Chapter 7

Phageome analysis of bifidobacteria-rich samples

Brian McDonnell[1], Eoghan Casey[1], Christian Milani[2,3], Gabriele Andrea Lugli[2], Alice Viappiani[3], Jennifer Mahony[1], Marco Ventura[2] & Douwe van Sinderen[1*]

[1]APC Microbiome Institute and School of Microbiology, Bioscience Institute, National University of Ireland, Cork, Ireland
[2] Laboratory of Probiogenomics, Department of Chemistry, Life Sciences, and Environmental Sustainability, University of Parma, Parma, Italy
[3]GenProbio srl, Parma, Italy

*Corresponding author: d.vansinderen@ucc.ie

Abstract

Bifidobacteria are important early colonisers of the human intestinal tract. The relative abundance of bifidobacterial species may be modulated, in part, by bacteriophage activity. Metagenomic studies of these populations is a crucial step in understanding this important interaction. This chapter outlines the technical instructions required to analyse the virome DNA sequences of a bifidobacteria-rich sample, e.g. an infant fecal sample.

Key words

Fecal, phage, metagenome, DNA, sequencing

1. Introduction

Bifidobacteria have been well established as early colonisers of the human gastrointestinal tract, particularly in vaginally delivered and breast-fed infants (1, 2). Persisting (though decreasing in abundance) in the human gut into adulthood (3, 4), the relative abundance of members of the *Bifidobacterium* genus in the gut can furthermore be associated with specific pathologies (for a recent review, see (5)).

Once established in the infant gut, population dynamics of bifidobacteria are likely to be modulated by the availability of nutrients such as human milk oligosaccharides (HMOs), which they metabolise (6, 7), as well as competition with other bacterial genera and indeed, phyla (4, 8). Other key modulators of the bifidobacterial population are likely to be (bacterio)phages. Phages are considered to be ubiquitous in the environment and are well known to exert considerable influence in the dairy environment, where they are responsible for slow and/or failed fermentations (9-11). Furthermore, bacteriophages have been developed as biocontrol agents in the areas of food preservation (12-15) and antibiotic resistant bacterial infections (16, 17). Estimates place phages as the second most abundant domain of life in the gut, after bacteria (for a recent review, see reference 18) and, as such, it is reasonable to assume that phages influence the abundance and diversity of bifidobacterial (amongst other) populations in the human gut.

After initially adsorbing to their host, phages may follow either the lytic (resulting in lysis of the host cell and release of phage progeny) or lysogenic (resulting in the phage genome being incorporated into that of its host) life cycles (19). Clearly, the lytic phage life cycle would seem to be more relevant in terms of modulating host populations. However, phages may also be 'induced' from the lysogenic cycle to follow the lytic cycle, which influences population dynamics in the gut (20). Recently, it has been shown that prophages (i.e. phage genomes present on the host genome) harboured by human gut-associated bifidobacteria have the

potential to be induced to form virions (21). Furthermore, these phages encode a novel receptor-binding protein (RBP) 'shufflon', potentially enabling them to infect a number of bifidobacterial strains/species, thereby substantially expanding their host range and detrimental impact on these populations (21).

At present there is a bottleneck in current techniques relating to bifidobacterial phages, in that plaque-forming units (or suitable permissive hosts) have not yet been identified (21, 22). This necessitates the examination of these phages through population analyses such as metagenomic studies of appropriate material, e.g. infant fecal samples (23). This chapter describes the necessary techniques allowing the researcher to firstly isolate virus-like particles (VLPs) from starting material, before extracting viral DNA (while minimising contamination from host DNA). Viral DNA is then quantified, sequenced and analysed using *in silico* techniques.

