



# Genome Characterization of a Novel Wastewater *Bacteroides fragilis* Bacteriophage (vB\_BfrS\_23) and its Host GB124

Mohammad A. Tariq<sup>1†</sup>, Fiona Newberry<sup>1,2†</sup>, Rik Haagmans<sup>1,2</sup>, Catherine Booth<sup>1</sup>, Tom Wileman<sup>1,2</sup>, Lesley Hoyles<sup>3</sup>, Martha R. J. Clokie<sup>4</sup>, James Ebdon<sup>5</sup> and Simon R. Carding<sup>1,2\*</sup>

<sup>1</sup> Gut Microbes and Health Research Programme, Quadram Institute Biosciences, Norwich Research Park, Norwich, United Kingdom, <sup>2</sup> Norwich Medical School, University of East Anglia, Norwich, United Kingdom, <sup>3</sup> Department of Biosciences, Nottingham Trent University, Nottingham, United Kingdom, <sup>4</sup> Department of Genetics and Genome Biology, Leicester University, Leicester, United Kingdom, <sup>5</sup> Environment and Public Health Research Group, School of Environment and Technology, University of Brighton, Brighton, United Kingdom

## OPEN ACCESS

### Edited by:

Andrew S. Lang,  
Memorial University of Newfoundland,  
Canada

### Reviewed by:

Alexander P. Hynes,  
McMaster University, Canada  
Ananda Shankar Bhattacharjee,  
Carl R. Woese Institute for Genomic  
Biology, University of Illinois  
at Urbana-Champaign, United States

### \*Correspondence:

Simon R. Carding  
Simon.Carding@quadram.ac.uk

†These authors share first authorship

### Specialty section:

This article was submitted to  
Virology,  
a section of the journal  
Frontiers in Microbiology

Received: 14 July 2020

Accepted: 05 October 2020

Published: 23 October 2020

### Citation:

Tariq MA, Newberry F,  
Haagmans R, Booth C, Wileman T,  
Hoyles L, Clokie MRJ, Ebdon J and  
Carding SR (2020) Genome  
Characterization of a Novel  
Wastewater *Bacteroides fragilis*  
Bacteriophage (vB\_BfrS\_23) and its  
Host GB124.  
Front. Microbiol. 11:583378.  
doi: 10.3389/fmicb.2020.583378

*Bacteroides* spp. are part of the human intestinal microbiota but can under some circumstances become clinical pathogens. Phages are a potentially valuable therapeutic treatment option for many pathogens, but phage therapy for pathogenic *Bacteroides* spp. including *Bacteroides fragilis* is currently limited to three genome-sequenced phages. Here we describe the isolation from sewage wastewater and genome of a lytic phage, vB\_BfrS\_23, that infects and kills *B. fragilis* strain GB124. Transmission electron microscopy identified this phage as a member of the *Siphoviridae* family. The phage is stable when held at temperatures of 4 and 60°C for 1 h. It has a very narrow host range, only infecting one host from a panel of *B. fragilis* strains ( $n = 8$ ). Whole-genome sequence analyses of vB\_BfrS\_23 determined it is double-stranded DNA phage and is circularly permuted, with a genome of 48,011 bp. The genome encodes 73 putative open reading frames. We also sequenced the host bacterium, *B. fragilis* GB124 (5.1 Mb), which has two plasmids of 43,923 and 4,138 bp. Although this phage is host specific, its isolation together with the detailed characterization of the host *B. fragilis* GB124 featured in this study represent a useful starting point from which to facilitate the future development of highly specific therapeutic agents. Furthermore, the phage could be a novel tool in determining water (and water reuse) treatment efficacy, and for identifying human fecal transmission pathways within contaminated environmental waters and foodstuffs.

**Keywords:** bacteriophage, *Bacteroides fragilis*, genomes, waste water, human

## INTRODUCTION

Bacteriophages (phage) are viruses that infect bacterial cells and as a result can influence their growth, fitness and response to stress (Casjens and Hendrix, 2015; Davies et al., 2016). They are estimated to be numerically the most abundant biological entity on earth numbering at least  $10^{31}$  (Hendrix et al., 1999; Bar-On et al., 2018). Phages are also a major constituent of the human

**Abbreviations:** BPRM, *Bacteroides* phage recovery medium; CDS, coding sequence; ORF, open reading frame.

microbiome and in particular, the intestinal microbiota where they can outnumber bacterial cells and human cells by up to 10:1 (Sender et al., 2016).

*Bacteroides* spp. are a dominant component of the intestinal bacteriome, accounting for between 5 and 40% of all anaerobes (Gorvitovskaia et al., 2016). In a recent study looking at 98 gut samples the relative abundance of *Bacteroides* spp. ranged from 0.37 to 98.82% (King et al., 2019). Although *Bacteroides fragilis* represents a smaller fraction of this group, it was present in all samples (King et al., 2019). *Bacteroides* spp. confer significant health benefits to their host including the digestion, processing and extraction of nutrients from complex plant-based polysaccharides, promoting colonic motility and angiogenesis, and the development of the gut-associated immune system (Hooper et al., 2003; Xu et al., 2003; Mazmanian et al., 2005).

