Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2019 doi:10.20944/preprints201906.0125.v1

Article

A Protocol for Extraction of Infective Viromes Suitable for Metagenomics Sequencing from Low Volume Fecal Samples

Ling Deng[1,*], Ronalds Silins[1], Josué L. Castro-Mejía[1], Witold Kot[2], Leon Jessen[3], Jonathan Thorsen[3], Shiraz Shah[3], Jakob Stokholm[3], Hans Bisgaard[3], Sylvain Moineau[4,5], and Dennis Sandris Nielsen[1,*]

[1] Section of Food Microbiology and Fermentation, Department of Food Science, Faculty of Science, University of Copenhagen, Rolighedsvej 26, 1958, Frederiksberg, Denmark
[2] Department of Environmental Science, Aarhus University, Frederiksborgvej 399, Roskilde, Denmark COPSAC,
[3] Copenhagen Prospective Studies on Asthma in Childhood, Herlev and Gentofte Hospital, University of Copenhagen, Ledreborg Alle 34, 2820, Gentofte, Denmark
[4] Département de biochimie, de microbiologie, et de bio-informatique, Faculté des sciences et de génie, Université Laval, Québec City, QC G1V 0A6, Canada
[5] Félix d'Hérelle Reference Center for Bacterial Viruses and Groupe de recherche en écologie buccale, Faculté de médecine dentaire, Université Laval, Québec City, QC G1V 0A6, Canada

* Correspondence: lingdeng@food.ku.dk (L.D); dn@food.ku.dk (D.S.N)

Received: date; Accepted: date; Published: date

Abstract: The human gut microbiome (GM) plays an important role in human health and diseases. However, while substantial progress has been made in understanding the role of bacterial inhabitants of the gut, much less is known regarding the viral component of the GM. Bacteriophages (phages) are viruses attacking specific host bacteria and likely play important roles in shaping the GM. Although metagenomic approaches have led to the discoveries of many new viruses, they largely remain uncultured as their hosts have not been identified, which hampers our understanding of their biological roles. Existing protocols for isolation of viromes generally require relatively high input volumes and are generally more focused on extracting nucleic acids of good quality and purity for down-stream analysis and less on purification of still infective viruses. Here we report the development of an efficient protocol requiring low sample input yielding purified viromes containing still infective phages which also are of sufficient purity for genome sequencing. We validated the method through spiking of known phages followed by plaque assays, qPCR and metagenomic sequencing. The protocol should facilitate the culturing of novel viruses from the gut as well as large scale studies on gut viromes.

Keywords: Isolation; Purification; Phage; T4; c2; phiX174; phi29

---

1. Introduction

During the past decades it has become apparent that the human gut microbiome (GM) has profound influence on health and diseases. While most studies investigating the human GM have focused on the bacterial component, there is an emerging understanding that non-bacterial members (archaea, eukaryotes and viruses) have deep impacts on GM structure and function [1-3], as well as host health [4-7], with especially the viruses playing a significant role.

Advances in metagenomics have led to a rapid and massive expansion in the known diversity of viral genomes, but most of these have no identified host and the knowledge of their characteristics is very limited [8-10]. While metagenomics is indispensable for discovery of new viral genomes,

© 2019 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2019 doi:10.20944/preprints201906.0125.v1

functional virology research requires isolation of cultivable viruses and their hosts. Development of efficient protocols for purification of infective viromes from fecal samples is thus essential for detailed studies coupling bacterial hosts and phages. Moreover, many of the reported methods for fecal virome extraction require gram-scale input and long processing time [11-14]. Importantly, these protocols are usually also constrained by the number of samples that can be processed in parallel, which makes large scale studies very tedious.

With the aim of enabling isolation and characterization of infective gut viromes for large scale studies and studies where limited input material is available (i.e. limited biobanked fecal samples or rodent fecal samples), we report here the development of an efficient protocol for extraction of infective viruses from low volume of fecal sample. The isolation of infective phages was validated by spiking the fecal samples with known phages from different viral families and determining phage recovery rates during purification by plaque assays and qPCR. Finally, the extracted viromes were analyzed by shot gun sequencing.

2. Materials and Methods

2.1. Sample collection and storage

Fecal samples were obtained from three anonymous healthy human infants aged around 1 year. The samples were collected in the home of the infants, mixed equally with 2 X SM buffer (400 mM NaCl, 20 mM MgSO4, 100 mM Tris-HCl, pH 7.5) containing 30% glycerol in 50 ml tubes and preserved in cooler bags with ice-packs (temperature 2-5 °C). Samples were delivered to the laboratory within 16 hours, upon reception immediately divided into smaller aliquots (0.5 g) and stored at − 80 °C until further use.

2.2. Virus stock production

The protocol was optimized and validated by spiking fecal samples with known viruses representing four common phage families namely *Podoviridae* (phage Φ29), *Myoviridae* (phage T4), *Siphoviridae* (phage c2) and *Microviridae* (phage ΦX174) (Table 1). *Lactococcus lactis* MG1363, the host of phage c2, was grown in M17 broth (Merck, Germany) containing 5mM $CaCl_2$ and 0.4% glycine at 30 °C. Phage Φ29 host *Bacillus subtilis* DSM 5547 was grown in TS broth (Merck, Germany) at 37°C while shaken at 225 rpm. The host of phage T4 *Escherichia coli* DSM 613 was grown in LB broth (Merck) at 37 °C while shaken at 225 rpm. *E. coli* ATCC 13706, the host of phage ΦX174, was grown in BHI broth (Merck) at 37 °C while shaken at 225 rpm.