2. Materials

Distilled or autoclaved water

Tris base

Hydrochloric acid

Sodium Hydroxide

Calibrated pH meter

Stir bar

Stir plate

Sodium chloride

Magnesium sulphate

Gelatin

Dithiothreitol (DTT; optional)

Proteinase K (as powder; optional)

Autoclave

15 mL centrifuge tube

20, 200, 1000 µL and 10 mL single channel pipettes, and tips

200 and 1000 µL multichannel pipettes, and tips

Vortex apparatus

Timer

Ice

Centrifuge with capacity for 1.5 mL, 15 mL, 50 mL tubes and 96-well plates

0.45 µm filters

Sterile disposable syringe

37°C incubator

DNase I

Heat block

Norgen phage DNA isolation kit

96 - 100% ethanol

Proteinase K

Isopropanol

Qubit fluorimeter

Qubit dsDNA HS assay kit

96-well storage plates, round well, 0.8 mL (midi plates)

Agencourt AMPure XP, 60 or 5 mL kit

Microseal ‘A’ film, ‘B’ adhesive seals, and ‘F’ foil seals, Bio-Rad or equivalent

Rubber roller (plate sealer)

Disposable nuclease-free multichannel reagent reservoirs

Hard-shell 96-well PCR plates, Bio-Rad or equivalent

Microplate shaker capable of 1800 rpm

Magnetic stand

Thermal cycler (PCR)

HT1 Hybridisation buffer (P/N: 20015892; Illumina, Inc.)

Agilent 2100 Bioanalyser (Agilent; optional) and High sensitivity DNA kit (Agilent; optional).

PhiX DNA control (Illumina, Inc.)

DNA sequencing system, e.g. MiSeq or NextSeq (Illumina, Inc.)

3. Methods

3.1 Tris-HCl (pH 7.5) buffer preparation

1. Prepare Tris-HCl buffer by weighing the appropriate amount of Tris base to achieve a concentration of 1 M in a final volume of, e.g., 500 mL.
2. Add the above to an appropriate container and sterile or distilled water to approximately 75% of the final desired volume.
3. Swirl the solution to mix.
4. Adjust the pH of the solution to 7.5 ± 0.5 by adding hydrochloric acid in a dropwise manner (see note 1), measuring the pH intermittently using a calibrated pH meter.
5. QS to volume using sterile or distilled water.
6. Autoclave the solution at 121°C for 15 minutes and allow to cool to room temperature prior to use.

3.2 SMG buffer preparation

1. Sodium chloride magnesium sulphate (SMG) buffer is prepared by weighing the appropriate amounts of sodium chloride, magnesium sulphate, Tris-HCl and gelatin to achieve the following concentrations in a final volume of, e.g. 500 mL: 200 mM (NaCl), 10 mM ($MgSO_4$), 50 mM (Tris-HCl) and 0.01 % w/v (gelatin).
2. Note: the above components can be scaled up or down depending on requirements.
3. Add the above components to an appropriate container and QS to volume using sterile or distilled water.
4. Swirl the solution to mix until the components are dissolved (see note 2).
5. Autoclave the solution at 121°C for 15 minutes and allow to cool to room temperature prior to use.

3.3 Proteinase K solution preparation

1. Weigh an appropriate amount of proteinase K powder to achieve a concentration of 20 mg/mL in a desired final volume (e.g., 20 mg in 1 mL sdH_2O).
2. Add the desired volume of sdH_2O and mix well by inversion until fully dissolved.
3. Filter sterilise the solution using 0.2 µm filters.
4. Aliquot into 100 µL amounts and store at -20°C until required.

3.4 Virus-like particle (VLP) isolation (adapted from (23) and (24))

1. Weigh 0.5 g of target environmental or complex sample into a suitable container (e.g. a 15 mL plastic centrifuge tube).
2. Add 1.2 mL of SMG buffer (as above) to the same container.
3. Vortex the solution at full speed for 2 minutes, ensuring that the sample is fully dissolved/dispersed/suspended in the buffer.
4. Incubate the solution/suspension on ice for 1 hour. Pre-cool the centrifuge chamber to 4°C during this time.
5. Centrifuge the tube at 2500 x *g* for 5 minutes (at 4°C), using an appropriate weight balance.
6. Being careful not to disturb the pellet, remove the tube from the centrifuge. Carefully pipette or decant the supernatant into a fresh centrifuge tube.
7. Centrifuge the supernatant at 5000 x *g* for 15 minutes (at 4°C), using an appropriate weight balance.
8. Being careful not to disturb the pellet, remove the tube from the centrifuge. Carefully pipette or decant the supernatant into a fresh container (e.g. a 1.5 mL tube).