*Bacteroides* spp. are also important clinical pathogens and can contribute to anaerobic infections (Perez-Brocal et al., 2015). *B. fragilis* is one of the most commonly isolated anaerobic pathogens from soft tissue infections and bacteremia (Shenoy et al., 2017). The capsular polysaccharide complex of *B. fragilis* consisting of two distinct polysaccharides is the primary mediator of intra-abdominal abscess formation (Tzianabos et al., 1993). Enterotoxigenic *B. fragilis* also produces metalloprotease toxins (fragilysin), which unlike pore-forming toxins, breakdown connective tissue through proinflammatory cytokine signaling leading to an increase in the permeability of the epithelial barrier, causing diarrheal diseases and acute inflammation (Wu et al., 2004; Kim et al., 2006).

Anaerobic bacteria and their phages have been proposed as candidates for indicators of fecal pollution as they do not replicate in estuarine waters and the phages are resistant to chlorine inactivation (Booth et al., 1979; Tartera and Jofre, 1987; McMinn et al., 2014, 2017; Dias et al., 2018). Despite the clinical importance of *B. fragilis*, and a study reporting up to  $3 \times 10^4$  PFU/100 mL of *B. fragilis* phages in sewage influent (Sun et al., 1997), only three complete genomes of *B. fragilis* phages have been described to date, two of which have been published with the third deposited under accession MN078104; all are virulent phages (Puig and Girones, 1999; Hawkins et al., 2008; Ogilvie et al., 2012). Here we describe the isolation of a new *B. fragilis* phage, vB\_BfrS\_23 from municipal wastewater and detail its genome characteristics. We also sequenced the host bacterium *B. fragilis* strain GB124 isolated from a United Kingdom municipal wastewater sample (Payan et al., 2005).

## MATERIALS AND METHODS

### Bacterial Culture and Growth Conditions

*Bacteroides fragilis* GB124 was used as the host reference strain for phage isolation and has been used to detect human fecal contamination in water sources (Payan et al., 2005; Ebdon et al., 2007), and to test the treatment efficacy of water reuse technologies (Purnell et al., 2015, 2016;

Dias et al., 2018). *Bacteroides* phage recovery medium (BPRM) was used to cultivate host GB124 and propagate the phage (**Supplementary Table 1**).

### Phage Isolation and Purification

Phage vB\_BfrS\_23 was isolated from 100 mL raw (untreated) municipal wastewater from a United Kingdom-based treatment plant. Wastewater was filtered with a 0.45  $\mu\text{m}$  PES membrane syringe filter (Sartorius UK Ltd.) and concentrated by centrifugation for 15 min at 5,000  $g$  using Amicon Ultra-15 10K centrifugal filter units. One milliliter of this concentrated sewage filtrate was mixed with 1 mL of mid-exponential growth phase ( $\text{OD}_{620\text{nm}}$  0.3–0.4) *B. fragilis* GB124 allowing 5 min for adsorption and was then added to semi-soft BPRM agar (0.35%) and poured on BPRM agar (1.5%) (**Supplementary Table 1**; Ebdon et al., 2007). After 18 h anaerobic incubation (5%  $\text{CO}_2$ , 5%  $\text{H}_2$  and 90%  $\text{N}$  at 37°C and  $\sim 25$  psi pressure) the plates were screened for plaques. A single plaque was picked using a sterile pipette and resuspended in 10 mL BPRM medium containing sub-cultured host ( $\text{OD}_{620\text{nm}}$  0.3–0.4). The suspension was incubated for 18 h to allow further propagation of the phages. The sample was filtered through a 0.22  $\mu\text{m}$  PES membrane filter (Sartorius UK Ltd.). The procedure was repeated a further three times to obtain a pure phage stock. This stock was used to further propagate and increase the phage titer. Fifty microliter was used to perform serial dilutions and was added to semi-soft BPRM agar (0.35%) with 200  $\mu\text{L}$  of mid-log phase ( $\text{OD}_{620\text{nm}}$  0.3–0.4) bacterial host. The plates were incubated for 16 h in an anaerobic cabinet (5%  $\text{CO}_2$ , 5%  $\text{H}_2$  and 90%  $\text{N}$  at 37°C and  $\sim 25$  psi pressure). Five milliliter SM buffer (100 mM NaCl, 8.1 mM  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$  and 50 mM Tris.HCl pH 7.4) was added to a plate of complete cell lysis and left on a mini gyratory shaker SSM3 (Stuart, United Kingdom) for 1 h. The top agar was harvested along with the buffer and transferred to a 50 mL tube (Corning, United Kingdom), after a brief vortex the tube was centrifuged at  $3,000 \times g$  for 10 min and the supernatant filtered through a PES membrane bottle top vacuum filter using  $\sim 100$  psi pressure (Millipore Millivac, Merck UK). The titer was evaluated using dilutions  $10^{-1}$  to  $10^{-9}$  and the titer adjusted to  $1 \times 10^8$  PFU/mL for temperature assays and was stored at 4°C for the duration of the experiments.