Table 1. Bacterial strains and their respective bacteriophages

Bacterial strain	Phage (Family)	Growth media	Source
Bacillus subtilis DSM 5547	Φ29 (*Podoviridae*)	TSB	Lab.stock
Escherichia coli DSM 613	T4 (*Myoviridae*)	LB medium	Lab.stock
Escherichia coli ATTC 13706	ΦX174 (*Microviridae*)	BHI Broth	Félix d'Hérelle Reference Center
Lactococcus lactis MG1363	c2 (*Siphoviridae*)	M17	Lab. Stock

For virus propagation, 100 µl of bacterial overnight culture was added to 2 × 10 ml of broth (Table 1), and grown for 2 hours at 37 °C while shaken at 225 rpm, except for *Lactococcus lactis* MG 1363 which was grown at 30 °C without shaking. After incubation, 50 µl of the respective phage stock lysate was added to one tube of each pair and both tubes were further incubated overnight. The following day the lysed cultures were transferred to a 50 ml tube and centrifuged at 5000 g for 30 minutes at 4 °C to remove cell derbis. The supernatant was recovered and filtered through a 0.45 µm syringe filter and stored at 4 °C. Infective phages in the filtrate were enumerated by plaque assay.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2019 doi:10.20944/preprints201906.0125.v1

2.3. Spiking of fecal samples with known phages

Fecal samples were diluted with 30 ml SM buffer (in a 50 ml centrifuge tube, Sarstedt) and spiked with each phage (Table 1) to a final concentration of 10^4 plaque forming units per milliter (PFU/ml) for each phage. Phage lysates were diluted with SM buffer to obtain the desired titer prior to spiking.

2.4. Plaque assay

To quantify the recovered phages at different purification steps, plaque assays were performed [11]. Prior to plaque assays, spot tests were applied to determine the optimal dilution level for plating. Briefly, 5 ml of media containing 0.5% agarose pre-warmed at 37 °C was mixed with 100 µl of the diluted phage sample and 200 µl of the bacterial culture and poured to the top of a pre-warmed agar plate (1.5%). The double layer plates were first solidified at room temperature and then incubated overnight at the corresponding growth temperature of the bacterial host. On the next day the phage plaques were counted and PFU/ml calculated.

2.5. Virome isolation from faeces

After spiking with known phages, samples were poured into a stomacher filter bag (Interscience BagPage, 100 ml). The mixture was homogenized (Stomacher 80, Seward, UK) for 120 seconds at the high level setting. Homogenized samples, from the other side of the filter in the bag, were transferred to 50 ml tubes and centrifuged at 5000 g for 30 min at 4 °C. After centrifugation, the supernatant was filtered through a 0.45 µm PES filter (Minisart® High Flow Syringe Filter) into the bottom of the outer tube of a Centriprep 50K device (Merck). Afterwards the filtrate was purified and concentrated using the Centriprep 50K device by centrifuging at 1500 g three times in row, first time for 30 min, second time for 10 min, and third time for 3 min. Extra centrifugation time was sometimes applied to allow the liquid level in the inner tube to be similar with outer tube. The liquid filtered into the inner tube was poured off after each centrifugation step. A volume of 200 µl SM buffer was added to the inner tube at the end and centrifuged for 3 min. After the final centrifugation, 140 µl of the concentrated virome solution remaining in the outer tube was collected. The Centriprep filter membrane was cut out and added to the virome solution before storing at -80 °C until nucleic acids extraction. The remaining volume was stored at 4°C for plaque assays.

2.6. Nucleic acid extraction of virome from faeces

The concentrated virome solution and the cut filter membrane was first treated by 1 µl of 100 time diluted Pierce™ Universal Nuclease (Thermofisher Scientific) for 5 min at room temperature, then the QIAmp viral RNA mini kit (Qiagen) was used for viral DNA/RNA extraction following the procedures described by the manufacturer with modifications as described [15]. Next, 10 µl of the extracted nucleic acids were amplified through Multiple Displacement Amplification (MDA) using the Genomephi V3 kit (GE Healthcare Life Sciences) following the instructions of the manufacturer, but the amplification time was shortened to 30 min (from 90 minutes). Finally, the amplified DNA was cleaned using a Genomic DNA Clean & Concentrator™ Kit (Zymo Research) following the manufacture's protocol.

