucleic acid, low input samples, such as

those diluted in collection media, represent a challenge with virome studies. After comparison and optimization of several methods, the following protocol was identified as the most universally applicable for isolation of phage and DNA viral sequences from both low (Monaco *et al.*, 2016) and high (Norman *et al.*, 2015) input samples.

## **Materials and Reagents**

### A. Stool aliquoting and pulverization

1. Versi-dry sheets (Thermo Fisher Scientific, Thermo Scientific™, catalog number: 74018-00)
2. Large Kim-wipes (KCWW, Kimberly-Clark, catalog number: 34721)
3. Extra heavy-duty aluminum foil (VWR, catalog number: 89107-734)
4. Dry ice
5. Liquid nitrogen and a Dewar along with Styrofoam cup
6. Pre-labeled screw capped tubes (STARSTEDT) in a freezer box (4 pre-labeled tubes per sample)
7. Stool scrapers (autoclaved in sets of 6) (Fisher Scientific, catalog number: 21-401-25B)
8. Sterilization pouches (small)
9. Autoclave bags
10. Bleach (5.25% solution of sodium hypochlorite)
11. 75% EtOH

### B. Virus-like particle (vlp) preparation

1. Sterile 1.5 ml and 2.0 ml screw-cap tubes
2. 1 ml luer-lok syringes (BD, catalog number: 309628)
3. 0.45 µm filters 13 mm diameter (EMD Millipore, catalog number: SLHV013SL)
4. 0.22 µm filters 13 mm diameter (EMD Millipore, catalog number: SLGV013SL)
5. 1 M Tris, pH 7.5 (Fisher Scientific, catalog number: MT-46-030-CM) (Tris 1 M pH 7.5, DNase-, RNase-, protease-free [6 x 1 L bottles])
6. 5 M NaCl (Promega, catalog number: V4221) (5 M NaCl 1 L bottle, DNase-, RNase-, protease-free; aliquoted in 1 ml aliquots and stored at -20 °C)
7. Saline magnesium (SM) buffer (Fisher Scientific, catalog number: 50-329-444) (SM buffer with gelatin pH 7.5, 100 ml)
8. 10% SDS (diluted from 20% SDS stock [Thermo Fisher Scientific, Ambion™, catalog number: AM9820] in RNase, DNase, Protease free H<sub>2</sub>O; aliquoted and stored at -20 °C)
9. Lysozyme (10 mg/ml) (EMD Millipore, catalog number: 71412) – aliquoted and stored at -20 °C
10. Turbo DNase I (2 U/µl) (Thermo Fisher Scientific, Ambion™, catalog number: AM2238)
11. BaseLine zero DNase (1 U/µl) (Epicentre, catalog number: DB0711K)
12. Chloroform (Fisher Scientific, catalog number: C298-500)

13. Phenol:chloroform:isoamyl alcohol (25:24:1) pH 8.0 (Fisher Scientific, catalog number: BP1752I-100)
14. QIAGEN DNeasy Blood and Tissue Kit (QIAGEN, catalog number: 69506)
15. CTAB (Sigma-Aldrich, catalog number: 52365)
16. CTAB/NaCl (see Recipes) [0.45 µm filtered + 0.22 µm filtered]

#### C. Library construction

1. GenomiPhi V2 DNA Amplification Kit (GE Healthcare, catalog number: 25-6600-31)
2. Covaris microTUBE AFA fiber snap-cap 50 µl (Covaris, catalog number: 520045)
3. NEBNext® Ultra™ DNA Library Prep Kit for Illumina (New England Biolabs, catalog number: E7370L, 96 rxns)
4. NEBNext® Multiplex Oligos for Illumina®  
Index Primers Set 1 (New England Biolabs, catalog number: E7335L)  
Index Primers Set 2 (New England Biolabs, catalog number: E7500L)
5. AMPure XP beads (Beckman Coulter, catalog number: A63881)
6. TE low EDTA
7. D1K reagents (Agilent Technologies, catalog number: 5067-5362)
8. D1K screen tape for TapeStation (Agilent Technologies, catalog number: 5067-5361)