### Transmission Electron Microscopy

Briefly, a small drop of phage suspension containing  $\sim 1 \times 10^7$  PFU/mL was applied to a formvar/carbon coated copper transmission electron microscopy (TEM) grid (Agar Scientific, Stansted, United Kingdom) and left for 1 min. Excess liquid was removed with filter paper. A small drop of 2% uranyl acetate (BDH 10288) was applied to the grid surface and left for a further 1 min after which it was removed with filter paper. Grids were left to thoroughly dry before viewing and imaging using a Talos F200c TEM with Gatan Oneview digital camera.

### Host Range Assay

In total, eight *B. fragilis* strains (**Supplementary Table 2**) were used to determine the host range and specificity of vB\_BfrS\_23. Bacterial strains were cultured in BPRM broth to exponential

phase (OD<sub>620</sub> 0.3–0.33) prior to incorporation into double BPRM agar overlays (Ebdon et al., 2007). Dilutions of vB\_BfrS\_23 were then spotted onto the double agar overlay and observed for plaques following 16 h in an anaerobic cabinet (5% CO<sub>2</sub>, 5% H<sub>2</sub> and 90% N at 37°C and ~25 psi pressure).

## One-Step Growth Curve and Eclipse Period

To determine the burst size and latency period of vB\_BfrS\_23, a one-step growth curve was carried out (Kropinski, 2018). Initially, 9.9 mL of *B. fragilis* GB124 was grown anaerobically (5% CO<sub>2</sub>, 5% H<sub>2</sub> and 90% N at 37°C and ~25 psi pressure) to mid-exponential phase and 0.5 OD<sub>620</sub>. One-hundred µL of 1 × 10<sup>7</sup> PFU/mL phage was then added for 5 min to allow phage adsorption. A 0.1 mL aliquot was then used to make ten-fold serial dilutions to a final dilution of 1 × 10<sup>1</sup>. As an adsorption control, a 1 mL aliquot from the 1 × 10<sup>3</sup> dilution flask aliquot was added to 50 µL of CHCl<sub>3</sub> and kept on ice for the duration of the experiment (less than 4 h). At various time points 0.1 mL was taken from each dilution and mixed with 200 µL of bacterial host suspension (in BPRM broth) and plated using 0.35% (w/v) BPRM agar. The data were normalized by multiplying the adsorption control and the value obtained from 1 × 10<sup>3</sup> PFU/mL flask by x10, 1 × 10<sup>2</sup> PFU/mL flask by x100 and 1 × 10<sup>1</sup> PFU/mL flask by x1000. The burst size was determined as previously described (Kropinski, 2018).

At each sampling point 475 µL was taken to determine the eclipse period. The sample was added to 25 µL of chloroform (5%v/v), vortexed for 10 s and kept on ice until the end of the experiment to allow the chloroform to settle. One hundred microliter was taken from each timepoint sample and added to 200 µL of bacterial host and plated using 0.35% (w/v) BPRM agar. The plates were incubated for 16 h in an anaerobic cabinet (5% CO<sub>2</sub>, 5% H<sub>2</sub> and 90% N at 37°C and ~25 psi pressure). Each one-step growth and eclipse experiment were repeated to give three biological replicates.

## Thermal Assay

The viability of vB\_BfrS\_23 at different temperatures was assessed by incubating phage preparations of known titers at 4, 24, 30, 37, 40, 45, 60, 70, or 80°C for 15, 30, or 60 min out of direct sunlight. Serial dilutions of the phage stocks were incubated with 200 µL of bacterial host culture in BPRM broth for 15 min at 37°C prior to mixing with 5 mL BPRM semi-soft agar (0.35%, w/v) and pouring onto BPRM agar plates and incubated for 18 h at 37°C in anaerobic cabinet (5% CO<sub>2</sub>, 5% H<sub>2</sub> and 90% N at 37°C and ~25 psi pressure). For accuracy, plaques were counted on plates containing between 30 and 300 plaques.

