Characterization of virus-like particles associated with the human faecal and caecal microbiota

Lesley Hoyles[a], Anne L. McCartney[b], Horst Neve[c], Glenn R. Gibson[b], Jeremy D. Sanderson[d], Knut J. Heller[c], Douwe van Sinderen[a,e]*

[a]School of Microbiology, University College Cork, Western Road, Cork, Ireland

[b]Food Microbial Sciences Unit, Department of Food and Nutritional Sciences, The University of Reading, Whiteknights, Reading RG6 6UR, UK

[c]Institute of Microbiology and Biotechnology, Max Rubner-Institut, Federal Research Institute of Nutrition and Food, Kiel, Germany

[d]Department of Gastroenterology, Guy's and St Thomas' NHS Foundation Trust and King's College London, London, UK

[e]Alimentary Pharmabiotic Centre, Biosciences Institute, National University of Ireland, Western Road, Cork, Ireland

Lesley Hoyles lesley.hoyles@reading.ac.uk

Anne L. McCartney a.l.mccartney@reading.ac.uk

Horst Neve horst.neve@mri.bund.de

Glenn R. Gibson g.r.gibson@reading.ac.uk

Jeremy D. Sanderson jeremy.sanderson@kcl.ac.uk

Knut J. Heller knut.heller@t-online.de

Douwe van Sinderen d.vansinderen@ucc.ie ***Correspondence and reprints**

ABSTRACT

This work represents an investigation into the presence, abundance and diversity of virus-like particles (VLPs) associated with human faecal and caecal samples. Various methodologies for the recovery of VLPs from faeces were tested and optimized, including successful down-stream processing of such samples for the purpose of an in-depth electron microscopic analysis, pulsed-field gel electrophoresis and efficient DNA recovery. The applicability of the developed VLP characterization method beyond the use of faecal samples was then verified using samples obtained from human caecal fluid.

Keywords: bacteriophages; microbial ecology; gastrointestinal tract; transmission electron microscopy; caecum; faeces.

1. INTRODUCTION

Viruses are the most numerous biological entities within the biosphere of our planet (being present at an estimated number of ~10^{31}), with bacteriophages representing the most abundant group of environmental viruses [1, 2]. Bacteriophages are ubiquitous viruses that infect bacterial cells and disrupt their metabolism. Multiple bacteriophage types can infect a specific microbial isolate, with most bacteriophages infecting only certain species or even strains of bacteria [3].

Although they are abundant and potentially important to microbial populations indigenous to different ecological niches within the human gastrointestinal tract and to host health, little attention has been paid to bacterial virus-like particle (VLP) assemblages and their interactions with the gastrointestinal microbiota and/or human host until recently. Various publications have highlighted the potential importance of bacteriophages in inflammation states, including Crohn's disease [4, 5, 6], and as therapeutic agents [7, 8]. Breitbart *et al.* [9] conducted the first metagenomic study on dsDNA-containing VLPs associated with the human faecal virome using a fresh sample from a healthy adult male. On the basis of previous estimates of gut microbial diversity, Breitbart *et al.* [9] predicted that there are two to five times more viral genotypes (~1,200 viral genotypes) present in the human gastrointestinal microbiota than the number of bacterial species, with the vast majority of these VLPs representing bacteriophages and prophages. RNA viruses present in human faeces have been found to be mostly associated with plant viruses, such as Pepper Mild Mottle Virus, with RNA bacteriophages making a minimal contribution to the diversity of the intestinal virome [10]. Using frozen samples from four adult female monozygotic twins and their mothers, Reyes *et al.* [11] demonstrated that each individual harbours a unique virome (ssDNA and dsDNA) regardless of their genetic relatedness to another individual, and that intrapersonal virome diversity is very low, with >95 % of virotypes retained over a one-year

period (between 71 and 2,773 viral genotypes identified among the samples). More recently, Minot *et al.* [12] confirmed inter-individual variation in the virome and, similar to modulation of the faecal bacteriome by foodstuffs, demonstrated that diet influences host faecal virome structure. Minot *et al.* [13] also confirmed the relative stability of an individual's virome, monitoring 24 faecal samples from one individual over a 2.5-year period.

The murine virome has been used to demonstrate enrichment of bacteriophage-encoded, antibiotic-resistance genes (related and unrelated to the drug used) after antibiotic therapy, with the adaptive capacity of the virome (specifically its bacteriophage component) suggested to protect gut bacteria, thereby preserving the microbiota's robustness during antibiotic stress [14].

It is clear from the aforementioned studies that the virome–bacteriome community in the gut is governed by complex and dynamic interactions in health, and that its balance may be disturbed when under stress, e.g. during antibiotic intervention [14]. Although metagenomic studies have greatly improved our understanding of the virome associated with the human and murine gut microbiomes, it has been notable how little bacteriophage material, and consequently DNA, has been isolated from samples when this information was included in a publication. Reyes *et al.* [11] reported the isolation of ~500 ng DNA from 2 to 5 g of frozen faeces, whereas Thurber *et al.* [15] stated that between 500 and 3000 ng of DNA could be isolated from 500 g of human faeces, though the publication they cite [9] provides no information regarding the amount of DNA isolated from the 500 g faecal sample examined in the original study. In addition, to the best of our knowledge, no attempts have been made to enumerate bacteriophages in faecal filtrates prior to CsCl purification. The aims of this study were to develop reliable and effective methods for the recovery and

characterization of VLPs in human samples, and to apply these methods to human faecal and caecal samples to demonstrate the methods' efficiency.