2.7. Virus quantification by quantitative real-time PCR (qPCR)

Phage T4 was also quantified by real-time qPCR using SYBR Green Master Mix (Roche) on 7500 Fast Real-Time PCR System (Applied Biosystems, USA). 5 pmol of forward and reverse primers (5'-CACAGAGGAACGGTCTTGTAAA-3' and 5'-GAGAAGCCCTCCAGAATCATAAA-3' targeting T4 genome) were added to 20 µl reactions, which were run using the following setup: initial stage at 50 °C for 2 min, hot start at 95 °C for 2 min, followed by 40 cycles of (i) 95 °C for 15 sec, (ii) 55 °C for 30 sec and (iii) 72 °C for 30 sec [16]. Serial five times dilutions of T4 genomic DNA was used to

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2019 doi:10.20944/preprints201906.0125.v1

131 generate standard curves. After the qPCR amplification, a melting curve analysis was performed in
132 order to distinguish putative nonspecific amplifications. Each reaction was performed in duplicates.

133 *2.8. Sequencing of fecal virome nucleic acids*

134 The concentration of the MDA amplified and cleaned DNA was measured by Qubit dsDNA HS
135 Assay Kit (ThermoFisher Scientific). Random shotgun libraries were constructed using the Nextera
136 XT kit (Illumina) and normalized by AMPure XP beads following the standard procedures described
137 by the manufactures. Constructed libraries were sequenced using 2 × 150 bp paired-end settings on
138 an Illumina NextSeq platform.

139 *2.9. Processing and analysis of the sequencing results*

140 The sequencing data obtained was processed and analyzed using a pipeline previously
141 described [17]. Briefly, the raw reads were trimmed using Trimmomatic v0.35 (>97%). As quality
142 control, presence of non-viral DNA was quantified using 50000 random forward-reads from each
143 sample, which were queried against the human genome, as well as all the bacterial and viral
144 genomes hosted at NCBI using Kraken2 [18]. For each sample, reads generated from virus-like
145 particles (VLPs)-derived DNA sequencing were subjected to within-sample *de novo* assembly using
146 Spades v3.5.0 [19] and contigs with a minimum length of 1000 nt were retained. Contigs generated
147 from all samples were pooled and de-replicated by multiple blasting and removing those contained
148 in over 90% of the length of another (90% similarity), as outlined previously [20]. Following
149 assembly and quality control, high-quality/de-replicated reads from all samples were merged and
150 recruited against all the assembled contigs at 95% similarity using Subread [21] and a
151 contingency-table of reads per Kbp of contig sequence per million reads sample (RPKM) was
152 generated. The taxonomy of contigs was determined by querying (USEARCH-ublast, e-value 10^{-3})
153 the viral contigs against a database containing taxon signature genes for virus orthologous groups
154 hosted at www.vogdb.org.

155 **3. Results and Discussion**

156 *3.1. Design of the experiments*

157 Since we aimed to isolate infective phages simultaneously with nucleic acids suitable for
158 downstream processing, caution was taken into designing a process where not only the phage
159 particles should be kept intact, but also the receptor-binding fibers used to bind to their bacterial
160 hosts. Taking advantage of the possibility to concentrate VLPs using Centriprep-filters, we chose an
161 approach where the low-input fecal samples (containing approximately 250 mg fecal matter) were

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2019 doi:10.20944/preprints201906.0125.v1

first diluted with 30 ml of buffer before the first homogenization step (Figure 1).

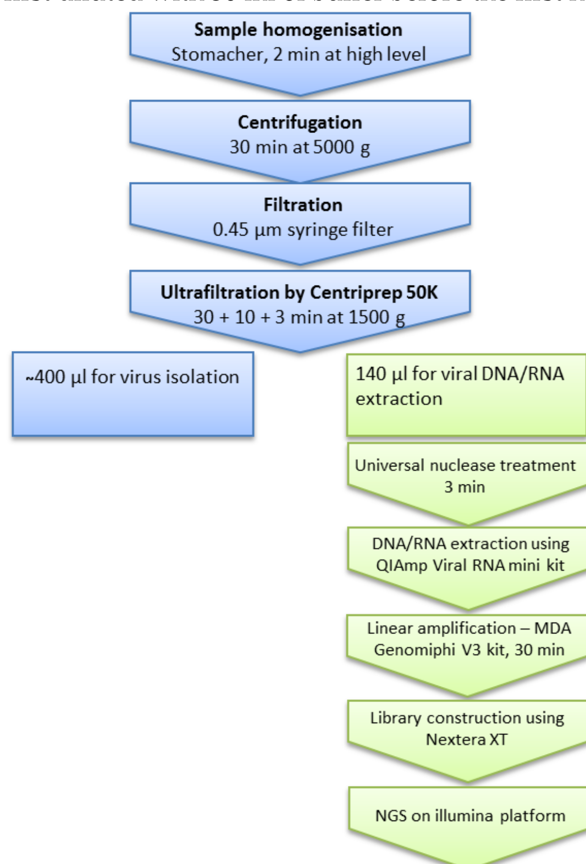


Figure 1. Overview of the virome extraction, amplification and sequencing procedures. A workflow for gut virome extraction and sequencing was established, the virome isolation part and sequencing part is in blue and green, respectively.