9. OPTIONAL: Add dithiothreitol (DTT) to the solution to a final concentration of 6.5 mM and incubate at 37°C for one hour. This optional step may be performed to encourage denaturation of proteins present in these complex samples.
10. Filter the solution using a 0.45 µm filter.
11. Treat the solution with 10 U/mL DNase I for one hour at room temperature.
12. Inactivate the DNase by incubating the solution at 75°C for 10 minutes. Allow to cool to room temperature (RT) prior to DNA extraction.

3.5 Extraction of bacteriophage (phage) viral DNA

1. DNA is extracted using the Norgen Phage DNA isolation kit (Norgen Biotek, Ontario, Canada; see note 3). Prepare Wash Solution A in advance by adding 90 mL of 96 - 100% ethanol to the supplied Wash Solution A bottle and mix well by inversion.
2. Add the entire VLP-containing sample (approximately 1 mL; obtained following procedure 3.4 as described above) to an appropriate centrifuge tube, e.g. one of 15 mL capacity.
3. Add 500 µL Lysis Buffer B (supplied) to the sample and vortex vigorously for 10 seconds.
4. OPTIONAL: Add 4 µL proteinase K (20 mg/mL, as prepared above) to the sample and incubate at 55°C for 15 minutes. Proteinase K is commonly used in this manner to assist in the lysis of the viral capsid proteins, allowing the DNA to be isolated.
5. Incubate the sample at 65°C for 15 minutes, mixing by inversion 2-3 times during that period.
6. Add 320 µL isopropanol to the sample, and vortex briefly to mix.

7. Assemble a spin column by insertion into a collection tube (provided). Apply 650 µL (maximum) of the sample to the centre of the column and centrifuge for 1 minute at 6000 x *g*.
8. Ensuring that the applied sample has completely passed through the column and into the collection tube, discard the flow-through, while retaining the collection tube.
9. Repeat steps 7 and 8 above until the remainder of the sample has passed through the column.
10. Wash the column by adding 400 µL of Wash Solution A (as prepared above) to the column and centrifuging for 1 minute at 6000 x *g*. Discard the flow-through and replace the spin column in the same collection tube.
11. Repeat step 10 above twice.
12. Perform a final centrifugation step at 14000 x *g* for 2 minutes without the addition of Wash Solution A in order to remove any residual liquid. Discard the collection tube and place the column into a fresh 1.7 mL elution tube (provided).
13. Elute the isolated DNA by adding 75 µL of Elution Buffer B to the centre of the column, and centrifuging for 1 minute at 6000 x *g*.
14. OPTIONAL: In order to increase DNA yield a second elution (in a separate elution tube) can be performed by repeating step 13.
15. Store the isolated DNA-containing solution at -80°C until required.

3.6 Quantification of viral DNA

1. Quantification of (presumed) viral DNA is performed using the Qubit 2.0 Fluorometer and Qubit dsDNA HS Assay Kit (Life Technologies, Bleiswijk, Netherlands; see note 3).

2. The dsDNA HS assay requires two standards; it is recommended that these are prepared on the day of use. Standards and samples are prepared in Qubit Assay tubes or similar, which must be labelled on the tube lids (rather than the sides). Carefully label the amount of tubes required to measure two standards and the number of samples prepared.
3. The Qubit dsDNA HS reagent (supplied) must be protected from light. Ensure that this is the case by wrapping the tube with aluminium foil or similar.
4. Prepare a 1:200 dilution of the HS Reagent in HS Buffer (supplied) to yield Qubit Working Solution. The amount of Qubit Working Solution prepared should be sufficient to allow a 180-199 µL volume to be added to each standard tube and sample tube, for a total volume of 200 µL in each tube.
5. Prepare Qubit standards by adding 190 µL of Qubit Working Solution to each tube, followed by 10 µL of each standard to the appropriate tube. Vortex briefly to mix, avoiding bubble formation.
6. Add 1-20 µL sample (depending on the amount available) to each sample tube. Add the appropriate volume of Qubit Working Solution so that the final volume in each sample tube is 200 µL.
7. Vortex each sample briefly to mix, avoiding bubble formation. Incubate each sample at RT for two minutes.
8. On the 'Home' screen of the Qubit 2.0 Fluorometer, press 'DNA', then press 'dsDNA High Sensitivity' (i.e. the assay type). The standards screen is then displayed.
9. Press 'Yes' to read new standards. Insert the 'Standard #1' tube into the sample chamber, close the lid, and press 'Read'. When the reading is complete (approximately three seconds), remove the 'Standard #1' tube.
10. Repeat step 9 above using the 'Standard #2' tube. The instrument should now display the 'Sample' screen.