2. MATERIALS AND METHODS

2.1 Processing of and isolation of VLPs from faecal and caecal samples

Faecal samples were obtained from six healthy adult (2 male, 4 female) donors of between 23 and 52 years of age, all of whom were members of the Department of Food and Nutritional Sciences, University of Reading. None of the donors had taken antibiotics, prebiotics and/or probiotics in the 3 months prior to sampling, and none had any history of gastrointestinal disorder. All donors gave informed oral consent for their faecal samples to be used for microbiological analyses. Ethical approval for the collection of caecal effluent from patients was obtained from St Thomas' Hospital Research Ethics Committee (06/Q0702/74) covering Guy's and St Thomas' Hospitals and transferred by agreement to London Bridge Hospital. Where available, clinical information for the colonoscopy patients is given in the text. Samples were collected at Reading (faeces) or transported from St Thomas' Hospital (caecum), and maintained under anaerobic conditions (faeces, MACS1000 anaerobic workstation, Don Whitley Scientific, gas composition 80 % N_2, 10 % H_2, 10 % CO_2; caecum, on ice in a gas jar with an anaerobic gas-generating pack; Oxoid Ltd) for a maximum of 2 h before processing. Caecal samples were collected during routine colonoscopy following preparation of the bowel with sodium picosulphate and a reduced fibre diet for 3 days. Liquid residue in the caecum was aspirated via the colonoscope suction channel into a standard 30 ml trap specimen container and immediately transferred to a gas jar.

Faecal homogenates (20 %, w/v, from 25 g of faecal material) were prepared in sterile TBT [0.1 μm filtered prior to autoclaving; 100 mM Tris/HCl, pH 8.0; 100 mM NaCl; 10 mM $MgCl_2.6H_2O$] or sterile 0.5 % 'Lab-Lemco'/6 % NaCl (LL [16]). Caecal homogenates (20 %, v/v, from 10 to 30 ml of caecal effluent) were prepared in LL. Faecal and caecal samples were placed into a filter stomacher bag and homogenized in a Stomacher 400 Lab System (Seward) for 120 s at low speed. The bag was removed from the stomacher and massaged

manually to further disperse any large particles remaining in the sample; the sample was then stomached for a further 120 s. The homogenate was kept on ice for 2 h to allow desorption of VLPs from solid material, then centrifuged at 11,180 ***g*** for 30 min at 10ºC. The supernatant was transferred to a fresh tube and centrifuged again at 11,180 ***g*** for 30 min at 10ºC. Supernatant [herein referred to as faecal filtrate (FF) for both faecal and caecal samples] was passed through sterile 0.45 μm cellulose acetate filters (Millipore) and the FF was then collected in a sterile container. An aliquot (10 μl) of each of the FFs was examined by epifluorescence microscopy (EFM) after viral particles had been stained with SYBR Gold (see below) to confirm that the samples were free of bacteria.

FF prepared in LL was used for enumeration of VLPs in faeces via EFM (see below) and 1 ml aliquots of FF prepared in TBT were for examination by transmission electron microscopy (TEM; see below).

Aliquots of FF and LL (two sets of three aliquots of 100 μl each: one for aerobic cultivation, one for anaerobic cultivation) were spread onto Columbia blood agar containing 5 % laked horse blood (Oxoid), and incubated aerobically and anaerobically. Sterility of the filtrates was confirmed by the absence of microbial growth on plates following incubation for 2 (aerobic) and 5 (anaerobic) days. Sterile brain–heart infusion broth was inoculated with 100 μl of FF and LL, and incubated aerobically for 2 days: aliquots (100 μl) were spread in triplicate on plates to confirm sterility of the broth cultures. In addition, 10-ml aliquots of FF were stored at 4°C for 6 months after collection, and were found to be free of bacteria when checked (by plating) at monthly intervals.

Poly(ethylene glycol) (PEG; BioUltra, 8000; Sigma) was added to 2 × 20 ml aliquots of the LL-prepared 0.45 μm-filtered FFs (final concentration of PEG, 10 %, w/v). The samples were shaken gently to dissolve the PEG, then left at 4°C for 16 h ([16] used 16 h as they observed an improved recovery of RNA bacteriophage compared to shorter incubations;

[17] used 10 h, as they determined this to be longer than the minimum time required to obtain a constant and stable white layer of viral precipitate). PEG-precipitated VLPs were then harvested by centrifugation at 4,500 ***g*** for 30 min in a swing-out-bucket rotor at 4°C. The supernatant was removed and the pellet resuspended in 1 ml TBT for pulsed-field gel electrophoresis (PFGE) or 5 ml TBT for purification of VLPs on a CsCl gradient [15] with ultracentrifugation performed at 100,000 ***g*** for 2 h at 10°C in a fixed-angle Type 50 Ti rotor (Beckman Coulter). VLPs recovered by CsCl (1.35 and 1.5 g/cm^3 fractions) were dialysed (12,000 Da cut-off) twice against 400 ml sterile TBT, passed through a sterile 0.45 μm cellulose acetate filter and stored at 4°C until DNA was extracted. An aliquot (5 μl) of each of the CsCl-purified samples was viewed using EFM after viral particles had been stained with SYBR Gold (see below) to confirm that the samples were free of bacteria.

2.2 Epifluorescence microscopy (EFM)

A combination of the methods of Thurber and Patel [15, 18] was used to prepare FFs (faecal only) for EFM. VLPs present in FF were not fixed in paraformaldehyde prior to enumeration, as Wen *et al.* [19] showed that aldehyde fixation leads to a rapid loss in viral abundance. The SYBR Gold concentration of 1–5× recommended by [15] did not work with FFs; therefore, the 400× concentration recommended by [18] for use with planktonic aquatic samples was used in this study.