## DNA Extraction

For Illumina sequencing, phage preparations (~10<sup>9</sup> PFU/mL) were incubated with RNase A (100 U Ambion) and Turbo DNase (2U Invitrogen) at 37°C for 30 min to remove bacterial chromosomal DNA. Nucleases were heat-inactivated at 65°C in the presence of 15 mM EDTA for 10min. The Norgen Phage DNA isolation kit (Geneflow Limited, Lichfield, United Kingdom) was

used to extract the phage DNA. For Nanopore sequencing, phage was PEG-precipitated (10% (w/v) PEG 8000 and 6% (w/v) NaCl), treated with DNase (4 U Turbo DNase; Invitrogen) and RNase A (100 U; Ambion) followed by treatment with SDS (0.5%, w/v) and 4 µL (80 µg of proteinase K 20 mg/mL, Ambion) treatment at 55°C for 1 h and heat inactivation at 75°C for 15 min. Lipids and proteins were removed by mixing the sample 1:1 with chloroform and vigorous shaking for a few seconds followed by centrifugation at 15,000 g at 20°C for 5 min. The upper aqueous phase was treated with NaCl (0.2 M final concentration) prior to mixing 1:1 with isopropanol and left in -20°C for 16 h. The sample was centrifuged at 13,000 g at 20°C for 1 h followed by two washes with 70% ethanol prior to resuspending the pellet in nuclease-free water (Invitrogen). Bacterial DNA was extracted from an overnight culture of GB124 grown in BPRM broth, the sample was centrifuged at 3,000 g for 20 min, the pellet was resuspended using 300 µL of TE buffer in accordance to the Promega Maxwell® RSC Cultured Cell DNA Kit (AS1620) protocol (FB211) and run on the Promega Maxwell® RSC Instrument (AS4500).

## DNA Sequencing

Phage and bacterial genomic DNA were sequenced using Illumina and MinION ONT sequencing platforms. For MinION sequencing, the standard ONT protocol and native barcoding kit EXP-NBD104 with the ligation sequencing kit SQK-LSK109 were used. In brief, 1 µg of high quality vB\_BfrS\_23 and *B. fragilis* GB124 DNA was end-repaired and dA-tailed using the NEBNext FFPE Repair Mix (M6630) and NEBNext End Repair/dA-tailing (E7546). The native barcode (EXP-NBD104 kit) was used to barcode the samples and they were ligated using NEB Blunt/TA Ligase Master Mix (M0367). The sequence adapters were ligated with NEBNext Quick Ligation Module (E6056) and the samples were primed and loaded using the Flow Cell Priming Kit (EXP-FLP001) on the MinION R9 4.1 FLO-MIN106. Samples were run for 72 h, and the raw reads were base-called using Guppy v3.5.1.<sup>1</sup> Adapters were removed using Porechop v0.2.3 (rrwick/Porechop, 2020).<sup>2</sup> Genomic bacterial DNA was also sequenced using the Illumina MiSeq system. Briefly, the Illumina Nextera XT (Illumina, Saffron Walden, United Kingdom) library preparation kit was used to prepare sequencing libraries prior to running on Illumina MiSeq 2 × 150-cycle v2 chemistry. Paired-end sequencing reads were provided as FASTQ files with the raw reads having their adapters removed using Trimmomatic (Bolger et al., 2014) prior to quality trimming using Sickle at -q 30 and -l 15 (Joshi and Fass, 2011).

## Phage Genome Assembly and Annotation

Illumina MiSeq and MinION generated sequences were assembled using Unicycler v0.4.8 (Wick et al., 2017) resulting in a single contiguous circular sequence of 48,011 bp. The genome was annotated using RAST (Aziz et al., 2008; Overbeek et al., 2014; Brettin et al., 2015). The putative functions of the coding

<sup>1</sup><https://nanoporetech.com>

<sup>2</sup><https://github.com/rrwick/Porechop>



regions (CDS) were predicted using NCBI-nr (June 15, 2020) and Conserved Domain Database (CDD) (June 15, 2020) searches using BlastP and tBlastn. For BlastP and tBlastn, hits were considered significant if the e-values were lower than  $1 \times 10^{-5}$  at  $\geq 60\%$  protein identity (Altschul et al., 1990). For CDD searches, only hits with an e-value of 0.01 or lower were considered significant (Marchler-Bauer et al., 2015).

## GB124 Genome Assembly, Quality Checks and Annotation

The genome was assembled using Illumina MiSeq and ONT MinION reads, using Unicycler (Wick et al., 2017). Following host assembly, the contig was annotated using Prokka v.1.14.6 (Seemann, 2014). Antimicrobial resistance genes were investigated with ABRicate v.0.9.8 (tseemann/abricate, 2020)<sup>3</sup> using Resfinder v3.2 (database version September 10, 2019) (Zankari et al., 2012), NCBI and AMRFinderPlus v3.8 (database version 2020-05-04.1) (Feldgarden et al., 2020). Insertion elements were predicted using ISfinder (Siguiet et al., 2006).<sup>4</sup> Significant hits (Score > 100 and e value <  $4 \times 10^{-11}$ ) in ISfinder were examined in the Prokka GenBank file and protein sequence submitted to BlastP. Suspected insertion sequence (IS) elements were visualized in Artemis 18.1.0 (Carver et al., 2012) and investigated for downstream Anti-Microbial Resistance (antimicrobial resistance) genes. ABRicate hits were considered significant if the coverage and identity were >90%.