11. Insert the first sample tube into the sample chamber, close the lid, and press ‘Read’. When the reading is complete (approximately three seconds), remove the sample tube.
12. Repeat step 11 above for the remainder of the sample tubes.
13. The instrument now displays the sample concentration(s) in ng/mL, corresponding to the concentration of the diluted sample. In order to calculate the concentration of the undiluted sample, apply the following formula, where ‘x’ is the number of microliters of sample added to the assay tube:

    Concentration of neat sample = Conc. given by Qubit * 200 / x
14. The above formula generates the concentration of the undiluted sample in ng/mL.

3.7 Preparation of DNA sample for sequencing

1. Prepare DNA samples for sequencing using the Nextera XT DNA Library Prep reference guide (Illumina Inc., CA, U.S.A.), Document #15031942; v05. Consult this guide for complete information relating to this protocol.
2. Ensure that the input DNA does not contain > 1 mM EDTA, and is free of organic contaminants. Ideally, the input sample should have an OD_{260nm} / OD_{280nm} (UV absorbance) ratio of between 1.8 and 2.0, and an OD_{260nm} / OD_{230nm} ratio of between 2.0 and 2.2. These ratios may be determined through the use of a spectrophotometer.
3. <u>‘Tagmentation’ of DNA sample</u>
    a. *Preparation of reagents*: Defrost the ATM and TD reagents on ice, and centrifuge (pulse) briefly. Check the NT reagent for precipitates - if present, vortex until fully resuspended.
    b. Add the following to each well of a 96-well PCR plate: 10 µL TD, 1 ng DNA (5 µL). Pipette up and down briefly to mix.
    c. Add 5 µL ATM to each well. Pipette up and down 10 times to mix.

d. Seal the plate using Microseal 'A' adhesive film and a rubber roller, and centrifuge the plate at 280 x *g* for 1 minute.

e. Run the sealed plate in the following PCR program with a 50 µL reaction volume:

Step	Temperature	Duration
Heated lid	100°C	∞
Incubation	55°C	10 min
Store	10°C	∞

f. When the sample temperature is at 10°C, immediately add 5 µL NT to each well. Pipette up and down 10 times to mix.

g. Centrifuge the plate at 280 x *g* for 1 minute. Incubate the plate at RT for 5 minutes.

4. Amplification of Libraries

a. *Preparation of reagents*: Defrost the index adapters. If provided in tubes, vortex to mix, then centrifuge briefly before use. If provided in plates, centrifuge briefly before use. Thaw NPM on ice for 20 minutes.

b. Depending on the index adapter kit in use, add either (i) 5 µL i7 then 5 µL i5 adapters (tubes); or (ii) 10 µL pre-paired i7 and i5 adapters to each sample.

c. Add 15 µL NPM to each well. Pipette up and down 10 times to mix, then seal the plate using Microseal 'A' adhesive film and a rubber roller.

d. Centrifuge at 280 x *g* at 20°C for 1 minute.

e. Place the plate on a PCR cycler and run the following program with a 50 µL reaction volume:

Step	Temperature	Duration
Heated lid	100°C	∞
Incubation	72°C	3 min

Incubation	95°C	30 sec
Cycle (x12)	95°C	10 sec
	55°C	30 sec
	72°C	30 sec
Incubation	72°C	5 min
Store	10°C	< 24 hr

f. Safe stopping point: store samples on the thermocycler for up to 24 hours, or at 2-8°C for up to 48 hours.