As seen from Figure 1, bacterial cells and other larger particles were then pelleted by centrifugation at a modest speed (5000 g), and the supernatant subsequently cleaned by gentle filtration through 0.45 µm pore polyethersulfone (PES) membrane filters. We chose to use 0.45 µm PES filters for easier filtration and maximal recovery of the phages while ensuring removal of bacterial cells [15]. Then, viral particles from the filtrate were concentrated by ultrafiltration using Centriprep 50K tubes to a final volume of approximately 550 µl. Depending on the centrifuge, 16 viromes can be isolated and concentrated simultaneously on a Beckman Allegra 25R refrigerated centrifuge and 24 on an Eppendorf 5920 centrifuge. The total processing time is in both cases less than 4 hours, with a hands-on time less than 2 hours making extraction of 48 samples feasible in one work day with less than 4 hours of hands-on time.

Cesium chloride (CsCl) density gradient centrifugation can yield VLPs of high purity, but was avoided here as it is known to damage phages with fragile tail structures [11,22]. Moreover, CsCl density gradient centrifugation is labor intensive and require lengthy centrifugation steps and consequently the number of samples that can be processed simultaneously is limited [11,12,23]. PEG/NaCl precipitation was also not selected here to concentrate viral particles as optimal PEG/NaCl concentration for precipitation is phage-dependent [24] and using this method could introduce bias into the viral populations after recovery. It has also been reported that chloroform can be added to disrupt the cell membrane, allowing further removal of bacteria and its debris, but enveloped viruses would be removed at the same time [12,13,20,22,23]. Therefore, we chose not to treat the virome samples with chloroform.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2019 doi:10.20944/preprints201906.0125.v1

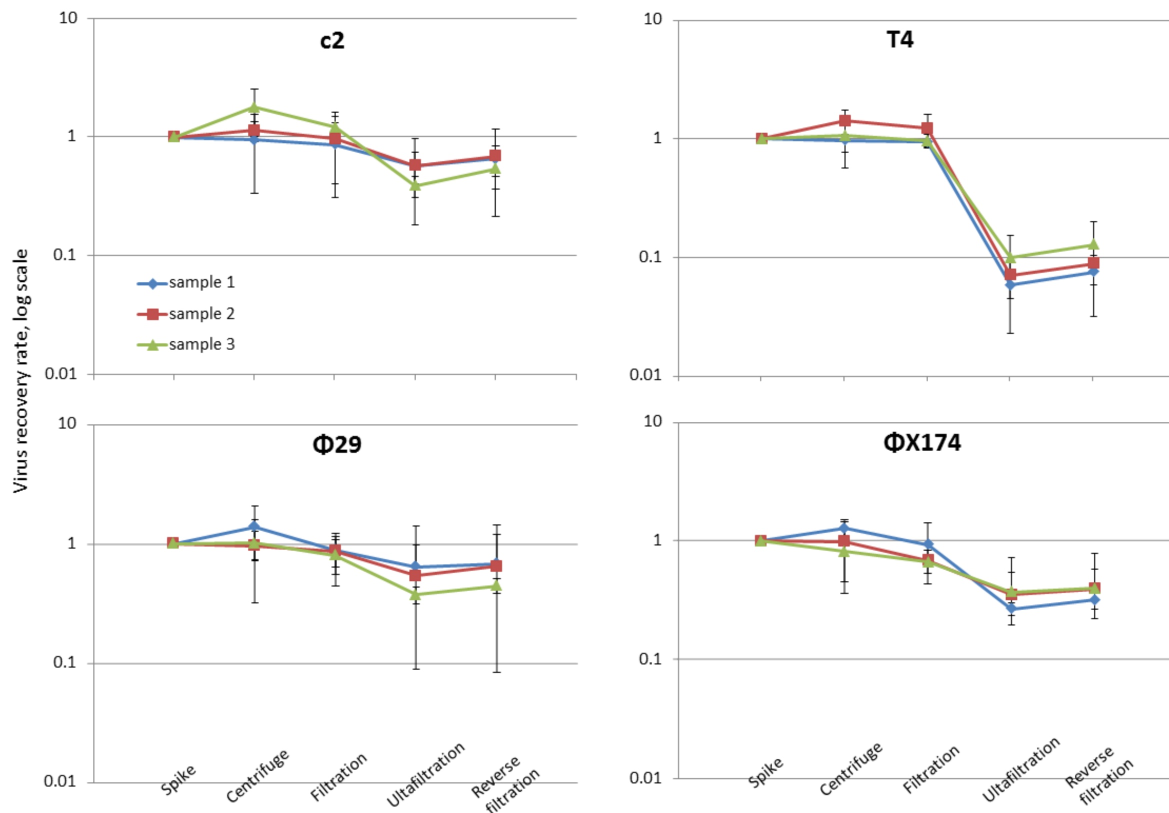
189 *3.2. Assessing the protocol design by recovery rates of spiked phages*