#### **Equipment**

1. Biosafety hood
2. 6 mortar/pestles, 100 ml capacity
3. Beaker
4. Microcentrifuge
5. PCR hood
6. -80 °C freezer
7. Covaris E210
8. PCR Thermocycler (Eppendorf)
9. NanoDrop Micro-Volume UV-Vis spectrophotometer
10. Agilent 2200 TapeStation (Agilent Technologies, model: 2200 TapeStation)
11. Loading tips for TapeStation (Agilent Technologies, catalog number: 5067-5153)
12. Dynamag™-spin magnet (Thermo Fisher Scientific, catalog number: 12320D)

#### **Procedure**

##### A. Stool aliquoting/pulverization

*Note: All steps were performed in a biosafety hood due to use of human feces and risk of viable infectious organisms.*

1. Washing and preparation of mortars and pestles
  - a. Thoroughly sterilize working area with 10% bleach followed by 75% EtOH and UV exposure for 10 min to minimize contamination.
  - b. Prepare 10% bleach beaker for scrapers.
  - c. Bleach mortar/pestles for 10 min in 10% bleach solution.
  - d. Place versi-dry sheets in hood.
  - e. Place samples on dry ice in hood.
  - f. With remaining dry ice, place open microtube freezer box stably on top of dry ice with empty pre-labeled tubes and place in hood.
  - g. Tear heavy duty aluminum foil into pieces around 10 inches wide. Then using a flat surface, tear this into 3 approximately equal squares. Do this twice (6 squares). Place these in the hood.
  - h. Once the 10 min bleach is finished, rinse all 6 mortars and pestles with dH<sub>2</sub>O and set on another versi-dry to dry slightly, and then set upright in hood on versi-dry. UV for 10 min along with all other equipment to be used.
  - i. When done, place pestles in mortars, and set all but one to side, stack aluminum foil.
  - j. Add large mound of gloves to hood.
  - k. Add liquid N<sub>2</sub> to Dewar.
2. Pulverization of stool
  - a. Double glove.  
*Note: This is not needed if not using human stool.*
  - b. Place large Kim-wipe on top of the versi-dry. Place mortar and pestle on it.
  - c. Open a sterile package of stool scrapers.
  - d. Place a piece of aluminum foil into mortar and mold to inner shape of mortar.
  - e. Select sample to use, bring out of dry ice and set on Kim-wipe.
  - f. Identify corresponding pre-labelled screw-cap tubes, and bring them to front of the box.
  - g. Fill Styrofoam cup 3/4 full with liquid N<sub>2</sub> from Dewar. Pour some liquid N<sub>2</sub> into the box of screw-top tubes, and pour enough into the mortar to fill it 1/2 way full. Place cup to side of mortar (recommend having some liquid N<sub>2</sub> remaining).
  - h. Take a stool scraper out of package.
  - i. Open sample tube.
  - j. Use stool scraper to carefully get sample out of tube into the mortar. Try to get the whole sample out at once.  
*Note: May need to warm sample tube in hands to slightly thaw around edges. Do not thaw longer than necessary to remove sample.*
  - k. Once out, place now empty sample tube into the box of tubes to re-freeze, consider adding small amount of liquid N<sub>2</sub> near the tube to aid with re-freezing.

- l. For large chunks of stool, use pestle to break apart. When all large chunks have been crushed to small pieces, then pulverize with pestle. Intermittently add more liquid N<sub>2</sub> to mortar as needed to keep sample from thawing out until pulverized completely.
- m. Let liquid N<sub>2</sub> boil off. Sample now has the texture of a powder.
- n. Immediately take a fresh tube and scrape approximately 200 mg sample into the tube, tap the tube onto clean side of cold mortar to settle contents to bottom, and place it back into microtube box. Repeat with the remaining 3 tubes. If any sample remains after aliquoting 200 mg to new tubes, scrape remaining sample into the original tube, tapping tube onto a clean cold portion of the mortar to settle contents to the bottom of the tube. If needed, remove aluminum foil from the mortar to get any remaining bits of sample.