Plasmids were identified using the PLSDB web server<sup>5</sup> (data v2020\_03\_04) and coding regions found using Prokka v1.14.6 (Seemann, 2014; Galata et al., 2019). The putative functions assigned by Prokka were checked using BlastP according to NCBI-nr (July 1, 2020) and Conserved Domain Database (July 1, 2020) (Altschul et al., 1990; Marchler-Bauer et al., 2015). Hits were considered significant if the e-values were lower than  $1 \times 10^{-5}$  at  $\geq 80\%$  protein identity. Plasmids were visualized using Brig v0.95 (Alikhan et al., 2011). The plasmids were screened for antimicrobial resistance and virulence genes using ABRicate v.0.9.8 and Resfinder, NCBI AMRFinderPlus and VFDB (Zankari et al., 2012; Chen et al., 2016; Feldgarden et al., 2020). The completeness and contamination of the seven-contig assembly were assessed using CheckM v1.0.18 (Parks et al., 2015). Average nucleotide identity with the type strain of *B. fragilis* was assessed using fastANI v1.2 (Jain et al., 2018) and *B. fragilis* CCUG 4856<sup>T</sup> (RefSeq assembly accession GCF\_005706655).

## vB\_BfrS\_23 Phage Genome Comparison of Large Terminase Subunit and Tail Fiber

The coding region for the genes were tblastx searched using default parameter and amino acid sequences sharing identity to the large terminase subunit and the tail fiber sequences were aligned using MAFFT v7 (Katoh and Standley, 2013). The L-INS-i algorithm with default parameters was used to improve accuracy. The alignment file was used to create a p-distance

analysis in MEGA7 (Kumar et al., 2016) following construction of a neighbor-joining tree on p-distance using 1,000 bootstrap analyses using default parameters (Saitou and Nei, 1987).

## vB\_BfrS\_23 and $\phi$ B124-14 Linear Genome Comparison Alignment

A detailed comparison of the vB\_BfrS\_23 with  $\phi$ 124-14 (Ogilvie et al., 2012), B40-8 and *Bacteroides* phage Barc2635 was performed using tBLASTx in Easyfig (Sullivan et al., 2011). The annotated GenBank file of vB\_BfrS\_23 was compared with the GenBank file for  $\phi$ B124-14, B40-8 and *Bacteroides* phage Barc2635.

## RESULTS

### Phage Isolation and Phenotypic Characterization

*Bacteroides fragilis* GB124 was used as a host for phage discovery and the starting point for the screening of a filtered raw wastewater sample. We identified and isolated a virulent phage capable of infecting GB124, that generated plaques that ranged in size between 0.5 and 2 mm (Figure 1A). TEM images revealed the presence of a non-contractile long tail  $\sim 150$  nm and an icosahedral head  $\sim 50$  nm in size consistent with vB\_BfrS\_23 belonging to the *Siphoviridae* family of the order *Caudovirales* (Figure 1B).

### vB\_BfrS\_23 Phage Characteristics

Phage vB\_BfrS\_23 was seen to infect and lyse only one of the eight *B. fragilis* strains tested, GB124, which was the host strain used for isolating vB\_BfrS\_23 (Supplementary Table 2).

The one-step growth curve experiment (Figure 1C) showed that the phage had a burst size of  $\sim 44$  phage/cell (mean of three independent experiments) and latency period of  $\sim 37$  min. The eclipse period was determined to be  $\sim 23$  min ( $n = 3$ ). It was also stable at temperatures between 4°C and 60°C (Figure 1D) with viability decreasing at 60°C with a more rapid inactivation seen at 70°C. No plaques were seen at 80°C. At 37°C plaques were of sizes up to 2 mm (Supplementary Figure 7). Interestingly, a slight increase in PFU/mL was seen between 40°C and 45°C, with the plaques being more uniform and smaller (0.5 mm) at 45°C (Supplementary Figure 7).

### Phage Genome and Phylogeny

vB\_BfrS\_23 is a double-stranded DNA phage of 48,011 bp with a GC content of 38.6%, containing 73 putative CDS (Figure 2 and Supplementary Table 3). It was most similar to the virulent phage  $\phi$ B124-14 (86% query coverage) followed by *Bacteroides* phage Barc2635 (85% query coverage) and then  $\phi$ B40-8 (73% query coverage). The linear genome comparison of the vB\_BfrS\_23, Barc2635,  $\phi$ B124-14 and  $\phi$ B40-8 phages is illustrated in Figure 3.

