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

Article

Accepted Version

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1016/j.resmic.2014.10.006

Publisher: Elsevier

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

2   **microbiota**

3

4   Lesley Hoyles[a], Anne L. McCartney[b], Horst Neve[c], Glenn R. Gibson[b], Jeremy D. Sanderson[d],

5   Knut J. Heller[c], Douwe van Sinderen[a,e]*

6

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

8   [b]Food Microbial Sciences Unit, Department of Food and Nutritional Sciences, The University

9   of Reading, Whiteknights, Reading RG6 6UR, UK

10   [c]Institute of Microbiology and Biotechnology, Max Rubner-Institut, Federal Research

11   Institute of Nutrition and Food, Kiel, Germany

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

13   College London, London, UK

14   [e]Alimentary Pharmabiotic Centre, Biosciences Institute, National University of Ireland,

15   Western Road, Cork, Ireland

16

17   Lesley Hoyles lesley.hoyles@reading.ac.uk

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

19   Horst Neve horst.neve@mri.bund.de

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

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

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

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

24

24 **ABSTRACT**

25   This work represents an investigation into the presence, abundance and diversity of

26 virus-like particles (VLPs) associated with human faecal and caecal samples. Various

27 methodologies for the recovery of VLPs from faeces were tested and optimized, including

28 successful down-stream processing of such samples for the purpose of an in-depth electron

29 microscopic analysis, pulsed-field gel electrophoresis and efficient DNA recovery. The

30 applicability of the developed VLP characterization method beyond the use of faecal samples

31 was then verified using samples obtained from human caecal fluid.

32

33 **Keywords:** bacteriophages; microbial ecology; gastrointestinal tract; transmission electron

34 microscopy; caecum; faeces.

35

36

37

37   **1. INTRODUCTION**

38   Viruses are the most numerous biological entities within the biosphere of our planet

39   (being present at an estimated number of $\sim 10^{31}$), with bacteriophages representing the most

40   abundant group of environmental viruses [1, 2]. Bacteriophages are ubiquitous viruses that

41   infect bacterial cells and disrupt their metabolism. Multiple bacteriophage types can infect a

42   specific microbial isolate, with most bacteriophages infecting only certain species or even

43   strains of bacteria [3].

44   Although they are abundant and potentially important to microbial populations

45   indigenous to different ecological niches within the human gastrointestinal tract and to host

46   health, little attention has been paid to bacterial virus-like particle (VLP) assemblages and

47   their interactions with the gastrointestinal microbiota and/or human host until recently.

48   Various publications have highlighted the potential importance of bacteriophages in

49   inflammation states, including Crohn's disease [4, 5, 6], and as therapeutic agents [7, 8].

50   Breitbart *et al.* [9] conducted the first metagenomic study on dsDNA-containing VLPs

51   associated with the human faecal virome using a fresh sample from a healthy adult male. On

52   the basis of previous estimates of gut microbial diversity, Breitbart *et al.* [9] predicted that

53   there are two to five times more viral genotypes ($\sim 1,200$ viral genotypes) present in the

54   human gastrointestinal microbiota than the number of bacterial species, with the vast majority

55   of these VLPs representing bacteriophages and prophages. RNA viruses present in human

56   faeces have been found to be mostly associated with plant viruses, such as Pepper Mild

57   Mottle Virus, with RNA bacteriophages making a minimal contribution to the diversity of the

58   intestinal virome [10]. Using frozen samples from four adult female monozygotic twins and

59   their mothers, Reyes *et al.* [11] demonstrated that each individual harbours a unique virome

60   (ssDNA and dsDNA) regardless of their genetic relatedness to another individual, and that

61   intrapersonal virome diversity is very low, with >95 % of virotypes retained over a one-year

62   period (between 71 and 2,773 viral genotypes identified among the samples). More recently,

63   Minot *et al.* [12] confirmed inter-individual variation in the virome and, similar to

64   modulation of the faecal bacteriome by foodstuffs, demonstrated that diet influences host

65   faecal virome structure. Minot *et al.* [13] also confirmed the relative stability of an

66   individual's virome, monitoring 24 faecal samples from one individual over a 2.5-year

67   period.

68   The murine virome has been used to demonstrate enrichment of bacteriophage-

69   encoded, antibiotic-resistance genes (related and unrelated to the drug used) after antibiotic

70   therapy, with the adaptive capacity of the virome (specifically its bacteriophage component)

71   suggested to protect gut bacteria, thereby preserving the microbiota's robustness during

72   antibiotic stress [14].

73   It is clear from the aforementioned studies that the virome–bacteriome community in

74   the gut is governed by complex and dynamic interactions in health, and that its balance may

75   be disturbed when under stress, e.g. during antibiotic intervention [14]. Although

76   metagenomic studies have greatly improved our understanding of the virome associated with

77   the human and murine gut microbiomes, it has been notable how little bacteriophage

78   material, and consequently DNA, has been isolated from samples when this information was

79   included in a publication. Reyes *et al.* [11] reported the isolation of ~500 ng DNA from 2 to 5

80   g of frozen faeces, whereas Thurber *et al.* [15] stated that between 500 and 3000 ng of DNA

81   could be isolated from 500 g of human faeces, though the publication they cite [9] provides

82   no information regarding the amount of DNA isolated from the 500 g faecal sample

83   examined in the original study. In addition, to the best of our knowledge, no attempts have

84   been made to enumerate bacteriophages in faecal filtrates prior to CsCl purification. The aims

85   of this study were to develop reliable and effective methods for the recovery and

86   characterization of VLPs in human samples, and to apply these methods to human faecal and

87   caecal samples to demonstrate the methods' efficiency.