*Notes:*

- a. *Sample left in aluminum foil thaws VERY quickly once out of the mortar.*
- b. *Do not screw on caps tightly as it is likely they still have some liquid N<sub>2</sub> present.*
- o. Place stool scraper into bleach beaker, aluminum foil into autoclave bag, pestle into mortar, deglove, take Styrofoam cup out of hood. Place Kim-wipe into an autoclave bag. If versidry is dirty, remove versidry and place it into the autoclave bag.
- p. Carefully place mortar and pestle in bleach bath.
- q. Repeat process for total of 6 samples.
- r. Store samples at -80 °C.
- s. Wash as above and repeat as needed.

B. VLP enrichment protocol

1. Pulverize as above ~200 mg of stool into sterile 2 ml screw cap tube on dry ice to keep frozen.
2. Add 400 µl cold SM buffer per sample. Keep on ice. Vortex on high speed for 5 min.

*Note: It is possible to use a 2 ml tube adapter plate to hold all sample tubes at once if available. Avoid touching tip to tube when adding reagents. Change tips between samples to avoid cross-contamination. Low speed centrifugation is used to avoid lysing bacteria. Goal is to get everything into solution. More SM buffer can be added if needed.*

3. Centrifuge for 10 min at 2,000 x g at 4 °C. Transfer supernatant to a clean 1.5 ml Eppendorf tube. Centrifuge a second time for 10 min at 2,000 x g at 4 °C.

*Note: Sometimes a third spin is needed to clarify supernatant. Final supernatant volume should be at least 200 µl because a fraction of the volume will be lost in the subsequent filtration. More SM buffer can be added if needed.*

4. Filter supernatant once through a 0.45 µm filter. Then filter twice through 0.22 µm filters. All filtration steps use 1 ml Luer-lok syringes and 13 mm diameter filters.

*Note: Filtering once through 0.22 µm filter is sufficient, but 2 x 0.22 µm filtration steps were used in this manuscript. If the filter clogs: pull the plunger out to relieve pressure, take off the filter and put a new one on. If the initial sample volume was low, may need to use a pipette to recover remaining sample from the old, clogged filter.*

5. Check final volume of filtrate. Bring to 200  $\mu$ l with cold SM buffer.
  - a. Take 200  $\mu$ l to a clean 1.5 ml tube on ice.
  - b. Store any remainder for later use at -80  $^{\circ}$ C.
6. Lysozyme/chloroform/DNase treatment
  - a. Add 20  $\mu$ l (10% volume) lysozyme (10 mg/ml stock) to each tube. Incubate at 37  $^{\circ}$ C 30 min.
  - b. Add 44  $\mu$ l (20% volume) chloroform, briefly vortex, and incubate 10 min at room temperature (RT; 15-25  $^{\circ}$ C).
  - c. Centrifuge at 2,500 x g 5 min at RT.
  - d. Collect aqueous phase and transfer to 2 ml screw cap tube.

*Note: Screw cap is preferable to prevent sample loss from the top popping open during the heating.*
  - e. Make DNase master mix.

	Per sample
Turbo DNase buffer	25 $\mu$ l
Turbo DNase I (2 U/ $\mu$ l)	2 $\mu$ l
Baseline zero (1 U/ $\mu$ l)	1 $\mu$ l
H <sub>2</sub> O	22 $\mu$ l
Final volume	50 $\mu$ l

7. Add 50  $\mu$ l DNase master mix to each tube. Incubate at 37  $^{\circ}$ C for 1 h.
8. Heat-inactivate DNase at 65  $^{\circ}$ C for 15 min followed by a quick spin to pull down condensation.
9. SDS/CTAB cleanup
  - a. Add 10  $\mu$ l SDS (10%) + 1  $\mu$ l proteinase K (10 mg/ml stock) to each tube and incubate at 56  $^{\circ}$ C for 20 min. At this point, pre-incubate CTAB/NaCl solution at 65  $^{\circ}$ C.
  - b. Add 35  $\mu$ l NaCl (5 M soln) + 28.1  $\mu$ l CTAB/NaCl (2.5% soln). Pulse vortex. Incubate at 65  $^{\circ}$ C for 10 min then perform a quick spin.