The filtration system (see [15] for specifics of the set-up employed), including a glass-graduated column, for collecting VLPs on filters was cleaned using 5 ml of 0.1 μm-filtered, sterile H_2O and 5 ml of 0.1 μm-filtered ethanol. A sterile pair of flat-tipped forceps was used to remove a 0.02-μm white Anodisc 25 membrane (Whatman) from its box, and the filter was fitted to the glass frit of the filtration system under a low vacuum [<10 psi (~62 kPa)]. Duplicate 2 ml aliquots of 0.1 μm filtered, sterile H_2O and 10 μl of sterile LL in 2 ml

166 0.1 μm filtered, sterile H_2O were used as negative controls (to assess whether there was any
167 contamination in the water, the LL, the filter tower or the filter surface) and passed through
168 filters under low vacuum before any of the FF aliquots were filtered in duplicate. For each
169 sample, a 10 μl aliquot of 0.45 μm-filtered FF was added to 2 ml of the 0.1 μm filtered,
170 sterile H_2O and the suspension gently mixed. The diluted sample was then introduced into the
171 filter system with the low vacuum maintained. The liquid was passed through the filter, and
172 VLPs were collected on the filter. The clip and glass column were carefully removed from
173 the filtration system and the filter was gently removed from the glass frit (still under
174 vacuum). While holding the filter with a forceps, the back of the filter was gently blotted dry
175 on a clean Kimwipe; the filter was then left to dry on a Kimwipe in a sterile Petri dish for 10
176 min in a dark box. The filtration system was cleaned with 5 ml of 0.1 μm-filtered, sterile H_2O
177 and 5 ml of 0.1 μm-filtered ethanol, and the next sample processed.

178 Filters were stained with 100 μl droplets of 400× SYBR Gold and dried as described
179 by [18], prior to being applied in pairs to sterile glass slides. The mountant (20 μl per
180 coverslip) ProLong® Gold antifade reagent (Invitrogen) was added to 25 mm glass coverslips.
181 Coverslips were picked up with sterile forceps, inverted and placed on the Anodisc filters on
182 the microscope slides. Slides were then left at room temperature in the dark for 16 h to allow
183 the antifade to cure. Slides prepared in this manner can be stored at room temperature for 1
184 week or at 4ºC for 2 months.

185 Slides were viewed at 1000× magnification under a Nikon Microphot-SA microscope
186 fitted with a B-2A (blue excitation) filter and attached to a CoolSNAP-Pro MONOCHROME
187 (Media Cybernetics Inc.) camera. Images were captured using Image-Pro PLUS version
188 4.5.0.19 (Media Cybernetics Inc.), and VLPs in 25 fields of view were enumerated by eye.
189 The number of VLPs per millilitre of FF was calculated using the following equation: mean
190 number of VLPs in 25 fields of view × 100 × 25760.205; where 100 represents taking the

sample back to per millilitre FF and 25760.205 represents the number of fields of view on a filter. To give an estimate of the number of VLPs per gram of faeces, results were multiplied by a factor of 5. The detection limit of the method was 1.03×10^5 VLPs/ml FF or 5.15×10^5 VLPs/g faeces.

2.3 TEM of FFs (faecal and caecal)

Aliquots (1 ml) of FF prepared in TBT (faeces) or LL (caecum) were used for TEM analysis within 1 week of collection: these filtrates had not been concentrated using PEG. Carbon films (~3×3 mm in size) were floated from mica-sheets into a drop of filtrate (100 µl). After an adsorption time of 5–10 min, samples were transferred into a drop of 1 % (v/v) of EM-grade glutaraldehyde (20 min) and subsequently into a drop of 2 % (w/v) uranyl acetate for negative staining (1–2 min). After two washes for a few seconds in drops of distilled water, samples were picked up with 400-mesh copper grids (Plano, Wetzlar, D). Electron micrographs were taken in a Tecnai 10 transmission electron microscope (FEI Company, Eindhoven, the Netherlands) at an accelerating voltage of 80 kV. Digital micrographs were taken with a Megaview G2 CCD camera (Olympus SIS, Münster, Germany). For estimation of VLP titers on the EM grids, a freshly prepared *Lactococcus lactis* bacteriophage preparation with defined titer and unique morphology (i.e. prolate-headed bacteriophage P001 [20]) was added in concentrations of 10^7 and 10^8 plaque-forming units (pfu) per ml to the faecal sample of donor 2 (with highest bacteriophage titer according to TEM analysis).

2.4 PFGE

Aliquots (40 µl) of PEG-precipitated samples resuspended in 1 ml TBT were used for PFGE. These were heated at 60 or 75°C as appropriate. An aliquot (25 µl) of each heated

(viscous) suspension and loading dye (5 μl) were combined and loaded into wells, which were sealed with molten agarose. Run conditions for PFGE were as described by Fuhrman *et al.* [21]: a 1 % agarose gel (SeaKem LE agarose) was made in 0.5× TBE [250 ml of 5× TBE buffer (27 g Sigma 7–9, 13.75 g boric acid, 5 ml 0.5 M EDTA, 500 ml of H_2O) added to 2.25 l H_2O] and a CHEF DR II apparatus (Bio-Rad) was run for 18 h at 6 V and 14°C in 0.5× TBE, with a 1–10 s switch time. Gels were stained with ethidium bromide (5 μg/ml) for 20 min and destained in distilled H_2O for 10 min, or with 1× SYBR Gold (Molecular Probes) made in 0.5× TBE. Bands of DNA were visualized under UV light (ethidium bromide) or using a Dark Reader DR89X Transilluminator (Integrated Scientific Solutions Inc.) (SYBR Gold).