88

88   **2. MATERIALS AND METHODS**

89   **2.1 Processing of and isolation of VLPs from faecal and caecal samples**

90   Faecal samples were obtained from six healthy adult (2 male, 4 female) donors of

91   between 23 and 52 years of age, all of whom were members of the Department of Food and

92   Nutritional Sciences, University of Reading. None of the donors had taken antibiotics,

93   prebiotics and/or probiotics in the 3 months prior to sampling, and none had any history of

94   gastrointestinal disorder. All donors gave informed oral consent for their faecal samples to be

95   used for microbiological analyses. Ethical approval for the collection of caecal effluent from

96   patients was obtained from St Thomas' Hospital Research Ethics Committee (06/Q0702/74)

97   covering Guy's and St Thomas' Hospitals and transferred by agreement to London Bridge

98   Hospital. Where available, clinical information for the colonoscopy patients is given in the

99   text. Samples were collected at Reading (faeces) or transported from St Thomas' Hospital

100   (caecum), and maintained under anaerobic conditions (faeces, MACS1000 anaerobic

101   workstation, Don Whitley Scientific, gas composition 80 % N_2, 10 % H_2, 10 % CO_2; caecum,

102   on ice in a gas jar with an anaerobic gas-generating pack; Oxoid Ltd) for a maximum of 2 h

103   before processing. Caecal samples were collected during routine colonoscopy following

104   preparation of the bowel with sodium picosulphate and a reduced fibre diet for 3 days. Liquid

105   residue in the caecum was aspirated via the colonoscope suction channel into a standard 30

106   ml trap specimen container and immediately transferred to a gas jar.

107   Faecal homogenates (20 %, w/v, from 25 g of faecal material) were prepared in sterile

108   TBT [0.1 μm filtered prior to autoclaving; 100 mM Tris/HCl, pH 8.0; 100 mM NaCl; 10 mM

109   $MgCl_2.6H_2O$] or sterile 0.5 % 'Lab-Lemco'/6 % NaCl (LL [16]). Caecal homogenates (20 %,

110   v/v, from 10 to 30 ml of caecal effluent) were prepared in LL. Faecal and caecal samples

111   were placed into a filter stomacher bag and homogenized in a Stomacher 400 Lab System

112   (Seward) for 120 s at low speed. The bag was removed from the stomacher and massaged

113   manually to further disperse any large particles remaining in the sample; the sample was then

114   stomached for a further 120 s. The homogenate was kept on ice for 2 h to allow desorption of

115   VLPs from solid material, then centrifuged at 11,180 *g* for 30 min at 10ºC. The supernatant

116   was transferred to a fresh tube and centrifuged again at 11,180 *g* for 30 min at 10ºC.

117   Supernatant [herein referred to as faecal filtrate (FF) for both faecal and caecal samples] was

118   passed through sterile 0.45 μm cellulose acetate filters (Millipore) and the FF was then

119   collected in a sterile container. An aliquot (10 μl) of each of the FFs was examined by

120   epifluorescence microscopy (EFM) after viral particles had been stained with SYBR Gold

121   (see below) to confirm that the samples were free of bacteria.

122   FF prepared in LL was used for enumeration of VLPs in faeces via EFM (see below)

123   and 1 ml aliquots of FF prepared in TBT were for examination by transmission electron

124   microscopy (TEM; see below).

125   Aliquots of FF and LL (two sets of three aliquots of 100 μl each: one for aerobic

126   cultivation, one for anaerobic cultivation) were spread onto Columbia blood agar containing

127   5 % laked horse blood (Oxoid), and incubated aerobically and anaerobically. Sterility of the

128   filtrates was confirmed by the absence of microbial growth on plates following incubation for

129   2 (aerobic) and 5 (anaerobic) days. Sterile brain–heart infusion broth was inoculated with 100

130   μl of FF and LL, and incubated aerobically for 2 days: aliquots (100 μl) were spread in

131   triplicate on plates to confirm sterility of the broth cultures. In addition, 10-ml aliquots of FF

132   were stored at 4°C for 6 months after collection, and were found to be free of bacteria when

133   checked (by plating) at monthly intervals.

134   Poly(ethylene glycol) (PEG; BioUltra, 8000; Sigma) was added to 2 × 20 ml aliquots

135   of the LL-prepared 0.45 μm-filtered FFs (final concentration of PEG, 10 %, w/v). The

136   samples were shaken gently to dissolve the PEG, then left at 4°C for 16 h ([16] used 16 h as

137   they observed an improved recovery of RNA bacteriophage compared to shorter incubations;

138   [17] used 10 h, as they determined this to be longer than the minimum time required to obtain

139   a constant and stable white layer of viral precipitate). PEG-precipitated VLPs were then

140   harvested by centrifugation at 4,500 *g* for 30 min in a swing-out-bucket rotor at 4°C. The

141   supernatant was removed and the pellet resuspended in 1 ml TBT for pulsed-field gel

142   electrophoresis (PFGE) or 5 ml TBT for purification of VLPs on a CsCl gradient [15] with

143   ultracentrifugation performed at 100,000 *g* for 2 h at 10°C in a fixed-angle Type 50 Ti rotor

144   (Beckman Coulter). VLPs recovered by CsCl (1.35 and 1.5 g/cm^3 fractions) were dialysed

145   (12,000 Da cut-off) twice against 400 ml sterile TBT, passed through a sterile 0.45 μm

146   cellulose acetate filter and stored at 4°C until DNA was extracted. An aliquot (5 μl) of each

147   of the CsCl-purified samples was viewed using EFM after viral particles had been stained

148   with SYBR Gold (see below) to confirm that the samples were free of bacteria.