5. Clean-up of libraries

a. *Preparation of reagents*: Incubate the AMPure XP beads at RT for 30 minutes. Thaw resuspension buffer (RSB) at RT and vortex briefly to mix (RSB may be stored at 2-8°C thereafter). Prepare fresh 80% ethanol (EtOH) using nuclease-free water.

b. Centrifuge the plate at 280 x *g* at 20°C for 1 minute. Transfer 50 µL supernatant from each well of the plate to corresponding wells of a fresh midi (96-well 0.8 mL deepwell storage) plate.

c. Add 30 µL AMPure XP beads to each midi plate sample well.

d. Seal the plate using Microseal 'B' adhesive seal and a rubber roller. Place the plate on a shaker at 1800 rpm for 2 minutes.

e. Incubate the plate at RT for 5 minutes, then place on a magnetic stand until the liquid is clear (approximately 2 minutes).

f. Without disturbing the beads, remove and discard the supernatant using a pipette. If beads are accidentally aspirated into pipette tips, re-dispense and repeat 2 minute incubation on magnetic stand.

g. Wash the beads by adding 200 µL 80% EtOH while the plate is on the magnetic stand. Do not mix. Incubate for 30 seconds, then remove and discard the supernatant.

h. Repeat step g. above once.

i. Remove residual EtOH using a 20 µL pipette and allow to air dry on the magnetic stand for 15 minutes.

j. Remove the plate from the magnetic stand. Add 52.5 µL RSB to the beads.

k. Seal the plate using Microseal 'B' plates and a rubber roller. Place the plate on a shaker at 1800 rpm for 2 minutes, then incubate at RT for 2 minutes.

l. Place the plate on a magnetic stand and wait until the liquid is clear (approximately 2 minutes).

m. Transfer 50 µL of the supernatant to a fresh 96-well PCR plate.

n. Safe stopping point: The plate may be sealed with either Microseal 'B' adhesive seal or Microseal 'F' foil seal and stored at -20°C for up to 7 days.

6. Library quality check

   a. The quality of the library may be checked using an Agilent 2100 Bioanalyser and the High Sensitivity DNA kit, by loading 1 µL undiluted library. Typical libraries show a broad size distribution of approximately 250 to 1000 bp, but fragment sizes as large as 1500 bp may be sequenced.

7. Normalisation of libraries

   a. *Preparation of reagents*: Bring LNA1 to RT, vortexing vigorously to ensure all precipitates are dissolved (**Note: LNA1 should be used in a fume hood**). Bring LNB1 to RT, vortexing vigorously with intermittent inversion, ensuring that all beads are resuspended and no beads remain at the bottom of the tube after inversion. Bring LNW1 and LNS1 to RT.

   b. Transfer 20 µL supernatant from each well of the PCR plate to the corresponding wells of a fresh midi plate.

c. Prepare LN master mix by combining [(46 µL LNA1 + 8 µL LNA2)*Number of samples] in a 15 mL conical tube, pipetting up and down 10 times to mix. Transfer this master mix to an appropriate trough.

d. Using a multichannel pipette, transfer 45 µL master mix to each well of the PCR plate.

e. Seal the plate using Microseal 'B' adhesive seals and place on a shaker at 1800 rpm for 30 minutes.

f. Place the plate on a magnetic stand until the liquid is clear (approximately 2 minutes). Without disturbing the beads, remove and discard the supernatant.

g. Wash the beads by adding 45 µL LNW1 to each well. Seal the plate using Microseal 'B' adhesive seals, and place on a shaker at 1800 rpm for 5 minutes.

h. Return the plate to the magnetic stand, and wait until the liquid is clear (approximately 2 minutes). Remove and discard the supernatant.

i. Repeat steps g. and h. above once.

j. Add 30 µL 0.1 N NaOH (freshly prepared within the last 7 days) to each well. Seal the plate using Microseal 'B' adhesive seals and place on a shaker at 1800 rpm for 5 minutes.

k. Add 30 µL LNS1 to each well of a fresh 96-well PCR plate labelled 'SGP' (storage plate).

l. Examine the midi plate to ensure that all samples are resuspended. If further resuspension is required, pipette each sample up and down 10 times to mix, seal the plate, and repeat the shaking portion of step j. above.

m. Place the plate on a magnetic stand and wait until the liquid is clear (approximately 2 minutes). Transfer 30 µL of the supernatant from each well of the midi plate to corresponding wells of the SGP plate.

n. Seal the sample plate and centrifuge at 1000 x *g* for 1 minute.

o. Safe stopping point: The plate may be stored at -20°C for up to 7 days.