*Note: Sample will turn cloudy after CTAB/NaCl is added.*
  - c. Add 200  $\mu$ l phenol:chloroform:isoamyl alcohol (25:24:1) pH 8.0. Pulse vortex. Centrifuge 8,000 x g for 5 min at RT.
  - d. Collect the aqueous fraction from step B9c. Add 200  $\mu$ l chloroform. Pulse vortex for 3-5 sec. Centrifuge 8,000 x g for 5 min at RT.
  - e. Collect the aqueous fraction from step B9d. This is final Virus Nucleic Acid.
  - f. Place in dry ice for transport and store at -80  $^{\circ}$ C.

*Note: All steps here below were performed in a PCR hood.*

10. Clean DNA using QIAGEN DNeasy Blood and Tissue Kit. Elute in 200  $\mu$ l elution buffer.

*Note: Start at the buffer AL step in the Purification of Total DNA from Animal Blood or Cells protocol. Other column clean-up kits may be used, but should remain consistent between samples used for comparison. We would recommend only clean-up kits that have a large size range of DNA retained.*

  - a. Add 200  $\mu$ l buffer AL. Mix thoroughly by vortexing.
  - b. Add 200  $\mu$ l ethanol (96-100%). Mix thoroughly by vortexing.

- c. Pipet 650  $\mu$ l of mixture into a DNeasy mini spin column placed in a 2 ml collection tube. Centrifuge at  $\geq 6,000 \times g$  for 1 min. Discard the flow-through. Repeat as necessary until all of sample is used.
- d. Add 500  $\mu$ l buffer AW1. Centrifuge for 1 min at  $\geq 6,000 \times g$ . Discard the flow-through.
- e. Add 500  $\mu$ l buffer AW2, and centrifuge for 2 min at 20,000  $\times g$ . Discard the flow-through and collection tube.
- f. Centrifuge for 1 min at 20,000  $\times g$  in a new collection tube.
- g. Transfer the DNeasy spin column to a new 1.5 ml or 2 ml microcentrifuge tube.
- h. Elute the DNA by adding 200  $\mu$ l buffer AE to the center of the DNeasy spin column membrane. Incubate for 1 min at RT. Centrifuge for 1 min at  $\geq 6,000 \times g$ .
- i. Recirculate the eluate through the column once (Add eluate to the center of the DNeasy spin column membrane. Incubate for 1 min at RT. Centrifuge for 1 min at  $\geq 6,000 \times g$ ).

### C. Phi29 polymerase DNA amplification

*Note: 2  $\mu$ l of each sample is used as template in 4 independent MDA reactions using the GenomiPhi V2 DNA Amplification Kit to reduce amplification bias. The 4 replicates are then pooled after MDA and quantified by NanoDrop.*

1. Add 8  $\mu$ l of GenomiPhi V2 DNA Amplification Kit sample buffer + 2  $\mu$ l of template to a clean 0.5 ml PCR tube.
2. Heat at 95  $^{\circ}$ C for 3 min then cool to 4  $^{\circ}$ C on ice. Keep cold.
3. Prepare master mix on ice:

	Per rxn
Reaction buffer	9 $\mu$ l
Enzyme mix	1 $\mu$ l
Final volume	10 $\mu$ l

4. Add 10  $\mu$ l of the above master mix per tube and return to thermocycler.
5. Heat to 30  $^{\circ}$ C for 2 h. Then heat-kill enzyme at 65  $^{\circ}$ C for 10 min. Cool to 4  $^{\circ}$ C. Can store at -20  $^{\circ}$ C.