149

150   **2.2 Epifluorescence microscopy (EFM)**

151   A combination of the methods of Thurber and Patel [15, 18] was used to prepare FFs

152   (faecal only) for EFM. VLPs present in FF were not fixed in paraformaldehyde prior to

153   enumeration, as Wen *et al.* [19] showed that aldehyde fixation leads to a rapid loss in viral

154   abundance. The SYBR Gold concentration of 1–5× recommended by [15] did not work with

155   FFs; therefore, the 400× concentration recommended by [18] for use with planktonic aquatic

156   samples was used in this study.

157   The filtration system (see [15] for specifics of the set-up employed), including a

158   glass-graduated column, for collecting VLPs on filters was cleaned using 5 ml of 0.1 μm-

159   filtered, sterile H$_2$O and 5 ml of 0.1 μm-filtered ethanol. A sterile pair of flat-tipped forceps

160   was used to remove a 0.02-μm white Anodisc 25 membrane (Whatman) from its box, and the

161   filter was fitted to the glass frit of the filtration system under a low vacuum [<10 psi (~62

162   kPa)]. Duplicate 2 ml aliquots of 0.1 μm filtered, sterile H$_2$O and 10 μl of sterile LL in 2 ml

163   0.1 μm filtered, sterile H_2O were used as negative controls (to assess whether there was any

164   contamination in the water, the LL, the filter tower or the filter surface) and passed through

165   filters under low vacuum before any of the FF aliquots were filtered in duplicate. For each

166   sample, a 10 μl aliquot of 0.45 μm-filtered FF was added to 2 ml of the 0.1 μm filtered,

167   sterile H_2O and the suspension gently mixed. The diluted sample was then introduced into the

168   filter system with the low vacuum maintained. The liquid was passed through the filter, and

169   VLPs were collected on the filter. The clip and glass column were carefully removed from

170   the filtration system and the filter was gently removed from the glass frit (still under

171   vacuum). While holding the filter with a forceps, the back of the filter was gently blotted dry

172   on a clean Kimwipe; the filter was then left to dry on a Kimwipe in a sterile Petri dish for 10

173   min in a dark box. The filtration system was cleaned with 5 ml of 0.1 μm-filtered, sterile H_2O

174   and 5 ml of 0.1 μm-filtered ethanol, and the next sample processed.

175   Filters were stained with 100 μl droplets of 400× SYBR Gold and dried as described

176   by [18], prior to being applied in pairs to sterile glass slides. The mountant (20 μl per

177   coverslip) ProLong® Gold antifade reagent (Invitrogen) was added to 25 mm glass coverslips.

178   Coverslips were picked up with sterile forceps, inverted and placed on the Anodisc filters on

179   the microscope slides. Slides were then left at room temperature in the dark for 16 h to allow

180   the antifade to cure. Slides prepared in this manner can be stored at room temperature for 1

181   week or at 4ºC for 2 months.

182   Slides were viewed at 1000× magnification under a Nikon Microphot-SA microscope

183   fitted with a B-2A (blue excitation) filter and attached to a CoolSNAP-Pro MONOCHROME

184   (Media Cybernetics Inc.) camera. Images were captured using Image-Pro PLUS version

185   4.5.0.19 (Media Cybernetics Inc.), and VLPs in 25 fields of view were enumerated by eye.

186   The number of VLPs per millilitre of FF was calculated using the following equation: mean

187   number of VLPs in 25 fields of view × 100 × 25760.205; where 100 represents taking the

188 sample back to per millilitre FF and 25760.205 represents the number of fields of view on a

189 filter. To give an estimate of the number of VLPs per gram of faeces, results were multiplied

190 by a factor of 5. The detection limit of the method was 1.03×10^5 VLPs/ml FF or 5.15×10^5

191 VLPs/g faeces.

192

193 **2.3 TEM of FFs (faecal and caecal)**

194   Aliquots (1 ml) of FF prepared in TBT (faeces) or LL (caecum) were used for TEM

195 analysis within 1 week of collection: these filtrates had not been concentrated using PEG.

196 Carbon films (~3×3 mm in size) were floated from mica-sheets into a drop of filtrate (100

197 μl). After an adsorption time of 5–10 min, samples were transferred into a drop of 1 % (v/v)

198 of EM-grade glutaraldehyde (20 min) and subsequently into a drop of 2 % (w/v) uranyl

199 acetate for negative staining (1–2 min). After two washes for a few seconds in drops of

200 distilled water, samples were picked up with 400-mesh copper grids (Plano, Wetzlar, D).

201 Electron micrographs were taken in a Tecnai 10 transmission electron microscope (FEI

202 Company, Eindhoven, the Netherlands) at an accelerating voltage of 80 kV. Digital

203 micrographs were taken with a Megaview G2 CCD camera (Olympus SIS, Münster,

204 Germany). For estimation of VLP titers on the EM grids, a freshly prepared *Lactococcus*

205 *lactis* bacteriophage preparation with defined titer and unique morphology (i.e. prolate-

206 headed bacteriophage P001 [20]) was added in concentrations of 10^7 and 10^8 plaque-forming

207 units (pfu) per ml to the faecal sample of donor 2 (with highest bacteriophage titer according

208 to TEM analysis).