2.5 Isolation of DNA from CsCl-purified VLPs

DNA was extracted from 500 μl portions of CsCl-purified VLP samples according to [15]. Prior to extracting DNA from purified VLPs, samples were treated with 20 μl of 1 mg/mL DNAse I (from bovine pancreas, 552 Kunitz/mg protein; made in water passed through a 0.1 μm filter prior to autoclaving) for 1 h at 37°C, and were then heated at 80°C for 10 min to inactivate the DNAse prior to DNA extraction. DNA was air-dried and resuspended in 50 μl of TE buffer. Sterility of samples (i.e. absence of bacterial DNA) was confirmed by negative PCR from samples (10 μl) with universal primers [22]. DNA from an in-house strain of *Bifidobacterium longum* (4-FAA1; [23]) was used as a positive control; sterile water was used as the negative control.

3. RESULTS AND DISCUSSION

3.1 Examination of VLPs in FFs (faeces) by using EFM

Previous studies examining the faecal virome have used 0.22 μm filters to prepare samples [9, 11, 12, 13]. Preliminary work performed with 0.45 μm-filtered faecal samples and 0.45 μm-filtered samples originating from *in vitro* fermentation systems inoculated with faeces stained with 4′,6-diamidino-2-phenylindole (DAPI) and viewed under an epifluorescence microscope suggested that the abundance of dsDNA VLPs in intestinal samples was high, and that these particles can be enumerated (unpublished observations). Contamination by bacteria was not observed in any of these samples, and confirmed by an absence of cultivable bacteria and no bacteria in TEM analyses, so we decided to continue using larger-pore filters for processing samples. Klieve & Swain [24] had previously used 0.45 μm-filtered samples to characterize VLPs present in rumen contents, and reported no problems with contamination by bacteria.

To examine the VLPs present in FFs (faeces) by EFM, the method described by [9] and [15] was used initially. However, it was found that we were unable to visualize many, if any, VLPs present in the samples using 1–5× SYBR Gold as the DNA/RNA stain. Increasing the concentration of SYBR Gold to 400×, as used by [18] for planktonic aquatic samples, allowed us to enumerate and detect VLPs present in the FFs prepared from faecal samples of six donors (Fig. 1 and Fig. 2). The samples were extracted in LL, though the method worked as well with samples extracted in TBT (data not shown). To determine the relative abundance of the predominant VLP morphologies in FFs, samples were not concentrated by CsCl centrifugation (Fig. 1). All FFs were found to be free of bacteria by EFM (Fig. 2), with the number of VLPs present in samples ranging from 2.4×10^8 to 1.12×10^9 VLPs/ml FF (mean 5.58×10^8 VLPs/ml FF) (equivalent to ~1.2×10^9 to 5.58×10^9 VLPs/g faeces, mean 2.94×10^9). Lepage *et al.* [4] enumerated VLPs in gut mucosal samples from 14 healthy individuals and

19 Crohn's disease patients and found on average 1.2×10^9 VLPs/biopsy (range 4.4×10^7–1.7×10^{10}), in agreement with the mean value we present here for faecal VLPs. Of note, Crohn's disease patients harboured significantly ($P = 0.024$) more VLPs than healthy individuals (2.9×10^9 vs 1.2×10^8 VLPs/biopsy) in the study of Lepage *et al.* [4].

It is generally accepted that there are around 10 bacteriophages for every microbial cell in environmental samples investigated to date [25]. Extrapolating this figure to the gut microbiota, from fluorescence *in situ* hybridization studies it is estimated that the faecal microbiota harbours ~10^{11} bacteria/g faeces in healthy adults [26]. Consequently, one would expect the presence of at least 10^{12} VLPs/g faeces. Enumerating VLPs in faeces via FFs (or any liquid medium) by EFM is highly subjective as a dot of very intense fluorescence may in fact represent a cluster of VLPs (Fig. 2), a phenomenon frequently encountered during this study. Patel *et al.* [18] stated that, to accurately enumerate VLPs by EFM, micro-adjustments using the fine focus of the microscope have to be made to ensure that all viruses in a particular grid-reticle box are counted. We agree with this statement, and furthermore add that many VLPs have very likely been lost during the preparation and filtering of samples, either by association with debris in the initial centrifugations or by being caught in the filters because of clogging or because the VLPs are too big to pass through the pores (e.g. members of the order *Megavirales* [27]). Even after centrifugation, the supernatants from several of the samples, while appearing relatively translucent, were highly viscous (perhaps due to host mucins) and clogged the 0.45 μm filters with less than 2 ml of sample being filtered. Consequently, the values we provide for the numbers of VLPs in FFs and faecal samples are a conservative estimate. We believe the true number of VLPs present in faeces to be higher, possibly between 10^{10} and 10^{12} VLPs/g faeces. Lepage *et al.* [4] determined there to be 10^{10} VLPs/mm^3 tissue in their study of mucosal VLPs.

3.2 Examination of VLP diversity in FFs by the use of TEM

VLPs were readily detected in the TBT-extracted faecal samples from all donors by TEM. Bacterial cells were never observed in any samples examined by TEM (limit of detection 10^6 per ml). It was striking how visibly different/distinct the VLP assemblages were between the donors, with no two donors sharing the same VLPs, at least on the basis of morphological appearance (Fig. 3 and Fig. 4; Supplementary Fig. 1). The vast majority of VLPs present in the samples appear to represent bacteriophages. Donor 1's VLP assemblage was predominated by small and large isometric-headed *Siphoviridae* with various tail length sizes (approx. 120 nm, 350 nm, 650 nm, or, in one extreme case, 1220 nm), with some small and large isometric-headed *Myoviridae* also present (Fig. 3a). Notably, donor 2's VLP assemblage was more diverse, predominated by numerous different morphotypes of *Myoviridae* with *Siphoviridae* also present (Fig. 3b). Interestingly, two detached *Myoviridae* tails of extreme length (480 nm) and thickness of the sheaths (40 nm) indicated the presence of giant *Myoviridae* phages (Fig. 3b). Sime-Ngando *et al*. [28] have reported isolation of bacteriophages with tails of 400 nm in length (heads 50–130 nm) from hypersaline lake samples, though images of these large bacteriophages are unavailable for direct comparison of the tails' structures with those found in this study. Donor 2's faecal sample contained distinct small isometric-headed *Myoviridae* phages with uncommon radial fibers (approx 150 nm in length) attached to the capsids (~70 nm diameter) (Fig. 4). These fibers are clearly extending the 110-nm tails of these phages, suggesting a primary role in bacteriophage adsorption. This unique bacteriophage morphotype was also documented in low numbers in Donor 4's faecal sample (not shown), and in caecal samples L16 (not shown) and L18 (Supplementary Fig. 2d). Fig. 4 also illustrates the apparent clustering of *Myoviridae* phages in samples. VLPs of various morphotypes derived from Donors 3, 4, 5, and 6 are shown in Supplementary Figure 1a–d. "Zeppelin"-like VLPs of constant thickness (56 nm) but