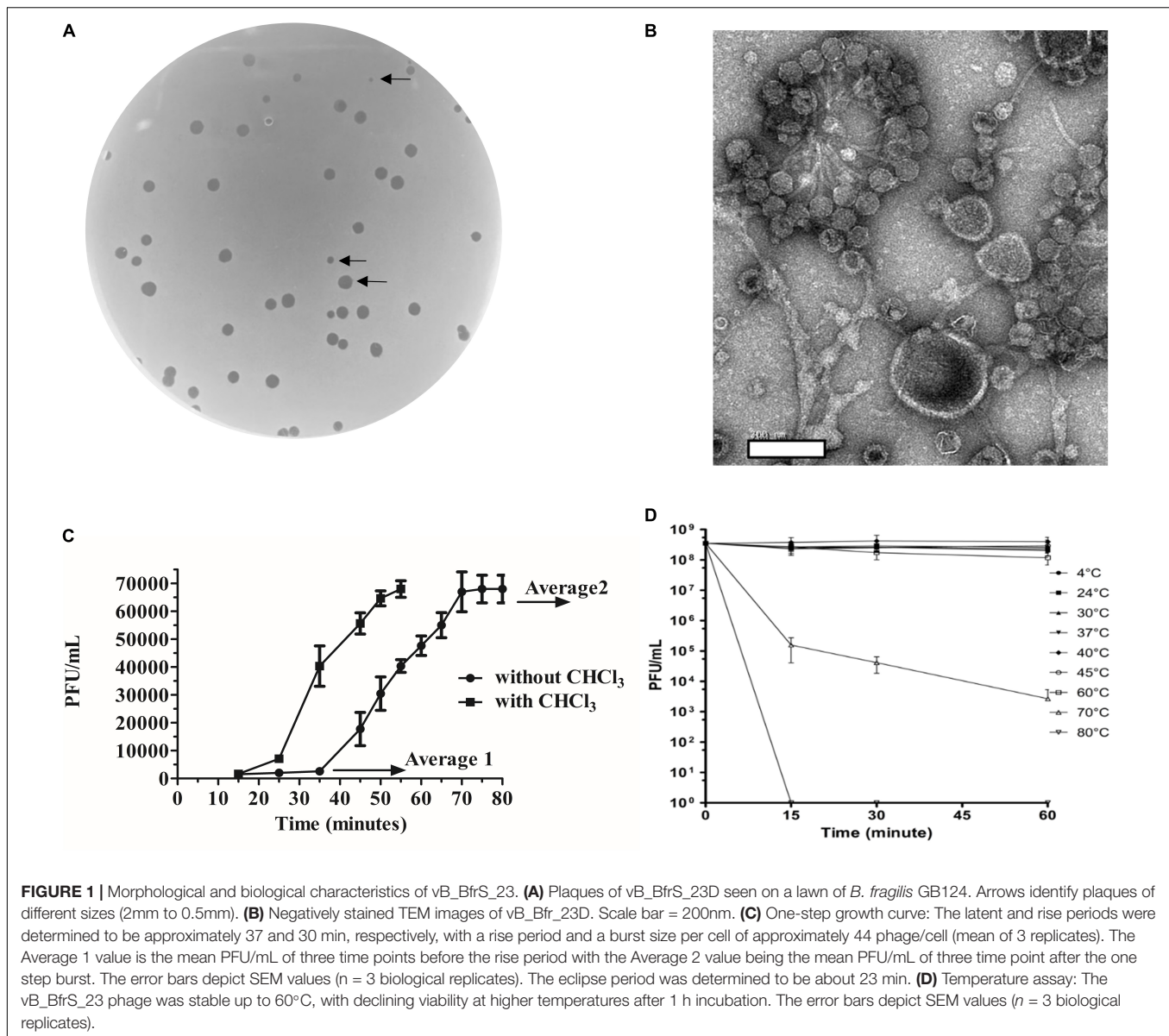
The terminase large subunit and the tail fiber proteins were used to generate a phylogenetic tree (Figure 4). Both the tail fiber (Figure 4A) and terminase large subunit (Figure 4B)