209

210 **2.4 PFGE**

211   Aliquots (40 μl) of PEG-precipitated samples resuspended in 1 ml TBT were used for

212 PFGE. These were heated at 60 or 75°C as appropriate. An aliquot (25 μl) of each heated

213   (viscous) suspension and loading dye (5 μl) were combined and loaded into wells, which

214   were sealed with molten agarose. Run conditions for PFGE were as described by Fuhrman *et*

215   *al.* [21]: a 1 % agarose gel (SeaKem LE agarose) was made in 0.5× TBE [250 ml of 5× TBE

216   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

217   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×

218   TBE, with a 1−10 s switch time. Gels were stained with ethidium bromide (5 μg/ml) for 20

219   min and destained in distilled H_2O for 10 min, or with 1× SYBR Gold (Molecular Probes)

220   made in 0.5× TBE. Bands of DNA were visualized under UV light (ethidium bromide) or

221   using a Dark Reader DR89X Transilluminator (Integrated Scientific Solutions Inc.) (SYBR

222   Gold).

223

224   **2.5 Isolation of DNA from CsCl-purified VLPs**

225   DNA was extracted from 500 μl portions of CsCl-purified VLP samples according to

226   [15]. Prior to extracting DNA from purified VLPs, samples were treated with 20 μl of 1

227   mg/mL DNAse I (from bovine pancreas, 552 Kunitz/mg protein; made in water passed

228   through a 0.1 μm filter prior to autoclaving) for 1 h at 37°C, and were then heated at 80°C for

229   10 min to inactivate the DNAse prior to DNA extraction. DNA was air-dried and

230   resuspended in 50 μl of TE buffer. Sterility of samples (i.e. absence of bacterial DNA) was

231   confirmed by negative PCR from samples (10 μl) with universal primers [22]. DNA from an

232   in-house strain of *Bifidobacterium longum* (4-FAA1; [23]) was used as a positive control;

233   sterile water was used as the negative control.

234

234   **3. RESULTS AND DISCUSSION**

235   **3.1 Examination of VLPs in FFs (faeces) by using EFM**

236   Previous studies examining the faecal virome have used 0.22 μm filters to prepare

237   samples [9, 11, 12, 13]. Preliminary work performed with 0.45 μm-filtered faecal samples

238   and 0.45 μm-filtered samples originating from *in vitro* fermentation systems inoculated with

239   faeces stained with 4′,6-diamidino-2-phenylindole (DAPI) and viewed under an

240   epifluorescence microscope suggested that the abundance of dsDNA VLPs in intestinal

241   samples was high, and that these particles can be enumerated (unpublished observations).

242   Contamination by bacteria was not observed in any of these samples, and confirmed by an

243   absence of cultivable bacteria and no bacteria in TEM analyses, so we decided to continue

244   using larger-pore filters for processing samples. Klieve & Swain [24] had previously used

245   0.45 μm-filtered samples to characterize VLPs present in rumen contents, and reported no

246   problems with contamination by bacteria.

247   To examine the VLPs present in FFs (faeces) by EFM, the method described by [9]

248   and [15] was used initially. However, it was found that we were unable to visualize many, if

249   any, VLPs present in the samples using 1−5× SYBR Gold as the DNA/RNA stain. Increasing

250   the concentration of SYBR Gold to 400×, as used by [18] for planktonic aquatic samples,

251   allowed us to enumerate and detect VLPs present in the FFs prepared from faecal samples of

252   six donors (Fig. 1 and Fig. 2). The samples were extracted in LL, though the method worked

253   as well with samples extracted in TBT (data not shown). To determine the relative abundance

254   of the predominant VLP morphologies in FFs, samples were not concentrated by CsCl

255   centrifugation (Fig. 1). All FFs were found to be free of bacteria by EFM (Fig. 2), with the

256   number of VLPs present in samples ranging from 2.4×10^8 to 1.12×10^9 VLPs/ml FF (mean

257   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).

258   Lepage *et al.* [4] enumerated VLPs in gut mucosal samples from 14 healthy individuals and

259   19 Crohn's disease patients and found on average 1.2×10^9 VLPs/biopsy (range 4.4×10^7–

260   1.7×10^{10}), in agreement with the mean value we present here for faecal VLPs. Of note,

261   Crohn's disease patients harboured significantly ($P = 0.024$) more VLPs than healthy

262   individuals (2.9×10^9 vs 1.2×10^8 VLPs/biopsy) in the study of Lepage *et al.* [4].

263   It is generally accepted that there are around 10 bacteriophages for every microbial

264   cell in environmental samples investigated to date [25]. Extrapolating this figure to the gut

265   microbiota, from fluorescence *in situ* hybridization studies it is estimated that the faecal

266   microbiota harbours $\sim 10^{11}$ bacteria/g faeces in healthy adults [26]. Consequently, one would

267   expect the presence of at least 10^{12} VLPs/g faeces. Enumerating VLPs in faeces via FFs (or

268   any liquid medium) by EFM is highly subjective as a dot of very intense fluorescence may in

269   fact represent a cluster of VLPs (Fig. 2), a phenomenon frequently encountered during this

270   study. Patel *et al.* [18] stated that, to accurately enumerate VLPs by EFM, micro-adjustments

271   using the fine focus of the microscope have to be made to ensure that all viruses in a

272   particular grid-reticle box are counted. We agree with this statement, and furthermore add

273   that many VLPs have very likely been lost during the preparation and filtering of samples,

274   either by association with debris in the initial centrifugations or by being caught in the filters

275   because of clogging or because the VLPs are too big to pass through the pores (e.g. members

276   of the order *Megavirales* [27]). Even after centrifugation, the supernatants from several of the

277   samples, while appearing relatively translucent, were highly viscous (perhaps due to host

278   mucins) and clogged the 0.45 μm filters with less than 2 ml of sample being filtered.