8. Dilution of libraries to starting concentration

    a. Calculate the molarity (nM) value of the library using the below formula, where 'bp' indicates the average library size:

$$\frac{\textbf{ng/µL} * 10^5}{\textbf{660 g/mol} * \textbf{bp}}$$

    Note: the library size can be calculated using a Bioanalyzer (as described in step 6 above), alternatively, a generic value of 600 bp may be used.

    b. Using the molarity value, calculate the volumes of RSB and library needed to dilute libraries to the starting concentration for the system in use. For most systems, this may be found in the 'denature and dilute libraries' guide.

3.8 DNA Sequencing

1. Combine the following volumes of sample DNA and freshly diluted 0.2 N NaOH in a microcentrifuge tube:
    a. 4 nM sample DNA (5 µl);
    b. 0.2 N NaOH (5 µl).
2. Vortex briefly to mix the sample solution, and centrifuge the sample solution at 280 x *g* for 1 minute.
3. Incubate for 5 minutes at RT to denature the DNA.
4. Combine 10 µL Denatured DNA and 990 µL pre-chilled (4°C) HT1 hybridisation buffer in a fresh tube. This results in a 20 pM denatured library in 1 mM NaOH. This results in a 20 pM denatured library in 1 mM NaOH.
5. Place the denatured DNA on ice until you are ready to proceed to final dilution.

6. Dilute the denatured DNA to achieve a 4 nM library concentration, using the below table to give a final volume of 600 µl of the desired input concentration.

Final concentration	**6 pM**	**8 pM**	**10 pM**	**12 pM**	**15 pM**	**20 pM**
20 pM library	180 µl	240 µl	300 µl	360 µl	450 µl	600 µl
Pre-chilled HT1	420 µl	360 µl	300 µl	240 µl	150 µl	0 µl

7. Invert several times to mix and then pulse centrifuge the DNA solution. Place the denatured and diluted DNA on ice until you are ready to your samples on Illumina reagent cartridge.
8. Add denatured PhiX DNA control (to a final concentration of 5%) to the denatured and diluted DNA.
9. Pooled samples, loaded in the sequencing cartridge of the corresponding Illumina sequencer, i.e. MiSeq or NextSeq sequencing system, are then sequenced.
10. Follow the instructions given by the software interface of the selected sequencing system.

3.9 *In silico* DNA sequence analysis

1. DNA quality filtering
    a. Demultiplexed reads obtained from DNA sequencing are collected in two .fastq format files for each index by using the BaseSpace Sequence Hub from the Illumina platform services manager (https://basespace.illumina.com).
    b. A filtering step is performed to improve the quality of the sequenced reads employing the fastq-mcf script (https://github.com/ExpressionAnalysis/ea-utils) (see Chapter 4 of this book for further details).
2. Phage genome assembly

a. Fastq files obtained from the DNA quality filtering step are used as input material to perform a *de novo* assembly of the phage genome sequence.
b. Filtered reads are then assembled into contigs. we suggest the use of an open source sequence assembler program such as SPAdes v3.14.0 (25) (see Chapter 5 for more detail).

3. Gene prediction
    a. Contigs of phage sequences obtained from the genome assembly are then collected in .fasta file format.
    b. The prediction of Open Reading Frames (ORFs) across the phage sequence is performed using a gene-finding program such as the Prokaryotic Dynamic Programming Genefinder Algorithm (Prodigal) v2.6.3 (26) (see Chapter 5 for more detail).
    c. Transfer RNA genes are identified using tRNAscan-SE v1.4 (27) (see Chapter 4 for further details).
4. Gene annotation
    a. Amino acid sequences of protein-coding genes are used to predict their function by using a reference database such as the RefSeq database of the National Center for Biotechnology Information (NCBI) or specific phage databases like PHAge Search Tool Enhanced Release (PHASTER) database (28).
    b. Similarity search of collected phage ORFs are managed by a multi-threading program such as BLASTP (29) or RAPSearch2 (30), in order to predict the function of each ORF based on the function of identified homologs (see Chapter 5 for more detail).
5. Phage contig selection and validation

a. Annotated contigs are screened to remove potential bacterial genomic DNA 'contaminants' present with the phage sequence by using a whole genome aligner such as MAUVE (31) or a local alignment search tools such as BLAST (29).