*Note: Incubation time should be optimized to the shortest time that allows sufficient amplification to minimize amplification bias.*

6. Pool the 4 independent MDA reactions from the same sample. Adjust volume to 200  $\mu$ l with DNase-free H<sub>2</sub>O.
7. Purify MDA product using QIAGEN DNeasy Blood and Tissue Kit. Elute in 100  $\mu$ l buffer AE, re-circulate through column once to increase yield.
  - a. Add 200  $\mu$ l buffer AL. Mix thoroughly by vortexing.
  - b. Add 200  $\mu$ l ethanol (96-100%). Mix thoroughly by vortexing.
  - c. Pipet the mixture into a DNeasy mini spin column placed in a 2 ml collection tube. Centrifuge at  $\geq 6,000 \times g$  for 1 min. Discard the flow-through.
  - d. Add 500  $\mu$ l buffer AW1. Centrifuge for 1 min at  $\geq 6,000 \times g$ . Discard the flow-through.

- e. Add 500  $\mu$ l buffer AW2, and centrifuge for 2 min at 20,000  $\times$  g. Discard the flow-through and collection tube.
- f. Centrifuge for 1 min at 20,000  $\times$  g.
- g. Transfer the spin column to a new 1.5 ml or 2 ml microcentrifuge tube.
- h. Elute the DNA by adding 100  $\mu$ l buffer AE to the center of the spin column membrane. Incubate for 1 min at RT. Centrifuge for 1 min at  $\geq$  6,000  $\times$  g.
- i. Recirculate the eluate through the column once.
- j. Check the concentration by NanoDrop.

D. Covaris fragmentation with Covaris E210

1. Dilute 200 ng DNA to 50  $\mu$ l total volume dH<sub>2</sub>O per sample in the Covaris snap-top tubes.  
*Note: Increasing to 500 ng DNA at this step did not increase number of different viral sequences.*
2. Pre-chill the Covaris.
3. Fragment nucleic acid using the 400 setting as per Covaris manual: Intensity 5, Duty cycle 5%, Cycles per burst 200, Treatment time 55 sec, Temp 7 °C, Water level 6, Sample volume 50  $\mu$ l.
4. Remove samples and store on ice if proceeding immediately to the next step (recommended) or store at -20 °C for later use.

E. NEBNext DNA library construction

*Note: Perform as per NEBNext<sup>®</sup> Ultra<sup>™</sup> DNA Library Prep Kit for Illumina manual protocol.*

1. End repair
  - a. Bring DNA up to 55.5  $\mu$ l with dH<sub>2</sub>O.
  - b. Add 6.5  $\mu$ l End repair reaction buffer.
  - c. Add 3  $\mu$ l End Prep enzyme mix.
  - d. In thermocycler:
    - 20 °C for 30 min
    - 65 °C for 30 min
    - 4 °C Hold
2. Adaptor ligation
  - a. Add 15  $\mu$ l Blunt/TA ligase master mix.
  - b. Add 2.5  $\mu$ l undiluted NEBNext adaptor.
  - c. Add 1  $\mu$ l ligation enhancer.
  - d. In thermocycler, incubate at 20 °C for 15 min.
3. Add 3  $\mu$ l USER enzyme to ligation mix and heat in thermocycler at 37 °C for 15 min.
4. AmPure bead size selection for 400-500 bp library product (300-400 bp insert).
  - a. Add 13.5  $\mu$ l nuclease-free H<sub>2</sub>O to the ligation reaction for a final volume of 100  $\mu$ l.
  - b. Add 40  $\mu$ l AmPure beads and mix well by pipetting up and down.
  - c. Incubate at RT for 5 min.



- d. Centrifuge tubes briefly. Place on magnet and incubate at RT for 5 min (until clear).
- e. Transfer supernatant to a clean tube.
- f. Add 20  $\mu$ l AmPure beads and mix well. Incubate at RT for 5 min.
- g. Centrifuge tubes briefly. Place on magnet and incubate at RT for 5 min.
- h. Leaving tubes on magnet, remove and discard supernatant, careful to not dislodge beads.
- i. With tubes on magnet, add 200  $\mu$ l freshly prepared 80% EtOH and incubate at RT for 30 sec.
- j. With tubes on magnet, remove EtOH wash.
- k. Repeat steps E4i and E4j two more times for a total of 3 washes.
- l. Air dry beads for 10 min at RT on the magnet with the tube top open.
- m. Remove tubes from magnet and add 28  $\mu$ l TE low EDTA, pH 8 to beads and resuspend by vortexing or mixing well by pipetting.
- n. Centrifuge tubes briefly, replace tubes on magnet and incubate at RT for 5 min, or until clear.
- o. Collect 23  $\mu$ l of DNA and transfer to a clean PCR tube. Be sure not to transfer any beads.