different lengths (370–630 nm) were detected in Donor 5 (Supplementary Fig. 1c) and Donor 6's samples (Supplementary Fig. 1d). Donor 6's sample was predominated by these "zeppelin"-like VLPs, with only a few bacteriophages detected. Sime-Ngando et al. [28] reported the presence of rods of up to 22–24 ×1000 nm: similar to the 'zeppelins' identified in the present study, these rods had no visible internal or external structures. The largest *Myoviridae* phages found in this study were documented for the sample of Donor 5 (Supplementary Fig. 1c), with the phages having prolate heads of 150 × 115 nm and tails of 400 nm in length. The sample from Donor 6 did not reveal the same level of diversity as seen for the other faecal samples and the reason for this is not known (Supplementary Fig. 1d). Enumeration of VLPs in the sample of Donor 2 (highest number of VLPs as detected by TEM) demonstrated there to be approx. 5×10^7 VLPs/ml FF.

For the majority of the caecal samples, a lower extent of viral biodiversity was documented. Sample L10 (no clinical data) was unique, containing exclusively small isometric-headed *Podoviridae* (50 nm diameter) and ~20-nm long appendages (Fig. 5b). Samples L08 (healthy) (Fig. 5a), and L02 and L03 (both IBD) (Suppl. Fig. 5b, c) appeared to exclusively contain *Myoviridae* phages. The greatest extent of morphological variation and the highest bacteriophage numbers were seen in caecal samples L16 (no clinical data) and L17 (no clinical data) (Fig. 5c, d) and – to a lesser extent – in samples L01 (diverticulosis) and L18 (no clinical data) (Supplementary Fig. 2a, d). A new type of a giant *Myoviridae* bacteriophage was exclusively found in sample L16 (Fig. 5c) with isometric heads of 125 nm in diameter, with tails of 340 nm in length and unique "curled" tail fibers of ~70-nm in length.

Lepage *et al.* [4] demonstrated that gut mucosal samples were predominated by morphotypes consistent with *Siphoviridae*, *Myoviridae* and *Podoviridae,* and that each

individual appeared to be colonized by one dominant bacteriophage family. Our TEM results for faecal and caecal VLP assemblages are consistent with these results.

3.3 PFGE of VLPs present in FFs

It was clear from the EFM and TEM analyses that, based on the number of VLPs present in our samples, it should be possible to isolate large amounts of VLP-derived DNA from faeces. Attempts to recover abundant VLP DNA from samples using the centrifugation method of Thurber *et al.* [15], and which was used by [9] and [11] to isolate VLP DNA from faecal samples, proved disappointing in terms of the recovered DNA yield. However, increasing the centrifugation speed to 100,000 ***g***, routinely used in the laboratory to purify lactococcal bacteriophages for preparation of high-quality DNA and used by Kulikov *et al.* [29] to recover bacteriophages from horse faeces, markedly improved recovery of faecal VLPs.

To concentrate VLPs from 20-ml quantities of FF so that they could be applied to CsCl gradients in 5 ml aliquots, it was decided to use PEG precipitation. This method of recovery has previously been used with, for example, marine samples and faecal samples to improve detection of F-specific coliphages in faecal material [16, 17], and allows large starting volumes of sample to be used for recovering VLPs from human faeces. The method can be scaled easily so that the VLPs from larger volumes of FF are precipitated for collection by centrifugation: we used 20 ml of FF from each donor here to demonstrate the efficacy of the method, but the entire FF derived for each donor's 25 g of starting material could have been used for PEG precipitation of VLPs.

PEG-precipitated samples resuspended in 1 ml TBT were examined using PFGE. Aliquots (40 μl) of the resuspended pellets were heated at 60°C [9] or 75°C ([24] used this temperature to inactivate nucleases) before PFGE. Profiles containing one or more bands

were observed for samples from all donors (Fig. 6). Following treatment at 60ºC, Donor 1's sample produced a faint band at 105 kb, with this band being more prominent following sample treatment at 75ºC; at the higher temperature, bands were also observed at 135 kb and just below 48.5 kb, with the most prominent band at 79 kb. Donor 2's 60ºC-treated sample produced a profile with a faint smear and a thin band at 84 kb; with the 75ºC-treated sample, the smear was more pronounced, with a thick band observed between 75 and 84 kb. Similar to Donor 2, the smear in Donor 4's sample became more pronounced following treatment at the higher temperature. For Donor 3's 60ºC-treated sample, two strong bands (at 66 and 97 kb) were observed; in the 75ºC-treated samples, bands were observed at 57 kb, 75 kb and 97 kb. The profile of Donor 5 was most affected by changing the temperature at which samples were treated prior to loading on the gel: the 60ºC-treated sample had three bands visible (most prominent at 100 kb, fainter bands at 48.5 kb and 66 kb), whereas the 75ºC-treated sample had these three bands with at least another 8 bands visible between 48.5 and 97.0 kb. Donor 6's sample had a faint band just below 48.5 kb after heating at 60ºC, whereas the sample heated at the higher temperature produced two prominent bands that were smaller than 48.5 kb. The results from the gel were reproducible (triplicate gels run; data not shown). It is unsurprising that heating at 75ºC produced more complex banding patterns, as we would expect the higher temperature to disassemble (some of) the capsids of VLPs resistant to heating at 60ºC, thereby releasing packaged DNA.