190 Most published virome extraction protocols from fecal samples did not consider the loss of
191 infectivity during the purification procedures [25-28] and those that did require rather high volumes
192 of fecal sample [11]. To estimate the loss of phage infectivity during our proposed extraction
193 procedure, phages T4 (Myoviridae), c2 (Siphoviridae), Φ29 (Podoviridae) and ΦX174 (Microviridae)
194 representing the four most abundant phage families in the human gut were spiked into fecal
195 samples and their recovery rates at the different steps of the extraction protocol were determined by
196 plaque-assays [29]. We first confirmed that no plaques were formed with the host strains (Table 1)
197 using the viromes prepared from non-spiked fecal samples (results not shown). The average final
198 recovery rates for infective phages were on average 63.1% (±6.4%), 9.8% (±2.2), 59.1% (±10.4%) and
199 29.4% (±9.2%) for c2, T4, Φ29 and ΦX174, respectively, with the majority loss of infectivity happening
200 during the concentration procedure (Figure 2). However, a helpful feature of the Centriprep
201 ultrafilter is that it allows reverse flow of the buffer through the membrane when the liquid level of
202 inner tube is higher than that of outer tube, which can wash the attached viruses off the filter
203 membrane. We observed a 2% to 10% increase in the final phage recovery rate after this step. The
204 highest loss (1 log) of infective particles was observed with phage T4. The loss of phage T4 infectivity
205 during extraction from fecal samples is a common challenge and may reflect damage of the fragile
206 fiber structure as suggested earlier [11]. Importantly, the recovery of approximately 10% of infective
207 phages for T4 phages here was at least an order of magnitude higher than previously published
208 protocols [11].

209 *3.3. Determination of T4 genome recovery rate by qPCR*

210 The reduction of infective T4 numbers may mainly be due to the damage of its fragile structure,
211 but it could also because the entire viral particles were lost during the purification process.
212 Therefore, T4-specific qPCR was performed to determine the recovery rate of T4 genomes, as the
213 genomes should still be present as long as the capsid was intact. In accordance with our previous
214 observation [11], the final recovery rates of T4 genomes is much higher when determined by qPCR.
215 For sample 1, 2 and 3, the recovery rate was 21.6% (±1.4%), 72.2% (±4.8%) and 65.4% (±2.6%)
216 respectively. The large increase for all the samples suggested that T4 phages mainly lost infectivity
217 during the purificationduring purification but its capsid was kept intact as its genome can still be
218 detected [11].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2019 doi:10.20944/preprints201906.0125.v1



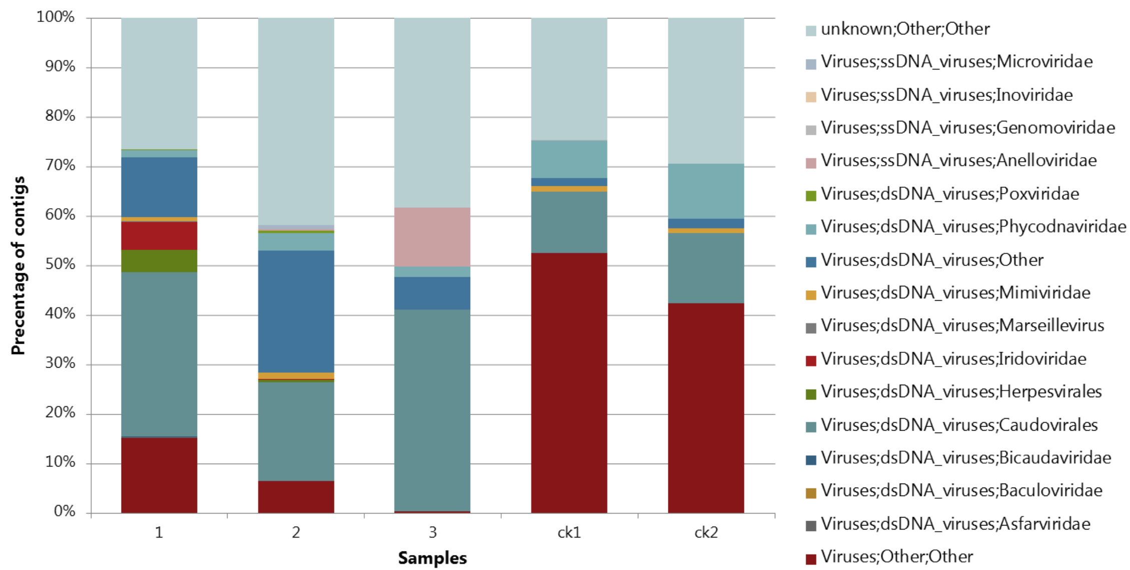
219

220 **Figure 2.** Infective phage recovery determined by plaque assays. The percentages of phages recovered
221 were determined by plaque assay at each different sampling point