279   Consequently, the values we provide for the numbers of VLPs in FFs and faecal samples are

280   a conservative estimate. We believe the true number of VLPs present in faeces to be higher,

281   possibly between 10^{10} and 10^{12} VLPs/g faeces. Lepage *et al.* [4] determined there to be 10^{10}

282   VLPs/mm^3 tissue in their study of mucosal VLPs.

283

284   **3.2 Examination of VLP diversity in FFs by the use of TEM**

285   VLPs were readily detected in the TBT-extracted faecal samples from all donors by

286   TEM. Bacterial cells were never observed in any samples examined by TEM (limit of

287   detection 10^6 per ml). It was striking how visibly different/distinct the VLP assemblages were

288   between the donors, with no two donors sharing the same VLPs, at least on the basis of

289   morphological appearance (Fig. 3 and Fig. 4; Supplementary Fig. 1). The vast majority of

290   VLPs present in the samples appear to represent bacteriophages. Donor 1's VLP assemblage

291   was predominated by small and large isometric-headed *Siphoviridae* with various tail length

292   sizes (approx. 120 nm, 350 nm, 650 nm, or, in one extreme case, 1220 nm), with some small

293   and large isometric-headed *Myoviridae* also present (Fig. 3a). Notably, donor 2's VLP

294   assemblage was more diverse, predominated by numerous different morphotypes of

295   *Myoviridae* with *Siphoviridae* also present (Fig. 3b). Interestingly, two detached *Myoviridae*

296   tails of extreme length (480 nm) and thickness of the sheaths (40 nm) indicated the presence

297   of giant *Myoviridae* phages (Fig. 3b). Sime-Ngando *et al*. [28] have reported isolation of

298   bacteriophages with tails of 400 nm in length (heads 50–130 nm) from hypersaline lake

299   samples, though images of these large bacteriophages are unavailable for direct comparison

300   of the tails' structures with those found in this study. Donor 2's faecal sample contained

301   distinct small isometric-headed *Myoviridae* phages with uncommon radial fibers (approx 150

302   nm in length) attached to the capsids (~70 nm diameter) (Fig. 4). These fibers are clearly

303   extending the 110-nm tails of these phages, suggesting a primary role in bacteriophage

304   adsorption. This unique bacteriophage morphotype was also documented in low numbers in

305   Donor 4's faecal sample (not shown), and in caecal samples L16 (not shown) and L18

306   (Supplementary Fig. 2d). Fig. 4 also illustrates the apparent clustering of *Myoviridae* phages

307   in samples. VLPs of various morphotypes derived from Donors 3, 4, 5, and 6 are shown in

308   Supplementary Figure 1a–d. "Zeppelin"-like VLPs of constant thickness (56 nm) but

309   different lengths (370–630 nm) were detected in Donor 5 (Supplementary Fig. 1c) and Donor

310   6's samples (Supplementary Fig. 1d). Donor 6's sample was predominated by these

311   "zeppelin"-like VLPs, with only a few bacteriophages detected. Sime-Ngando et al. [28]

312   reported the presence of rods of up to 22–24 ×1000 nm: similar to the 'zeppelins' identified

313   in the present study, these rods had no visible internal or external structures. The largest

314   *Myoviridae* phages found in this study were documented for the sample of Donor 5

315   (Supplementary Fig. 1c), with the phages having prolate heads of 150×115 nm and tails of

316   400 nm in length. The sample from Donor 6 did not reveal the same level of diversity as seen

317   for the other faecal samples and the reason for this is not known (Supplementary Fig. 1d).

318   Enumeration of VLPs in the sample of Donor 2 (highest number of VLPs as detected by

319   TEM) demonstrated there to be approx. 5×10^7 VLPs/ml FF.

320   For the majority of the caecal samples, a lower extent of viral biodiversity was

321   documented. Sample L10 (no clinical data) was unique, containing exclusively small

322   isometric-headed *Podoviridae* (50 nm diameter) and ~20-nm long appendages (Fig. 5b).

323   Samples L08 (healthy) (Fig. 5a), and L02 and L03 (both IBD) (Suppl. Fig. 5b, c) appeared to

324   exclusively contain *Myoviridae* phages. The greatest extent of morphological variation and

325   the highest bacteriophage numbers were seen in caecal samples L16 (no clinical data) and

326   L17 (no clinical data) (Fig. 5c, d) and – to a lesser extent – in samples L01 (diverticulosis)

327   and L18 (no clinical data) (Supplementary Fig. 2a, d). A new type of a giant *Myoviridae*

328   bacteriophage was exclusively found in sample L16 (Fig. 5c) with isometric heads of 125 nm

329   in diameter, with tails of 340 nm in length and unique "curled" tail fibers of ~70-nm in

330   length.

331   Lepage *et al.* [4] demonstrated that gut mucosal samples were predominated by

332   morphotypes consistent with *Siphoviridae*, *Myoviridae* and *Podoviridae*, and that each

333   individual appeared to be colonized by one dominant bacteriophage family. Our TEM results

334   for faecal and caecal VLP assemblages are consistent with these results.

335

336   **3.3 PFGE of VLPs present in FFs**

337   It was clear from the EFM and TEM analyses that, based on the number of VLPs

338   present in our samples, it should be possible to isolate large amounts of VLP-derived DNA

339   from faeces. Attempts to recover abundant VLP DNA from samples using the centrifugation

340   method of Thurber *et al.* [15], and which was used by [9] and [11] to isolate VLP DNA from

341   faecal samples, proved disappointing in terms of the recovered DNA yield. However,

342   increasing the centrifugation speed to 100,000 *g*, routinely used in the laboratory to purify

343   lactococcal bacteriophages for preparation of high-quality DNA and used by Kulikov *et al.*

344   [29] to recover bacteriophages from horse faeces, markedly improved recovery of faecal

345   VLPs.

346   To concentrate VLPs from 20-ml quantities of FF so that they could be applied to

347   CsCl gradients in 5 ml aliquots, it was decided to use PEG precipitation. This method of

348   recovery has previously been used with, for example, marine samples and faecal samples to

349   improve detection of F-specific coliphages in faecal material [16, 17], and allows large

350   starting volumes of sample to be used for recovering VLPs from human faeces. The method

351   can be scaled easily so that the VLPs from larger volumes of FF are precipitated for

352   collection by centrifugation: we used 20 ml of FF from each donor here to demonstrate the

353   efficacy of the method, but the entire FF derived for each donor's 25 g of starting material

354   could have been used for PEG precipitation of VLPs.