Using this method, it was possible to visualize VLP DNA when stained with SYBR Gold (Fig. 6) or ethidium bromide (Supplementary Fig. 3). It was clear that increasing the temperature at which the samples were heated prior to loading onto the gel revealed greater diversity in the samples, and that each individual harboured a unique VLP assemblage (Fig. 4). This is in agreement with the findings of the metagenomic studies of [11] and [12]. The sample from Donor 6 did not reveal the same level of diversity as seen for the other donors,

in agreement with the results shown by TEM (Supplementary Fig. 1d). This donor's sample was predominated by "zeppelin"-like VLPs that may not have released their nucleid acids by the conditions employed here.

Heating PEG-precipitated samples allowed us to generate PFGE images that demonstrated that each individual harbours a unique VLP assemblage, and encouraged us to continue with studies to improve recovery of VLP DNA from human faeces. It also showed that our method of recovery of VLPs present in faeces was superior to that used previously by [9] to generate a viral assemblage fingerprint. In that study, the authors stated that the limited amount of DNA recovered using tangential flow filtration with a 500 g faecal sample from a 33-year-old made it necessary to enhance the sample bands on the gel relative to the ladder. In contrast, no enhancement of the bands detected by SYBR Gold (Fig. 6) or ethidium bromide (Supplementary Fig. 3) was required using the methodology described herein, although SYBR Gold was, as expected, far more sensitive than ethidium bromide. PEG precipitation of VLPs present in 20 ml of FF, as used in this study, equates to the extraction of VLPs from ~4 g faeces (sample losses are seen during filtration, and vary from donor to donor due to the differences in viscosity of faecal supernatant obtained after removal of most bacteria and debris from samples after centrifugation), with 50 μl of the resuspended PEG precipitate representing VLPs isolated from ~200 mg of faeces.

The current study used a crude method of extracting DNA for PFGE. Using the method of Rohozinski *et al.* [30], in which VLPs are embedded in agarose blocks, with CsCl-purified VLPs prior to heating may allow better 'fingerprinting' of VLPs in faecal samples. This could, for example, be used as an inexpensive means of determining the effect of freeze–thawing faecal samples prior to the recovery of VLPs from faecal samples. All of the metagenomic studies conducted to date on the human faecal virome have used samples that have been frozen prior to recovery of VLPs from samples.

412

413 **3.4 Isolation of high-quality DNA from FFs**

414 Reyes *et al.* [11] and Thurber *et al.* [15] (based on the work of [9]) have reported
415 recovery of ~500 ng (from 2–5 g of frozen faeces) and 500–3000 ng (from 500 g of fresh
416 faeces), respectively. The method of Reyes *et al.* [11] is closest to that presented herein for
417 the recovery of VLPs from faeces [i.e. they hand-filtered samples, whereas Breitbart *et al.* [9]
418 used tangential flow filtration to process their sample]. As stated above, 20 ml of FF equates
419 to ~4 g faeces. We processed 500 μl portions of 3.5–4 ml CsCl-purified VLPs for DNA
420 extraction, and resuspended the DNA in 50 μl of TE (Fig. 7). If we had processed the entire
421 PEG-precipitated, CsCl-purified VLP sample for each donor, we would have recovered
422 between ~1000 and 1800 ng from ~4 g of faeces (theoretically ~6000–11000 ng from 25 g of
423 faeces). This increase in recovery of DNA is thought to be mainly due to the use of 0.45 μm
424 filters rather than 0.22 μm filters to process samples, supported by our observation that
425 passing FFs through 0.22 μm filters led to VLP counts by EFM that were approximately half
426 of those of the 0.45 μm-filtered samples. Increasing filter size and using PEG precipitation to
427 increase the volume of sample that can be processed would allow greater sampling of the
428 faecal virome in future metagenomics studies.

429 In addition to demonstrating the utility of the method with faecal samples, we have
430 successfully isolated VLP DNA from human caecal effluent (Fig. 7). For each of the samples,
431 20 ml of 0.45 μm-filtered caecal filtrate prepared in LL was PEG-precipitated, with the VLPs
432 then purified on CsCl gradients. These 20-ml aliquots equated to ~4 ml caecal effluent from
433 each of the three samples examined. We started with 30, 30 and 10 ml, respectively, of
434 effluent from caecal samples L07 (healthy), L08 (no clinical data) and L10 (no clinical data).
435 Therefore, if the filtrate from entire homogenates of these samples had been processed, we
436 would have expected to isolate between ~600 and 8000 ng VLP DNA from caecal effluent.

Consequently, it should be possible to conduct metagenomics studies of the VLP assemblages associated with the human caecum using the methods described herein.

3.5 Conclusions

We have presented a series of methods for enumerating and characterizing VLPs present in human faecal and caecal samples. On the basis of enumeration of VLPs using EFM, there are ~10^8 VLPs/ml FF, with faeces thought contain up to 10^{12} VLPs per gram of sample. TEM analysis of faecal VLPs indicates that an individual is colonized by one dominant bacteriophage family, with *Myoviridae* and *Siphoviridae* representing the main families of bacteriophage detected in faeces. PFGE analysis of PEG-concentrated samples has demonstrated that each individual harbours a unique VLP population, and that the temperature at which samples are treated greatly affects diversity observed. Using 0.45 μm filters to prepare samples, it is possible to isolate twice as much DNA as reported previously from ~4 g of faeces. The inclusion of PEG precipitation in the methodology means that VLPs from large volumes of FF can be concentrated, allowing recovery of microgram quantities of VLP DNA from faecal samples. In addition, PEG precipitation and CsCl purification can be used to recover VLPs from human caecal samples, with nanogram quantities of VLP DNA being recovered from the processed samples.