b. In cases where each sequenced sample belongs to a single phage genome, resulting sequences after bacterial genome filtering should result in the identification of a single contig, representing the actual phage genome sequence.

c. Employing BLASTN (29), the alignment of nucleotide sequences retrieved from the contig edges allows identification of circular phage sequences (https://blast.ncbi.nlm.nih.gov).

d. By removing the overlapping sequence among contig edges, it is possible to determine the actual length of the phage genome.

6. Phage sequence identification

a. Host prediction based on Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) can be performed employing databases with a collection of spacer sequences such as CRISPRCasdb (32).

b. Taxonomic classification of spacer sequences retrieved from (bifido)bacterial genomes allows the identification of putative host for each phage.

4 References

1. Lim ES, Zhou Y, Zhao G, Bauer IK, Droit L, Ndao IM, et al. Early life dynamics of the human gut virome and bacterial microbiome in infants. Nat Med. 2015;21(10):1228-34.
2. Backhed F, Roswall J, Peng Y, Feng Q, Jia H, Kovatcheva-Datchary P, et al. Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life. Cell Host Microbe. 2015;17(6):852.
3. Matsuki T, Watanabe K, Fujimoto J, Miyamoto Y, Takada T, Matsumoto K, et al. Development of 16S rRNA-gene-targeted group-specific primers for the detection and identification of predominant bacteria in human feces. Appl Environ Microbiol. 2002;68(11):5445-51.
4. Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, et al. Human gut microbiome viewed across age and geography. Nature. 2012;486(7402):222-7.
5. Mahony J, Lugli GA, van Sinderen D, Ventura M. Impact of gut-associated bifidobacteria and their phages on health: two sides of the same coin? Appl Microbiol Biotechnol. 2018;102(5):2091-9.
6. James K, Bottacini F, Contreras JIS, Vigoureux M, Egan M, Motherway MO, et al. Metabolism of the predominant human milk oligosaccharide fucosyllactose by an infant gut commensal. Sci Rep. 2019;9(1):15427.
7. James K, Motherway MO, Bottacini F, van Sinderen D. *Bifidobacterium breve* UCC2003 metabolises the human milk oligosaccharides lacto-N-tetraose and lacto-N-neo-tetraose through overlapping, yet distinct pathways. Sci Rep. 2016;6:38560.
8. Koenig JE, Spor A, Scalfone N, Fricker AD, Stombaugh J, Knight R, et al. Succession of microbial consortia in the developing infant gut microbiome. Proc Natl Acad Sci U S A. 2011;108 Suppl 1:4578-85.
9. Moineau S, Borkaev M, Holler BJ, Walker SA, Kondo JK, Vedamuthu ER, et al. Isolation and characterization of lactococcal bacteriophages from cultured buttermilk plants in the United States. Journal of Dairy Science. 1996;79(12):2104-11.
10. Lavelle K, Murphy J, Fitzgerald B, Lugli GA, Zomer A, Neve H, et al. A decade of *Streptococcus thermophilus* phage evolution in an Irish dairy plant. Appl Environ Microbiol. 2018;84(10):e02855-17.
11. Garneau JE, Moineau S. Bacteriophages of lactic acid bacteria and their impact on milk fermentations. Microb Cell Fact. 2011;10 Suppl 1:S20.
12. Guenther S, Loessner MJ. Bacteriophage biocontrol of *Listeria monocytogenes* on soft ripened white mold and red-smear cheeses. Bacteriophage. 2011;1(2):94-100.
13. Shahin K, Bouzari M. Bacteriophage application for biocontrolling *Shigella flexneri* in contaminated foods. J Food Sci Technol. 2018;55(2):550-9.
14. Guenther S, Herzig O, Fieseler L, Klumpp J, Loessner MJ. Biocontrol of *Salmonella Typhimurium* in RTE foods with the virulent bacteriophage FO1-E2. Int J Food Microbiol. 2012;154(1-2):66-72.
15. Bao H, Zhang P, Zhang H, Zhou Y, Zhang L, Wang R. Bio-Control of *Salmonella Enteritidis* in Foods Using Bacteriophages. Viruses. 2015;7(8):4836-53.
16. Kabwe M, Brown T, Speirs L, Ku H, Leach M, Chan HT, et al. Novel Bacteriophages Capable of Disrupting Biofilms From Clinical Strains of *Aeromonas hydrophila*. Front Microbiol. 2020;11:194.
17. Tkhilaishvili T, Wang L, Tavanti A, Trampuz A, Di Luca M. Antibacterial Efficacy of Two Commercially Available Bacteriophage Formulations, Staphylococcal Bacteriophage and PYO Bacteriophage, Against Methicillin-Resistant *Staphylococcus aureus*: Prevention and Eradication of Biofilm Formation and Control of a Systemic Infection of *Galleria mellonella* Larvae. Front Microbiol. 2020;11:110.
18. Shkoporov AN, Hill C. Bacteriophages of the Human Gut: The "Known Unknown" of the Microbiome. Cell Host Microbe. 2019;25(2):195-209.
19. Sturino JM, Klaenhammer TR. Engineered bacteriophage-defence systems in bioprocessing. Nat Rev Microbiol. 2006;4(5):395-404.