<sup>3</sup><https://github.com/tseemann/abricate>

<sup>4</sup><http://www-is.biotoul.fr>

<sup>5</sup><https://ccb-microbe.cs.uni-saarland.de/plsdb/>





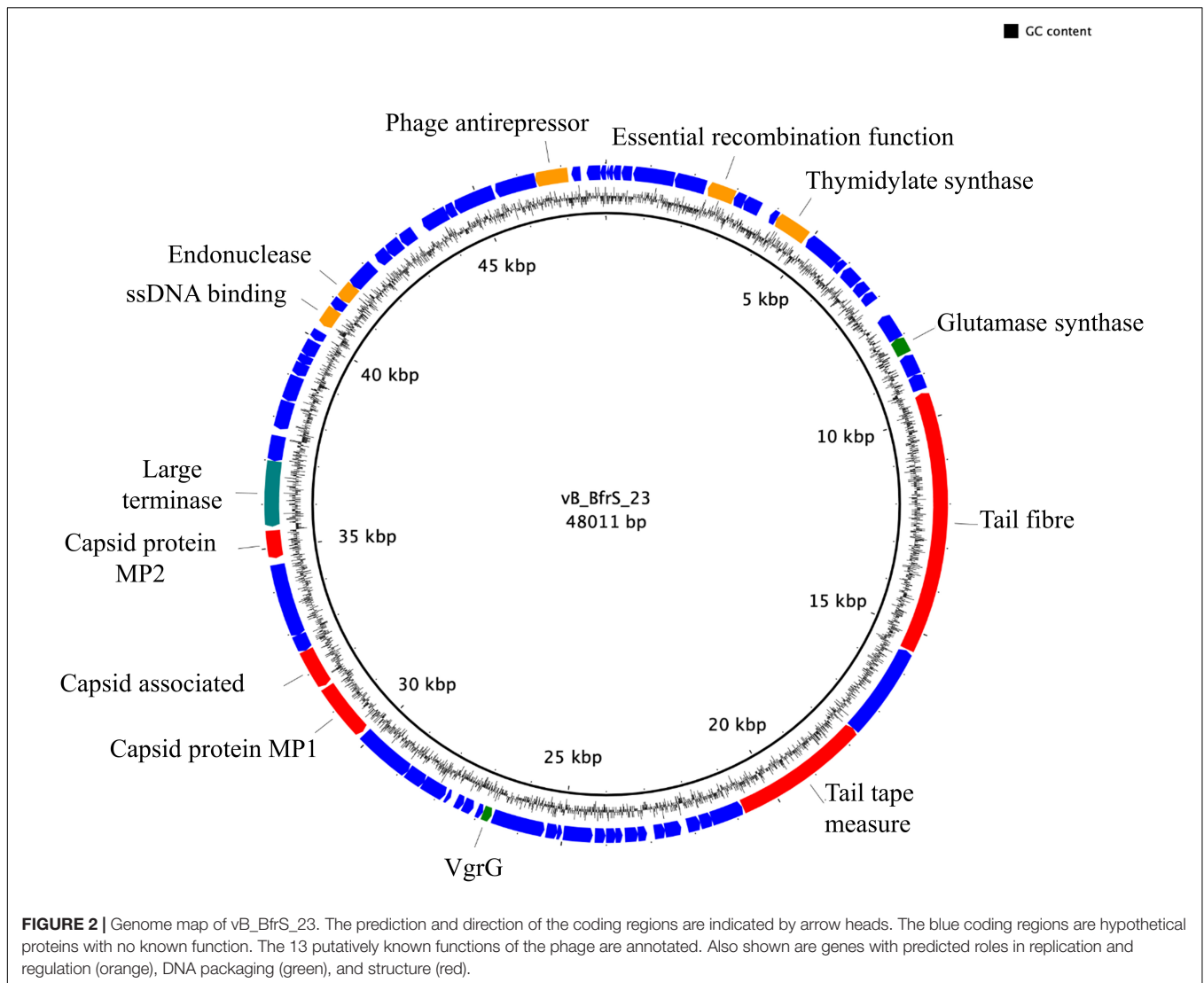
were shown to be most similar to  $\phi$ B124-14. BlastP<sup>6</sup> revealed 13 of the CDS had a putative function and 8 CDS contained conserved domain signatures. Most of the CDS were assignable to genome structure and replication/regulation, with the remainder associated with lysis and DNA structure. Putative CDS of similar function clustered together to form modules. However, 6 putative proteins identified in phage  $\phi$ B124-14 were not found within the vB\_BfrS\_23 genome (Ogilvie et al., 2012). Ten putative CDS showed no homology to any protein within the database, with 27 sharing the highest sequence similarity to genes in  $\phi$ B124-14, 27 to Barc2635 and 8 to  $\phi$ B40-8 (the following to prophage regions) and 1 to a *Bacteroides ovatus* phage (Supplementary Table 3).

Like  $\phi$ B124-14, Barc2635 and  $\phi$ B40-8, vB\_BfrS\_23 lacked an obvious virulent genome module and only contained 1

putative protein that alluded to a strictly lytic life cycle (CD 18). CD18 exhibited closest homology to a putative peptidase protein identified within  $\phi$ B124-14 and contained a peptidase superfamily domain. The peptidase protein appeared to reside within a cluster of unassigned proteins, suggesting it may be a putative virulent module.

Five CDs were assigned a predicted function relating to virus replication and regulation. CD11 encoded a putative thymidylate synthase, which is a key enzyme in the synthesis of 2'-deoxythymidine-5'-monophosphate, an essential precursor for DNA replication. A conserved domain region identified within the protein suggested it encodes a ThyA-like enzyme as reported in the  $\phi$ B124-14 genome (Ogilvie et al., 2012). CD7 (recombination protein) and CD70 (anti-repressor) were also encoded within the replication and regulation genome module, promoting transcription of phage genes (Lemire et al., 2011).