5. PCR amplification

- a. Add the following to each tube:

Adaptor ligated DNA	23 $\mu$ l
NEBNext high fid 2x PCR mix	25 $\mu$ l
Index primer	1 $\mu$ l
Universal primer	1 $\mu$ l

- b. In a thermocycler:
  - 98 °C for 30 sec
  - 98 °C for 10 sec
  - 65 °C for 30 sec
  - 72 °C for 30 sec } 7 cycles
  - 72 °C for 5 min
  - 4 °C hold

6. PCR cleanup

- a. Add 50  $\mu$ l AmPure beads and mix well.
- b. Incubate at RT x 5 min.
- c. Centrifuge tubes briefly, place on magnet and incubate at RT for 3-5 min.
- d. Leaving tubes on magnet, remove supernatant, careful to not dislodge beads.
- e. With tubes on magnet, add 200  $\mu$ l freshly prepared 80% EtOH and incubate at RT for 30 sec.
- f. With tubes on magnet, remove EtOH wash.
- g. Repeat steps E6e and E6f once.
- h. Air dry beads for 10 min at RT.

- i. Remove tubes from magnet and add 33  $\mu$ l TE low EDTA to beads and resuspend. Incubate at RT for 2 min.
  - j. Replace tubes on magnet and incubate at RT for 4 min.
  - k. Collect 30  $\mu$ l of DNA and transfer to a clean PCR tube.
7. Assay for library yield and quality on TapeStation HS DNA as per TapeStation HS DNA protocol.
  8. Pool samples (~12 samples/run, equimolar) to a final concentration of 10 nM and verify concentration on TapeStation HS DNA.
  9. Submit for Illumina Miseq as per facility protocol and recommendations (we used loading concentration 7 pM, 1% PhiX spike-in, Std flowcell, 2 x 250 bp run).

*Note: This protocol is also applicable for Illumina Hiseq.*

### **Data analysis**

Sequence analysis methods are rapidly evolving due to advances in both hardware speed and software coding. Many sequence processing software tools are open-source (such as BBTools, <http://jgi.doe.gov/data-and-tools/bbtools/>), as are statistical analysis and graphing packages in R (<https://www.r-project.org/>). We used VirusSeeker (Zhao *et al.*, 2017), a customized automated bioinformatics pipeline based on VirusHunter (Zhao *et al.*, 2013), to detect sequences sharing nucleotide and amino acid sequence similarity to known viruses (Figure 1 below). We recommend using a stringent protocol for viral sequence identification, such as VirusSeeker, that removes low-quality sequences, repeat sequences, and non-specific viral 'hits.' Similarly stringent methods have identified novel viral sequences (Zhao *et al.*, 2013). Basic steps in the analysis protocol are shown in Figure 1. Custom viral databases can be generated after downloading sequences corresponding to all viral genomes from the NCBI database (make note of the date of download as new sequences are frequently added). Deduplication is recommended to minimize amplification bias, and taxon-assigned sequences should be normalized to account for variations in sequencing depth between samples. Novel viral sequences identified or viral sequences of interest should be validated by real-time qPCR using primers specific to the viral sequence. Additionally, sequences can be *de novo* assembled into longer contigs and compared to the NCBI nr/nt databases to better identify phylogeny of viral sequences of interest (Monaco *et al.*, 2016). Phylogenetic trees comparing sequences of interest to known related viral sequences can be made using free software such as FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>).