458 **4. ACKNOWLEDGEMENTS**

459 Dr Bernard M. Mackey (University of Reading) is thanked for the use of
460 epifluorescence microscopy facilities, as is Angela Back (MRI Kiel) for her expert technical
461 support in electron microscopy.

462

463

5. REFERENCES

[1] Breitbart M, Rohwer F. Here a virus, there a virus, everywhere the same virus? Trends Microbiol 2005;13:278-84.

[2] Edwards RA, Rohwer F. Viral metagenomics. Nat Rev Microbiol 2005;3:504-10.

[3] Rohwer F. Global phage diversity. Cell 2003;113:141.

[4] Lepage P, Colombet J, Marteau P, Sime-Ngando T, Doré J, LeClerc M. Dysbiosis in inflammatory bowel disease: a role for bacteriophages? Gut 2008;57:424-25.

[5] Vishwakarma V, Periaswamy B, Bhusan Pati N, Slack E, Hardt WD, Suar M. A novel phage element of *Salmonella* enterica serovar Enteritidis P125109 contributes to accelerated type III secretion system 2-dependent early inflammation kinetics in a mouse colitis model. Infect Immun 2012;80:3236-246.

[6] Wagner J, Maksimovic J, Farries G, Sim WH, Bishop RF, Cameron DJ, et al. Bacteriophages in gut samples from pediatric Crohn's disease patients: metagenomic analysis using 454 pyrosequencing. Inflamm Bowel Dis 2013;19:1598-608.

[7] Watanabe R, Matsumoto T, Sano G, Ishii Y, Tateda K, Sumiyama Y., et al. Efficacy of bacteriophage therapy against gut-derived sepsis caused by *Pseudomonas aeruginosa* in mice. Antimicrob Agents Chemother 2006;51:446-52.

[8] Meader E, Mayer MJ, Steverding D, Carding SR, Narbad A. Evaluation of bacteriophage therapy to control *Clostridium difficile* and toxin production in an *in vitro* human colon model system. Anaerobe 2013;22:25-30.

[9] Breitbart M, Hewson I, Felts B, Mahaffy JM, Nulton J, Salamon P, et al. Metagenomic analyses of an uncultured viral community from human feces. J Bacteriol 2003;185:6220-223.

[10] Zhang T, Breitbart M, Lee WH, Run JQ, Wei CL, Soh SW, et al. RNA viral community in human feces: prevalence of plant pathogenic viruses. PLoS Biol 2006;4:e3.

[11] Reyes A, Haynes M, Hanson N, Angly FE, Heath AC, Rohwer F, et al. Viruses in the faecal microbiota of monozygotic twins and their mothers. Nature 2010;466:334-38.

[12] Minot S, Sinha R, Chen J, Li H, Keilbaugh SA, Wu GD, et al. The human gut virome: inter-individual variation and dynamic response to diet. Genome Res 2011;21:1616-625.

[13] Minot S, Bryson A, Chehoud C, Wu GD, Lewis JD, Bushman FD Rapid evolution of the human gut virome. Proc Natl Acad Sci U S A 2013;110:12450-5.

[14] Modi SR, Lee HH, Spina CS, Collins JJ. Antibiotic treatment expands the resistance reservoir and ecological network of the phage metagenome. Nature 2013;499:219-22.

[15] Thurber RV, Haynes M, Breitbart M, Wigley L, Rohwer F. Laboratory procedures to generate viral metagenomes. Nat Protoc 2009;4:470-83.

[16] Jones TH, Johns MW. Improved detection of F-specific RNA coliphages in fecal material by extraction and polyethylene glycol precipitation. Appl Environ Microbiol 2009;75:6142-6.

[17] Colombet J, Robin A, Lavie L, Bettarel Y, Cauchie HM, Sime-Ngando T. Virioplankton 'pegylation': use of PEG (polyethylene glycol) to concentrate and purify viruses in pelagic ecosystems. J Microbiol Methods 2007;71:212-9.

[18] Patel A, Noble RT, Steele JA, Schwalbach MS, Hewson I, Fuhrman JA. Virus and prokaryote enumeration from planktonic aquatic environments by epifluorescence microscopy with SYBR Green I. Nat Protoc 2007;2:269-76.

[19] Wen K, Ortmann AC, Suttle CA. Accurate estimation of viral abundance by epifluorescence microscopy. Appl Environ Microbiol 2004;70:3862-7.

[20] Mills S, Ross PR, Neve H, Coffey A. Bacteriophage and anti-phage mechanisms in lactic acid bacteria, in Salminen S, von Wright A, Lahtinen S, Ouwehand AC (Eds). Lactic Acid Bacteria: Microbial and Functional Aspects, 4th edition, Taylor & Francis, London, 2011, pp 165-86.