355   PEG-precipitated samples resuspended in 1 ml TBT were examined using PFGE.

356   Aliquots (40 µl) of the resuspended pellets were heated at 60°C [9] or 75°C ([24] used this

357   temperature to inactivate nucleases) before PFGE. Profiles containing one or more bands

358   were observed for samples from all donors (Fig. 6). Following treatment at 60ºC, Donor 1's

359   sample produced a faint band at 105 kb, with this band being more prominent following

360   sample treatment at 75ºC; at the higher temperature, bands were also observed at 135 kb and

361   just below 48.5 kb, with the most prominent band at 79 kb. Donor 2's 60ºC-treated sample

362   produced a profile with a faint smear and a thin band at 84 kb; with the 75ºC-treated sample,

363   the smear was more pronounced, with a thick band observed between 75 and 84 kb. Similar

364   to Donor 2, the smear in Donor 4's sample became more pronounced following treatment at

365   the higher temperature. For Donor 3's 60ºC-treated sample, two strong bands (at 66 and 97

366   kb) were observed; in the 75ºC-treated samples, bands were observed at 57 kb, 75 kb and 97

367   kb. The profile of Donor 5 was most affected by changing the temperature at which samples

368   were treated prior to loading on the gel: the 60ºC-treated sample had three bands visible

369   (most prominent at 100 kb, fainter bands at 48.5 kb and 66 kb), whereas the 75ºC-treated

370   sample had these three bands with at least another 8 bands visible between 48.5 and 97.0 kb.

371   Donor 6's sample had a faint band just below 48.5 kb after heating at 60ºC, whereas the

372   sample heated at the higher temperature produced two prominent bands that were smaller

373   than 48.5 kb. The results from the gel were reproducible (triplicate gels run; data not shown).

374   It is unsurprising that heating at 75ºC produced more complex banding patterns, as we would

375   expect the higher temperature to disassemble (some of) the capsids of VLPs resistant to

376   heating at 60ºC, thereby releasing packaged DNA.

377   Using this method, it was possible to visualize VLP DNA when stained with SYBR

378   Gold (Fig. 6) or ethidium bromide (Supplementary Fig. 3). It was clear that increasing the

379   temperature at which the samples were heated prior to loading onto the gel revealed greater

380   diversity in the samples, and that each individual harboured a unique VLP assemblage (Fig.

381   4). This is in agreement with the findings of the metagenomic studies of [11] and [12]. The

382   sample from Donor 6 did not reveal the same level of diversity as seen for the other donors,

383   in agreement with the results shown by TEM (Supplementary Fig. 1d). This donor's sample

384   was predominated by "zeppelin"-like VLPs that may not have released their nucleid acids by

385   the conditions employed here.

386   Heating PEG-precipitated samples allowed us to generate PFGE images that

387   demonstrated that each individual harbours a unique VLP assemblage, and encouraged us to

388   continue with studies to improve recovery of VLP DNA from human faeces. It also showed

389   that our method of recovery of VLPs present in faeces was superior to that used previously

390   by [9] to generate a viral assemblage fingerprint. In that study, the authors stated that the

391   limited amount of DNA recovered using tangential flow filtration with a 500 g faecal sample

392   from a 33-year-old made it necessary to enhance the sample bands on the gel relative to the

393   ladder. In contrast, no enhancement of the bands detected by SYBR Gold (Fig. 6) or ethidium

394   bromide (Supplementary Fig. 3) was required using the methodology described herein,

395   although SYBR Gold was, as expected, far more sensitive than ethidium bromide. PEG

396   precipitation of VLPs present in 20 ml of FF, as used in this study, equates to the extraction

397   of VLPs from ~4 g faeces (sample losses are seen during filtration, and vary from donor to

398   donor due to the differences in viscosity of faecal supernatant obtained after removal of most

399   bacteria and debris from samples after centrifugation), with 50 μl of the resuspended PEG

400   precipitate representing VLPs isolated from ~200 mg of faeces.

401   The current study used a crude method of extracting DNA for PFGE. Using the

402   method of Rohozinski *et al.* [30], in which VLPs are embedded in agarose blocks, with CsCl-

403   purified VLPs prior to heating may allow better 'fingerprinting' of VLPs in faecal samples.

404   This could, for example, be used as an inexpensive means of determining the effect of

405   freeze–thawing faecal samples prior to the recovery of VLPs from faecal samples. All of the

406   metagenomic studies conducted to date on the human faecal virome have used samples that

407   have been frozen prior to recovery of VLPs from samples.

408

409 **3.4 Isolation of high-quality DNA from FFs**

410   Reyes *et al.* [11] and Thurber *et al.* [15] (based on the work of [9]) have reported

411 recovery of ~500 ng (from 2–5 g of frozen faeces) and 500–3000 ng (from 500 g of fresh

412 faeces), respectively. The method of Reyes *et al.* [11] is closest to that presented herein for

413 the recovery of VLPs from faeces [i.e. they hand-filtered samples, whereas Breitbart *et al.* [9]

414 used tangential flow filtration to process their sample]. As stated above, 20 ml of FF equates

415 to ~4 g faeces. We processed 500 μl portions of 3.5–4 ml CsCl-purified VLPs for DNA

416 extraction, and resuspended the DNA in 50 μl of TE (Fig. 7). If we had processed the entire

417 PEG-precipitated, CsCl-purified VLP sample for each donor, we would have recovered

418 between ~1000 and 1800 ng from ~4 g of faeces (theoretically ~6000–11000 ng from 25 g of

419 faeces). This increase in recovery of DNA is thought to be mainly due to the use of 0.45 μm

420 filters rather than 0.22 μm filters to process samples, supported by our observation that

421 passing FFs through 0.22 μm filters led to VLP counts by EFM that were approximately half

422 of those of the 0.45 μm-filtered samples. Increasing filter size and using PEG precipitation to

423 increase the volume of sample that can be processed would allow greater sampling of the

424 faecal virome in future metagenomics studies.