222 *3.3. Assessing the protocol by sequencing and bioinfomatic analysis*

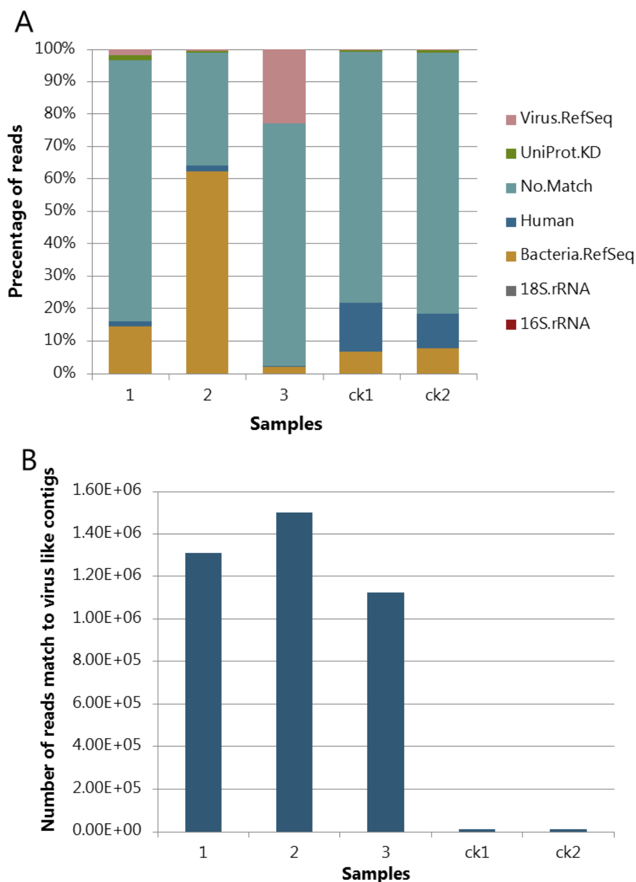
223 After the VLPs were concentrated from the fecal samples, the viral DNA was extracted and
224 amplified by MDA to include ssDNA viruses in library construction and sequencing. Only half hour
225 incubation was performed instead of 1.5 hours as described in the standard protocol for MDA to
226 limit the selective amplification of ssDNA which is known to increase with incubation time [27]. As
227 seen from Suppl. Figure 1 only a minor fraction of the metavirome sequences were derived from
228 human, fungi or bacterial genomes, indicating that the method is selective in separating viral
229 particles and larger particles such as bacteria. No 16S rRNA gene fragments were detected in 50000
230 reads in any of the samples underlining that the protocol is efficient in removing bacterial cells and
231 genomic fragments. However, as seen from figure 4A the fecal sample from infant 2 was found to
232 contain a rather high fraction of reads aligning to bacterial genomes, but a closer analysis of the
233 results showed that many of these reads matched to putative prophage sequences in Bacteroides
234 dorei. The B. dorei cell size has been reported to be 1.6 to 4.2 μm by 0.8 to 1.2 μm [30], meaning that it
235 should not pass through the 0.45 μm filter. Moreover, the samples from infant 1 and 3 showed very
236 few hits to bacterial genomes, reflecting that the there was no systematic bacterial contamination due
237 to the extraction protocol. Moreover the detected bacterial hits may reflect that the abundance of
238 induced prophages varied among different samples. Negative controls (SM buffer control, ck1 and
239 ck2 (figure 3 and figure 4) were also sequenced. As seen from figure 4B, the number of reads
240 matching to viral like sequences were less than 1 percentage of the true samples and with a
241 composition much different from the fecal samples (Figure 3) where Caudovirales was the dominant
242 order as found in most of the gut virome of infants [20,29].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2019 doi:10.20944/preprints201906.0125.v1



243

244 **Figure 3.** Taxonomic distribution (relative abundance) of the sequenced virome. The relative distribution is
245 described at the taxonomical level of orders. Taxonomy of contigs was determined by querying the viral contigs
246 against a database containing taxon signature genes for virus orthologous group hosted at www.vogdb.org.
247 The unknown;other;other category is the contigs that have no relation to any known classified sequences.



248

249 **Figure 4.** (**A**) Distribution of sequencing reads into the different taxonomic categories viral, human, bacterial,
250 and unknown origin. To check the presence of non-viral DNA sequences, 50000 random forward reads were
251 evaluated according to their match to a range of viral, bacterial and human reference genome and protein
252 databases as described [17]. No reads (in 50000 reads) were matched 16S rRNA gene sequences in all the
253 samples. (**B**) Counts of sequencing reads to the assembled virus like contigs. At least 10 times coverage/contig

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2019 doi:10.20944/preprints201906.0125.v1

254 was applied here as the threshold for counting. Numbers 1-3 refer to virome extracted from feces from infant
255 1-3. ck1 and ck2 refer to co-extracted blank (SM buffer) samples.

256 We have also applied the extraction method on fecal samples from adults (data not show) as
257 well as virome preparation for vaginal swabs. Good qualities of sequencing were in all cases
258 obtained (manuscripts in preparation), suggesting that the protocol described here should be widely
259 applicable.

260 In summary we here describe a protocol for extraction of still infective viromes from low
261 volume fecal samples suitable for metagenomic sequencing. The protocol has a relatively high
262 throughput allowing extraction of up to 48 viromes within one working day and with less than 4
263 hours of hands-on time.

264 **Author Contributions:** Conceptualization, L.D., J.L.C.M. and D.N.; Formal analysis, L.D., R.S., and J.L.C.M.;
265 Funding acquisition, H.B., S.M. and D.N.; Investigation, L.D., R.S., J.L.C.M. and W.K.; Methodology, L.D., R.S.,
266 J.L.C.M. and W.K.; Project administration, L.J., J.T, S.S, J.S., H.B., S.M. and D.N.; Supervision, L.D. and D.N.;
267 Writing – original draft, L.D.; Writing – review & editing, L.D., R.S., J.L.C.M., L.J., J.T., S.S., J.S., S.M., and D.N..