20. Reyes A, Wu M, McNulty NP, Rohwer FL, Gordon JI. Gnotobiotic mouse model of phage-bacterial host dynamics in the human gut. Proc Natl Acad Sci U S A. 2013;110(50):20236-41.
21. Mavrich TN, Casey E, Oliveira J, Bottacini F, James K, Franz C, et al. Characterization and induction of prophages in human gut-associated *Bifidobacterium* hosts. Sci Rep. 2018;8(1):12772.
22. Ventura M, Lee JH, Canchaya C, Zink R, Leahy S, Moreno-Munoz JA, et al. Prophage-like elements in bifidobacteria: insights from genomics, transcription, integration, distribution, and phylogenetic analysis. Appl Environ Microbiol. 2005;71(12):8692-705.
23. Milani C, Casey E, Lugli GA, Moore R, Kaczorowska J, Feehily C, et al. Tracing mother-infant transmission of bacteriophages by means of a novel analytical tool for shotgun metagenomic datasets: METAnnotatorX. Microbiome. 2018;6(1):145.
24. Kleiner M, Hooper LV, Duerkop BA. Evaluation of methods to purify virus-like particles for metagenomic sequencing of intestinal viromes. BMC Genomics. 2015;16(1):7.
25. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19(5):455-77.
26. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics. 2010;11:119.
27. Lowe TM, Chan PP. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. Nucleic Acids Res. 2016;44(W1):W54-7.
28. Arndt D, Marcu A, Liang Y, Wishart DS. PHAST, PHASTER and PHASTEST: Tools for finding prophage in bacterial genomes. Brief Bioinform. 2019;20(4):1560-7.
29. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. Journal of Molecular Biology. 1990;215(3):403-10.
30. Zhao Y, Tang H, Ye Y. RAPSearch2: a fast and memory-efficient protein similarity search tool for next-generation sequencing data. Bioinformatics. 2012;28(1):125-6.
31. Darling AC, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res. 2004;14(7):1394-403.
32. Pourcel C, Touchon M, Villeriot N, Vernadet JP, Couvin D, Toffano-Nioche C, et al. CRISPRCasdb a successor of CRISPRdb containing CRISPR arrays and cas genes from complete genome sequences, and tools to download and query lists of repeats and spacers. Nucleic Acids Res. 2020;48(D1):D535-D44.

5 Notes

1. While adjusting the pH, ensure that the volume of the solution does not exceed the final desired volume. If this occurs, discard the solution and repeat - reducing the amount of water added to the container in step 2 above. To ensure an accurate pH reading, stir the solution continuously during pH adjustment, preferably using an appropriately sized stir bar and stir plate.
2. The gelatin added to the solution may not fully dissolved until after autoclaving.
3. For additional guidelines regarding the use of kits, see the corresponding user guidelines supplied by the manufacturer.