425   In addition to demonstrating the utility of the method with faecal samples, we have

426 successfully isolated VLP DNA from human caecal effluent (Fig. 7). For each of the samples,

427 20 ml of 0.45 μm-filtered caecal filtrate prepared in LL was PEG-precipitated, with the VLPs

428 then purified on CsCl gradients. These 20-ml aliquots equated to ~4 ml caecal effluent from

429 each of the three samples examined. We started with 30, 30 and 10 ml, respectively, of

430 effluent from caecal samples L07 (healthy), L08 (no clinical data) and L10 (no clinical data).

431 Therefore, if the filtrate from entire homogenates of these samples had been processed, we

432 would have expected to isolate between ~600 and 8000 ng VLP DNA from caecal effluent.

433   Consequently, it should be possible to conduct metagenomics studies of the VLP

434   assemblages associated with the human caecum using the methods described herein.

435

436   **3.5 Conclusions**

437   We have presented a series of methods for enumerating and characterizing VLPs

438   present in human faecal and caecal samples. On the basis of enumeration of VLPs using

439   EFM, there are ~10^8 VLPs/ml FF, with faeces thought contain up to 10^{12} VLPs per gram of

440   sample. TEM analysis of faecal VLPs indicates that an individual is colonized by one

441   dominant bacteriophage family, with _Myoviridae_ and _Siphoviridae_ representing the main

442   families of bacteriophage detected in faeces. PFGE analysis of PEG-concentrated samples

443   has demonstrated that each individual harbours a unique VLP population, and that the

444   temperature at which samples are treated greatly affects diversity observed. Using 0.45 μm

445   filters to prepare samples, it is possible to isolate twice as much DNA as reported previously

446   from ~4 g of faeces. The inclusion of PEG precipitation in the methodology means that VLPs

447   from large volumes of FF can be concentrated, allowing recovery of microgram quantities of

448   VLP DNA from faecal samples. In addition, PEG precipitation and CsCl purification can be

449   used to recover VLPs from human caecal samples, with nanogram quantities of VLP DNA

450   being recovered from the processed samples.

451

452

453

453 **4. ACKNOWLEDGEMENTS**

457

458

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.

483   [11] Reyes A, Haynes M, Hanson N, Angly FE, Heath AC, Rohwer F, et al. Viruses in the

484   faecal microbiota of monozygotic twins and their mothers. Nature 2010;466:334-38.

485   [12] Minot S, Sinha R, Chen J, Li H, Keilbaugh SA, Wu GD, et al. The human gut virome:

486   inter-individual variation and dynamic response to diet. Genome Res 2011;21:1616-625.

487   [13] Minot S, Bryson A, Chehoud C, Wu GD, Lewis JD, Bushman FD Rapid evolution of the

488   human gut virome. Proc Natl Acad Sci U S A 2013;110:12450-5.

489   [14] Modi SR, Lee HH, Spina CS, Collins JJ. Antibiotic treatment expands the resistance

490   reservoir and ecological network of the phage metagenome. Nature 2013;499:219-22.

491   [15] Thurber RV, Haynes M, Breitbart M, Wigley L, Rohwer F. Laboratory procedures to

492   generate viral metagenomes. Nat Protoc 2009;4:470-83.

493   [16] Jones TH, Johns MW. Improved detection of F-specific RNA coliphages in fecal

494   material by extraction and polyethylene glycol precipitation. Appl Environ Microbiol

495   2009;75:6142-6.

496   [17] Colombet J, Robin A, Lavie L, Bettarel Y, Cauchie HM, Sime-Ngando T. Virioplankton

497   'pegylation': use of PEG (polyethylene glycol) to concentrate and purify viruses in pelagic

498   ecosystems. J Microbiol Methods 2007;71:212-9.

499   [18] Patel A, Noble RT, Steele JA, Schwalbach MS, Hewson I, Fuhrman JA. Virus and

500   prokaryote enumeration from planktonic aquatic environments by epifluorescence

501   microscopy with SYBR Green I. Nat Protoc 2007;2:269-76.

502   [19] Wen K, Ortmann AC, Suttle CA. Accurate estimation of viral abundance by

503   epifluorescence microscopy. Appl Environ Microbiol 2004;70:3862-7.

504   [20] Mills S, Ross PR, Neve H, Coffey A. Bacteriophage and anti-phage mechanisms in

505   lactic acid bacteria, in Salminen S, von Wright A, Lahtinen S, Ouwehand AC (Eds). Lactic

506   Acid Bacteria: Microbial and Functional Aspects, 4th edition, Taylor & Francis, London,

507   2011, pp 165-86.

508 [21] Fuhrman, JA, Griffith JF, Schwalbach, MS. Prokaryotic and viral diversity patterns in

509 marine plankton. Ecol Res 2002;17:183-94.

510 [22] Hoyles L, Collins MD, Falsen E, Nikolaitchouk N, McCartney AL. Transfer of members

511 of the genus *Falcivibrio* to the genus *Mobiluncus*, and amended description of the genus

512 *Mobiluncus*. Syst Appl Microbiol 2004;27:72-83.

513 [23] Hoyles, L. *In vitro* examination of the effect of orlistat on the ability of the faecal

514 microbiota to utilize dietary lipids. 2009, PhD thesis, University of Reading, UK.

515 [24] Klieve AV, Swain RA. Estimation of ruminal bacteriophage numbers by pulsed-field gel

516 electrophoresis and laser densitometry. Appl Environ Microbiol 1993;59:2299-303.