514 [21] Fuhrman, JA, Griffith JF, Schwalbach, MS. Prokaryotic and viral diversity patterns in
515 marine plankton. Ecol Res 2002;17:183-94.
516 [22] Hoyles L, Collins MD, Falsen E, Nikolaitchouk N, McCartney AL. Transfer of members
517 of the genus *Falcivibrio* to the genus *Mobiluncus*, and amended description of the genus
518 *Mobiluncus*. Syst Appl Microbiol 2004;27:72-83.
519 [23] Hoyles, L. *In vitro* examination of the effect of orlistat on the ability of the faecal
520 microbiota to utilize dietary lipids. 2009, PhD thesis, University of Reading, UK.
521 [24] Klieve AV, Swain RA. Estimation of ruminal bacteriophage numbers by pulsed-field gel
522 electrophoresis and laser densitometry. Appl Environ Microbiol 1993;59:2299-303.
523 [25] Brüssow H, Hendrix RW. Phage genomics: small is beautiful. Cell 2002;108:13-6.
524 [26] Hoyles L, McCartney AL. What do we mean when we refer to *Bacteroidetes*
525 populations in the human gastrointestinal microbiota? FEMS Microbiol Lett 2009;299:175-
526 83.
527 [27] Colson P, Fancello L, Gimenez G, Armougom F, Desnues C, Fournous G, et al.
528 Evidence of the megavirome in humans. J Clin Virol 2013;57:191-200.
529 [28] Sime-Ngando T, Lucas S, Robin A, Tucker KP, Colombet J, Bettarel Y, et al. Diversity
530 of virus-host systems in hypersaline Lake Retba, Senegal. Environ Microbiol 2011;13:1956-
531 1972.
532 [29] Kulikov EE, Isaeva AS, Rotkina AS, Manykin AA, Letarov AV. Diversity and dynamics
533 of bacteriophages in horse feces. Mikrobiologiya 2007;76:271-78.
534 [30] Rohozinski J, Girton LE, Van Etten JL. *Chlorella* viruses contain linear nonpermuted
535 double-stranded DNA genomes with covalently closed hairpin ends. Virology 1989;168:363-
536 69.
537

538 **Fig. 1.** Number of VLPs detected in FFs by using EFM. Values are shown as mean + SD (*n* =
539 2 per donor). Passing samples through a 0.22 μm filter reduced the number of VLPs present
540 in samples by approximately half (data not shown). White bars, per ml FF (faeces); grey bars,
541 per g faeces.
542
543

Fig. 2. Image showing faecal VLPs as they appear under an epifluorescence microscope (×1000 magnification). Bacteriophages appear as 'pinpricks' of light when stained with SYBR Gold. The brightest 'pinpricks' (black arrows) represent clusters of up to 30 VLPs. The black dot (shown in the white circle) on the images is due to a scratch on the microscope lens. (a) Contamination on filter, which gives an idea of the size and appearance of a bacterium in comparison with VLPs. (b) Negative control, 0.1 μm-filtered, autoclaved H_2O. (c) Negative control, sterile 0.5 % 'Lab-Lemco'/6 % NaCl. (d) Donor 1, 0.22 μm-filtered sample; (e) Donor 1 0.45 μm-filtered sample. (f) Donor 2, 0.22 μm-filtered sample; (g) Donor 2, 0.45 μm-filtered sample. Scale bar, 10 μm.

554 **Fig. 3.** Transmission electron micrographs revealing the diversity of VLPs found in faeces.
555 (a) Donor 1 (female, 41 years) and (b) Donor 2 (female, 36 years) after extraction of VLPs in
556 TBT buffer.
557

Fig. 4. Transmission electron micrographs of VLPs found in faeces of Donor 2 (female, 36 years) after extraction of VLPs in TBT buffer showing *Myoviridae* phages with radial whiskers attached to the capsids (top) and clusters of *Myoviridae* phages adsorbing to membrane vesicle material (bottom).

563 **Fig. 5.** Transmission electron micrographs revealing the diversity of VLPs found in caecal
564 effluents. Sample (a) L08, (b) L10, (c) L16 and (d) L17 after extraction of VLPs in LL. No
565 clinical data were available for these individuals.
566

Fig. 6. Use of PFGE to examine VLP populations in PEG-precipitated FFs (faeces), and demonstration that increasing the temperature at which samples are heated prior to loading onto the gel can affect the diversity uncovered. The gel was stained with SYBR Gold and visualized as described in Methods. This is a crude (but inexpensive) method of examining VLP populations in human faeces, and demonstrates that each individual harbours a unique VLP profile. Ladder, lambda ladder (#340; New England Biolabs).

574 **Fig. 7.** Isolation of high-quality DNA from CsCl-purified samples of human gastrointestinal
575 VLPs. VLP preparations were CsCl-purified and DNA extracted as described in Methods.
576 Aliquots (10 μl) of sample were run on a 0.8 % agarose gel at 90 V for 30 min. Lane 1,
577 molecular size ladder; lane 2, ϕKPLN1 positive control (a bacteriophage isolated from the
578 human caecum that infects *Klebsiella pneumoniae* subsp. *pneumoniae* K2 strains; L. Hoyles,
579 unpublished data), 63 ng DNA/μl; lane 3, VLP extract from faeces of Donor 4, 3.1 ng
580 DNA/μl; lane 4, VLP extract from faeces of Donor 3, 4.5 ng DNA/μl; lane 5, VLP extract
581 from caecal sample L07 (healthy), 0.2 ng DNA/μl; lane 6, VLP extract from caecal sample
582 L08 (no clinical data), 2.7 ng DNA/μl; lane 7, VLP extract from caecal sample L10 (no
583 clinical data), 3.8 ng DNA/μl. The two black arrows highlight the presence of RNA in the
584 sample from Donor 3, confirmed by treating the CsCl-purified sample with 20 μl of 1 mg/mL
585 RNAse A (bovine pancreas, ≥70 Kunitz/mg protein, prepared in 0.1 μm-filtered, sterile H_2O)
586 with the DNAse in a second DNA extraction (not shown). VLPs and DNA were extracted
587 from human caecum samples (1:4 LL, v/v) as described in Methods as part of a study
588 examining the microbiota associated with the human caecum (L. Hoyles, unpublished data).
589 With the exception of the sample in lane 5, DNA visible to the naked eye was extracted from
590 all samples of gastrointestinal origin.
591