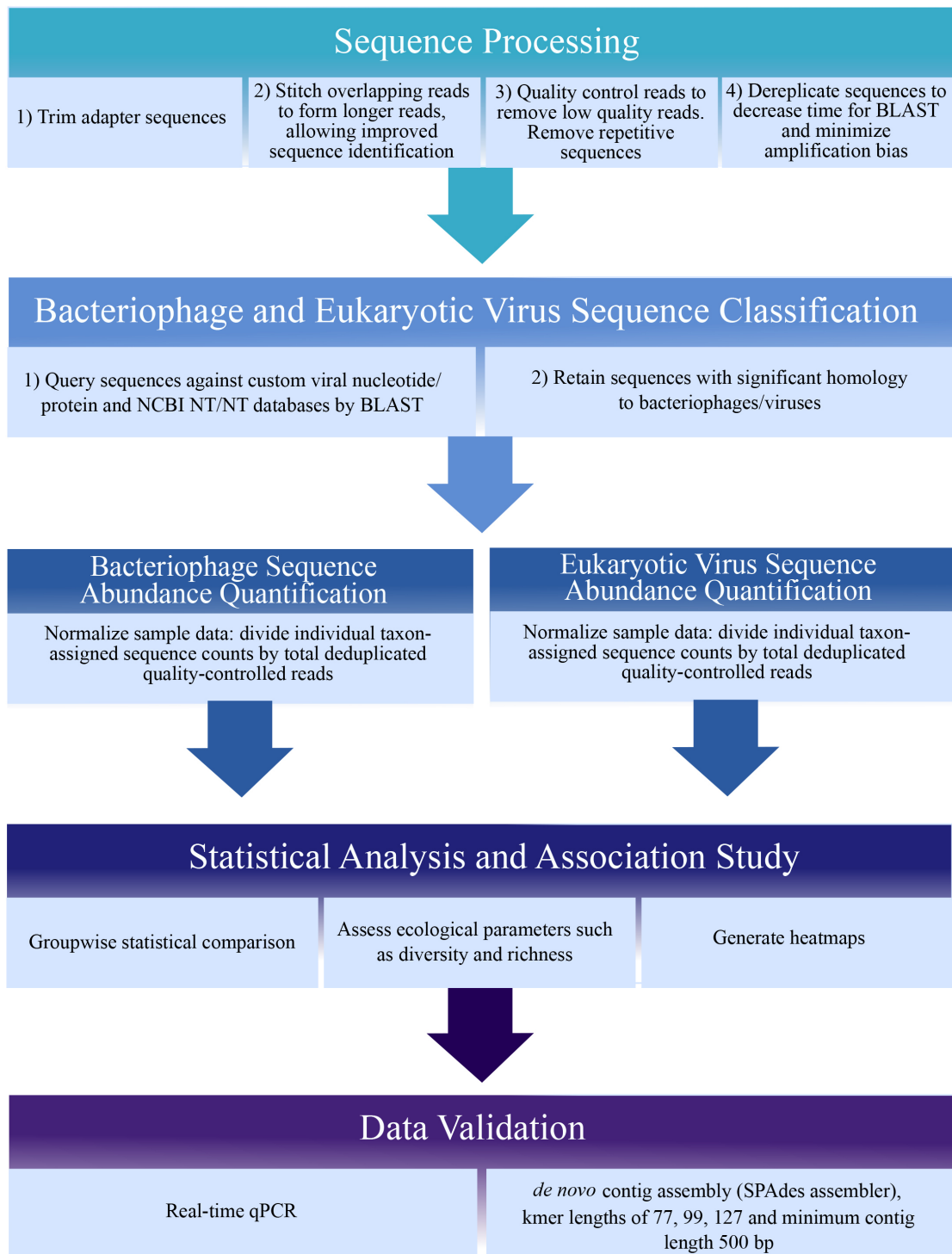


Figure 1. Sequence analysis schematic

### Notes

Stool pulverization is performed in order to evenly distribute microbial matter in the sample. The use of SM buffer containing gelatin stabilizes bacteriophage populations after freezing for further characterization of bacteriophage of interest, including culturing. Due to risk of contamination,

reagents should be used only for fecal microbiome studies. Use the same reagents for all Illumina Miseq runs (*i.e.*, same bottle of sterile water, kits, *etc.*) as many reagents may be contaminated by microbial DNA, and this ensures even contamination across runs.