517 [25] Brüssow H, Hendrix RW. Phage genomics: small is beautiful. Cell 2002;108:13-6.

518 [26] Hoyles L, McCartney AL. What do we mean when we refer to *Bacteroidetes*

519 populations in the human gastrointestinal microbiota? FEMS Microbiol Lett 2009;299:175-

520 83.

521 [27] Colson P, Fancello L, Gimenez G, Armougom F, Desnues C, Fournous G, et al.

522 Evidence of the megavirome in humans. J Clin Virol 2013;57:191-200.

523 [28] Sime-Ngando T, Lucas S, Robin A, Tucker KP, Colombet J, Bettarel Y, et al. Diversity

524 of virus-host systems in hypersaline Lake Retba, Senegal. Environ Microbiol 2011;13:1956-

525 1972.

526 [29] Kulikov EE, Isaeva AS, Rotkina AS, Manykin AA, Letarov AV. Diversity and dynamics

527 of bacteriophages in horse feces. Mikrobiologiya 2007;76:271-78.

528 [30] Rohozinski J, Girton LE, Van Etten JL. *Chlorella* viruses contain linear nonpermuted

529 double-stranded DNA genomes with covalently closed hairpin ends. Virology 1989;168:363-

530 69.

531

531   **Fig. 1.** Number of VLPs detected in FFs by using EFM. Values are shown as mean + SD (n =

532   2 per donor). Passing samples through a 0.22 μm filter reduced the number of VLPs present

533   in samples by approximately half (data not shown). White bars, per ml FF (faeces); grey bars,

534   per g faeces.

535

536

536   **Fig. 2.** Image showing faecal VLPs as they appear under an epifluorescence microscope

537   (×1000 magnification). Bacteriophages appear as 'pinpricks' of light when stained with

538   SYBR Gold. The brightest 'pinpricks' (black arrows) represent clusters of up to 30 VLPs.

539   The black dot (shown in the white circle) on the images is due to a scratch on the microscope

540   lens. (a) Contamination on filter, which gives an idea of the size and appearance of a

541   bacterium in comparison with VLPs. (b) Negative control, 0.1 μm-filtered, autoclaved H_2O.

542   (c) Negative control, sterile 0.5 % 'Lab-Lemco'/6 % NaCl. (d) Donor 1, 0.22 μm-filtered

543   sample; (e) Donor 1 0.45 μm-filtered sample. (f) Donor 2, 0.22 μm-filtered sample; (g)

544   Donor 2, 0.45 μm-filtered sample. Scale bar, 10 μm.

545

545   **Fig. 3.** Transmission electron micrographs revealing the diversity of VLPs found in faeces.

546   (a) Donor 1 (female, 41 years) and (b) Donor 2 (female, 36 years) after extraction of VLPs in

547   TBT buffer.

548

548   **Fig. 4.** Transmission electron micrographs of VLPs found in faeces of Donor 2 (female, 36

549   years) after extraction of VLPs in TBT buffer showing *Myoviridae* phages with radial

550   whiskers attached to the capsids (top) and clusters of *Myoviridae* phages adsorbing to

551   membrane vesicle material (bottom).

552

552 **Fig. 5.** Transmission electron micrographs revealing the diversity of VLPs found in caecal

553 effluents. Sample (a) L08, (b) L10, (c) L16 and (d) L17 after extraction of VLPs in LL. No

554 clinical data were available for these individuals.

555

555   **Fig. 6.** Use of PFGE to examine VLP populations in PEG-precipitated FFs (faeces), and

556   demonstration that increasing the temperature at which samples are heated prior to loading

557   onto the gel can affect the diversity uncovered. The gel was stained with SYBR Gold and

558   visualized as described in Methods. This is a crude (but inexpensive) method of examining

559   VLP populations in human faeces, and demonstrates that each individual harbours a unique

560   VLP profile. Ladder, lambda ladder (#340; New England Biolabs).

561

561   **Fig. 7.** Isolation of high-quality DNA from CsCl-purified samples of human gastrointestinal

562   VLPs. VLP preparations were CsCl-purified and DNA extracted as described in Methods.

563   Aliquots (10 μl) of sample were run on a 0.8 % agarose gel at 90 V for 30 min. Lane 1,

564   molecular size ladder; lane 2, ϕKPLN1 positive control (a bacteriophage isolated from the

565   human caecum that infects *Klebsiella pneumoniae* subsp. *pneumoniae* K2 strains; L. Hoyles,

566   unpublished data), 63 ng DNA/μl; lane 3, VLP extract from faeces of Donor 4, 3.1 ng

567   DNA/μl; lane 4, VLP extract from faeces of Donor 3, 4.5 ng DNA/μl; lane 5, VLP extract

568   from caecal sample L07 (healthy), 0.2 ng DNA/μl; lane 6, VLP extract from caecal sample

569   L08 (no clinical data), 2.7 ng DNA/μl; lane 7, VLP extract from caecal sample L10 (no

570   clinical data), 3.8 ng DNA/μl. The two black arrows highlight the presence of RNA in the

571   sample from Donor 3, confirmed by treating the CsCl-purified sample with 20 μl of 1 mg/mL

572   RNAse A (bovine pancreas, ≥70 Kunitz/mg protein, prepared in 0.1 μm-filtered, sterile H_2O)

573   with the DNAse in a second DNA extraction (not shown). VLPs and DNA were extracted

574   from human caecum samples (1:4 LL, v/v) as described in Methods as part of a study

575   examining the microbiota associated with the human caecum (L. Hoyles, unpublished data).

576   With the exception of the sample in lane 5, DNA visible to the naked eye was extracted from

577   all samples of gastrointestinal origin.

578