268 **Funding:** This work is supported by the Joint Programming Initiative 'Healthy Diet for a Healthy Life'. The
269 funding agencies supporting this work are: The Danish Agency for Science and Higher Education (5195-00002B
270 and 5195-00003B), Institut National de la Recherche Agronomique (INRA) and the Canadian Institutes of
271 Health Research (Team grant on Intestinal Microbiomics, Institute of Nutrition, Metabolism, and Diabetes).
272 S.M. holds the Tier 1 Canada Research Chair in Bacteriophages.

273 **Acknowledgments:** We would like to thank Marie-Agnès Petit from Micalis Institute, INRA, France and Finn
274 Vogensen from Department of Food Science, University of Copenhagen for valuable discussion.

275 **Conflicts of Interest:** The funders had no role in the design of the study; in the collection, analyses, or
276 interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

277 **References**

278 1. Andersen, L.O.; Vedel Nielsen, H.; Stensvold, C.R. Waiting for the human intestinal Eukaryotome. The
279 ISME journal 2013, 7, 1253-1255, doi:10.1038/ismej.2013.21.
280 2. Gaci, N.; Borrel, G.; Tottey, W.; O'Toole, P.W.; Brugere, J.F. Archaea and the human gut: new beginning of
281 an old story. World journal of gastroenterology 2014, 20, 16062-16078, doi:10.3748/wjg.v20.i43.16062.
282 3. Shkoporov, A.N.; Hill, C. Bacteriophages of the Human Gut: The "Known Unknown" of the Microbiome.
283 Cell host & microbe 2019, 25, 195-209, doi:10.1016/j.chom.2019.01.017.
284 4. Scarpellini, E.; Ianiro, G.; Attili, F.; Bassanelli, C.; De Santis, A.; Gasbarrini, A. The human gut microbiota
285 and virome: Potential therapeutic implications. Digestive and liver disease : official journal of the Italian
286 Society of Gastroenterology and the Italian Association for the Study of the Liver 2015, 47, 1007-1012,
287 doi:10.1016/j.dld.2015.07.008.
288 5. 5. Ogilvie, L.A.; Jones, B.V. The human gut virome: a multifaceted majority. Frontiers in microbiology
289 2015, 6, 918, doi:10.3389/fmicb.2015.00918.
290 6. 6. Reyes, A.; Semenkovich, N.P.; Whiteson, K.; Rohwer, F.; Gordon, J.I. Going viral: next-generation
291 sequencing applied to phage populations in the human gut. Nature reviews. Microbiology 2012, 10,
292 607-617, doi:10.1038/nrmicro2853.
293 7. 7. Fung, T.C.; Olson, C.A.; Hsiao, E.Y. Interactions between the microbiota, immune and nervous
294 systems in health and disease. Nature neuroscience 2017, 20, 145-155, doi:10.1038/nn.4476.
295 8. 8. Edwards, R.A.; Rohwer, F. Viral metagenomics. Nature reviews. Microbiology 2005, 3, 504-510,
296 doi:10.1038/nrmicro1163.
297 9. 9. Hurwitz, B.L.; U'Ren, J.M.; Youens-Clark, K. Computational prospecting the great viral unknown.
298 FEMS microbiology letters 2016, 363, doi:10.1093/femsle/fnw077.
299 10. 10. Paez-Espino, D.; Eloe-Fadrosh, E.A.; Pavlopoulos, G.A.; Thomas, A.D.; Huntemann, M.; Mikhailova,
300 N.; Rubin, E.; Ivanova, N.N.; Kyrpides, N.C. Uncovering Earth's virome. Nature 2016, 536, 425-430,
301 doi:10.1038/nature19094.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2019 doi:10.20944/preprints201906.0125.v1