## **Recipes**

### 1. CTAB/NaCl

3.5 ml 5 M NaCl (final conc. 0.7 M)

2.5 g CTAB (final conc. 10%)

12 ml nuclease-free H<sub>2</sub>O

Final volume: 25ml

Heat to 55 °C to dissolve

Add water to final volume of 25 ml. Then 0.45 µm followed by 0.22 µm filter

## **Acknowledgments**

CLM was supported by NIH training grant 5T32AI007172-34, and the work was funded by R24 ODO19793, R01 OD011170, R01 AI111918, and R01 DK101354. DSK is supported by the Burroughs Wellcome Fund. The stool pulverization protocol is adapted from a protocol generously provided by the Jeffrey Gordon laboratory at Washington University in St. Louis. The VLP protocol was adapted from a protocol for phage isolation (Reyes *et al.*, 2013). We would like to thank Brian Keller, M.D., Ph. D for critical review of this manuscript.

## **References**

1. Duerkop, B. A. and Hooper, L. V. (2013). [Resident viruses and their interactions with the immune system.](#) *Nat Immunol* 14(7): 654-659.
2. Handley, S. A., Thackray, L. B., Zhao, G., Presti, R., Miller, A. D., Droit, L., Abbink, P., Maxfield, L. F., Kambal, A., Duan, E., Stanley, K., Kramer, J., Macri, S. C., Permar, S. R., Schmitz, J. E., Mansfield, K., Brenchley, J. M., Veazey, R. S., Stappenbeck, T. S., Wang, D., Barouch, D. H. and Virgin, H. W. (2012). [Pathogenic simian immunodeficiency virus infection is associated with expansion of the enteric virome.](#) *Cell* 151(2): 253-266.
3. Mokili, J. L., Rohwer, F. and Dutilh, B. E. (2012). [Metagenomics and future perspectives in virus discovery.](#) *Curr Opin Virol* 2(1): 63-77.
4. Monaco, C. L., Gootenberg, D. B., Zhao, G., Handley, S. A., Ghebremichael, M. S., Lim, E. S., Lankowski, A., Baldrige, M. T., Wilen, C. B., Flagg, M., Norman, J. M., Keller, B. C., Luevano, J. M., Wang, D., Boum, Y., Martin, J. N., Hunt, P. W., Bangsberg, D. R., Siedner, M. J., Kwon, D. S. and Virgin, H. W. (2016). [Altered virome and bacterial microbiome in human](#)

- [immunodeficiency virus-associated acquired immunodeficiency syndrome](#). *Cell Host Microbe* 19(3): 311-322.
5. Norman, J. M., Handley, S. A., Baldrige, M. T., Droit, L., Liu, C. Y., Keller, B. C., Kambal, A., Monaco, C. L., Zhao, G., Fleshner, P., Stappenbeck, T. S., McGovern, D. P., Keshavarzian, A., Mutlu, E. A., Sauk, J., Gevers, D., Xavier, R. J., Wang, D., Parkes, M. and Virgin, H. W. (2015). [Disease-specific alterations in the enteric virome in inflammatory bowel disease](#). *Cell* 160(3): 447-460.
  6. Reyes, A., Wu, M., McNulty, N. P., Rohwer, F. L. and Gordon, J. I. (2013). [Gnotobiotic mouse model of phage-bacterial host dynamics in the human gut](#). *Proc Natl Acad Sci U S A* 110(50): 20236-20241.
  7. Virgin, H. W. (2014). [The virome in mammalian physiology and disease](#). *Cell* 157(1): 142-150.
  8. Zhao, G., Krishnamurthy, S., Cai, Z., Popov, V. L., Travassos da Rosa, A. P., Guzman, H., Cao, S., Virgin, H. W., Tesh, R. B. and Wang, D. (2013). [Identification of novel viruses using VirusHunter--an automated data analysis pipeline](#). *PLoS One* 8(10): e78470.
  9. Zhao, G., Wu, G., Lim, E.S., Droit, L., Krishnamurthy, S., Barouch, D.H., Virgin, H.W., and Wang, D. (2017). [VirusSeeker, a computational pipeline for virus discovery and virome composition analysis](#). *Virology* 503: 21-30.