302 11. 11. Castro-Mejia, J.L.; Muhammed, M.K.; Kot, W.; Neve, H.; Franz, C.M.; Hansen, L.H.; Vogensen, F.K.;
303 Nielsen, D.S. Optimizing protocols for extraction of bacteriophages prior to metagenomic analyses of
304 phage communities in the human gut. Microbiome 2015, 3, 64, doi:10.1186/s40168-015-0131-4.
305 12. 12. Reyes, A.; Haynes, M.; Hanson, N.; Angly, F.E.; Heath, A.C.; Rohwer, F.; Gordon, J.I. Viruses in the
306 faecal microbiota of monozygotic twins and their mothers. Nature 2010, 466, 334.
307 13. 13. Minot, S.; Bryson, A.; Chehoud, C.; Wu, G.D.; Lewis, J.D.; Bushman, F.D. Rapid evolution of the
308 human gut virome. Proceedings of the National Academy of Sciences 2013, 110, 12450-12455.
309 14. 14. Minot, S.; Sinha, R.; Chen, J.; Li, H.; Keilbaugh, S.A.; Wu, G.D.; Lewis, J.D.; Bushman, F.D. The human
310 gut virome: inter-individual variation and dynamic response to diet. Genome research 2011, 21, 1616-1625.
311 15. 15. Conceicao-Neto, N.; Zeller, M.; Lefrere, H.; De Bruyn, P.; Beller, L.; Deboutte, W.; Yinda, C.K.;
312 Lavigne, R.; Maes, P.; Van Ranst, M., et al. Modular approach to customise sample preparation procedures
313 for viral metagenomics: a reproducible protocol for virome analysis. Scientific reports 2015, 5, 16532,
314 doi:10.1038/srep16532.
315 16. 16. Muhammed, M.K.; Krych, L.; Nielsen, D.S.; Vogensen, F.K. A high-throughput qPCR system for
316 simultaneous quantitative detection of dairy Lactococcus lactis and Leuconostoc bacteriophages. PloS one
317 2017, 12, e0174223, doi:10.1371/journal.pone.0174223.
318 17. 17. Rasmussen, T.S.; de Vries, L.; Kot, W.; Hansen, L.H.; Castro-Mejía, J.L.; Vogensen, F.K.; Hansen, A.K.;
319 Nielsen, D.S. Mouse vendor influence on the bacterial and viral gut composition exceeds the effect of diet.
320 bioRxiv 2019, 10.1101/588160, 588160, doi:10.1101/588160.
321 18. 18. Wood, D.E.; Salzberg, S.L. Kraken: ultrafast metagenomic sequence classification using exact
322 alignments. Genome biology 2014, 15, R46.
323 19. 19. Nurk, S.; Meleshko, D.; Korobeynikov, A.; Pevzner, P.A. metaSPAdes: a new versatile metagenomic
324 assembler. Genome research 2017, 27, 824-834.
325 20. 20. Reyes, A.; Blanton, L.V.; Cao, S.; Zhao, G.; Manary, M.; Trehan, I.; Smith, M.I.; Wang, D.; Virgin,
326 H.W.; Rohwer, F. Gut DNA viromes of Malawian twins discordant for severe acute malnutrition.
327 Proceedings of the National Academy of Sciences 2015, 112, 11941-11946.
328 21. 21. Liao, Y.; Smyth, G.K.; Shi, W. The Subread aligner: fast, accurate and scalable read mapping by
329 seed-and-vote. Nucleic acids research 2013, 41, e108-e108.
330 22. 22. Thurber, R.V.; Haynes, M.; Breitbart, M.; Wegley, L.; Rohwer, F. Laboratory procedures to generate
331 viral metagenomes. Nature protocols 2009, 4, 470.
332 23. 23. Minot, S.; Sinha, R.; Chen, J.; Li, H.; Keilbaugh, S.A.; Wu, G.D.; Lewis, J.D.; Bushman, F.D. The human
333 gut virome: inter-individual variation and dynamic response to diet. Genome research 2011, 21, 1616-1625,
334 doi:10.1101/gr.122705.111.
335 24. 24. Vajda, B. Concentration and purification of viruses and bacteriophages with polyethylene glycol.
336 Folia microbiologica 1978, 23, 88-96.
337 25. 25. Shkoporov, A.N.; Ryan, F.J.; Draper, L.A.; Forde, A.; Stockdale, S.R.; Daly, K.M.; McDonnell, S.A.;
338 Nolan, J.A.; Sutton, T.D.S.; Dalmasso, M., et al. Reproducible protocols for metagenomic analysis of
339 human faecal phageomes. Microbiome 2018, 6, 68, doi:10.1186/s40168-018-0446-z.
340 26. 26. Conceicao-Neto, N.; Yinda, K.C.; Van Ranst, M.; Matthijnssens, J. NetoVIR: Modular Approach to
341 Customize Sample Preparation Procedures for Viral Metagenomics. Methods in molecular biology 2018,
342 1838, 85-95, doi:10.1007/978-1-4939-8682-8_7.
343 27. 27. Roux, S.; Solonenko, N.E.; Dang, V.T.; Poulos, B.T.; Schwenck, S.M.; Goldsmith, D.B.; Coleman, M.L.;
344 Breitbart, M.; Sullivan, M.B. Towards quantitative viromics for both double-stranded and single-stranded
345 DNA viruses. PeerJ 2016, 4, e2777.
346 28. 28. Kleiner, M.; Hooper, L.V.; Duerkop, B.A. Evaluation of methods to purify virus-like particles for
347 metagenomic sequencing of intestinal viromes. BMC genomics 2015, 16, 7.
348 29. 29. Lim, E.S.; Zhou, Y.; Zhao, G.; Bauer, I.K.; Droit, L.; Ndao, I.M.; Warner, B.B.; Tarr, P.I.; Wang, D.;
349 Holtz, L.R. Early life dynamics of the human gut virome and bacterial microbiome in infants. Nature
350 medicine 2015, 21, 1228-1234, doi:10.1038/nm.3950.
351 30. Bakir, M.A.; Sakamoto, M.; Kitahara, M.; Matsumoto, M.; Benno, Y. Bacteroides dorei sp. nov.,
352 isolated from human faeces. International journal of systematic and evolutionary microbiology 2006, 56,
353 1639-1643, doi:10.1099/ijs.0.64257-0.
354
355