<sup>6</sup><https://blast.ncbi.nlm.nih.gov/Blast.cgi>

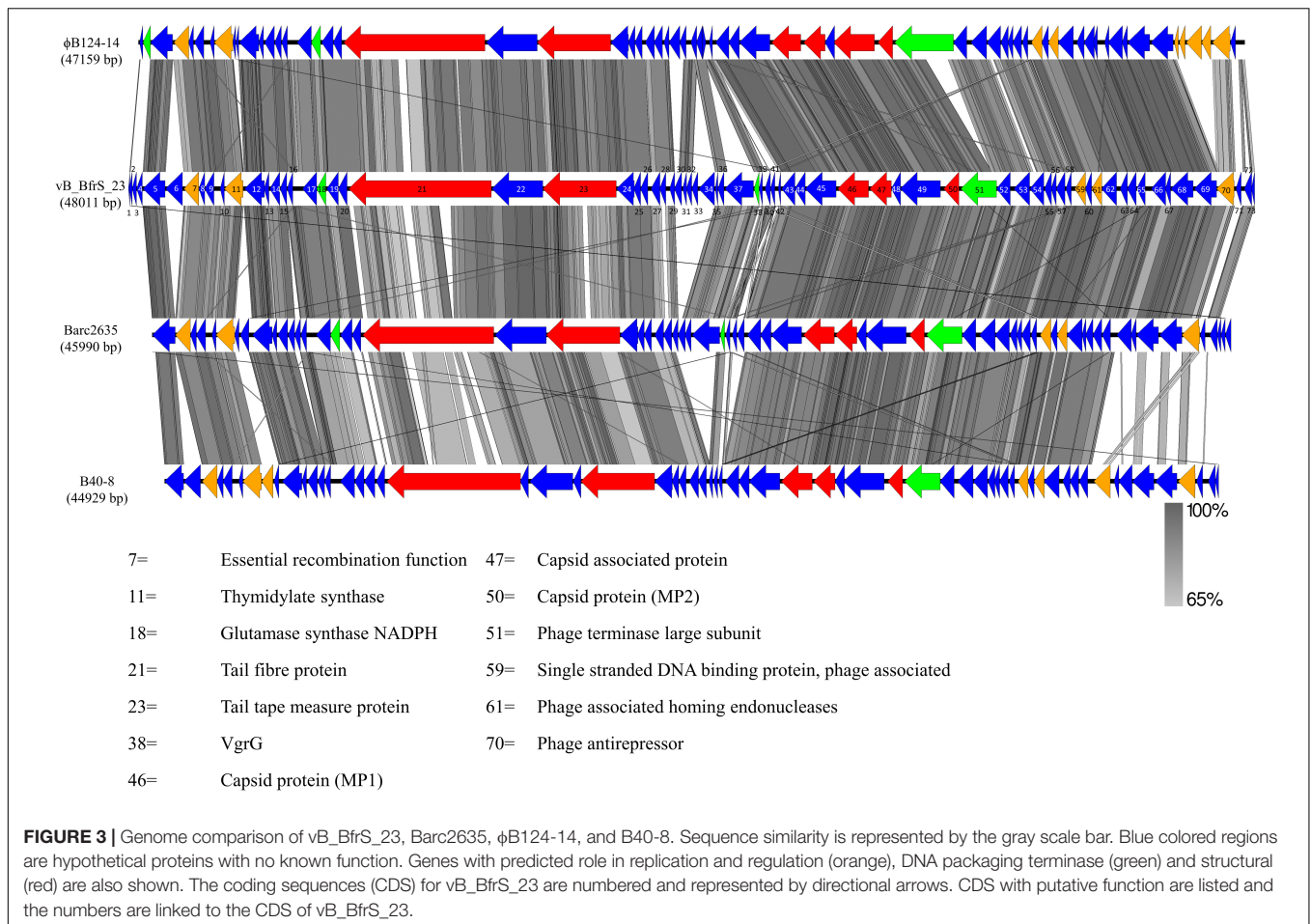


## *B. fragilis* GB124 Genome Assembly and Annotation

The genome was assembled into 7 contigs of >100 bp (N50 4,986,460 bp). CheckM analysis showed the genome to be 99.26% complete with no contamination. It shared 99.03% average nucleotide identity with *B. fragilis* CCUG 4856<sup>T</sup>, confirming GB124 as an authentic strain of *B. fragilis* (Chun et al., 2018). Two contigs were complete assemblies for plasmids which were identified using PLSDB and were named PBf1 and PBf2. PBf1 consisted of 4148 bp, was circular and contained eight predicted open reading frames (ORFs; **Figures 5A–C** and **Supplementary Figure 4**). PLSDB revealed an exact match to *B. xylanisolvans* strain H207 plasmid unnamed2 (NZ\_CP041232.1). A further Blastn search showed a 100% identity and query cover match to one other plasmid, *B. ovatus* strain 3725 D1 iv plasmid unnamed3. PBf1 contained a YoeB toxin, a toxic component of a type II toxin-antitoxin system that helps to maintain plasmid stability by post-segregation killing or genetic addition

(**Figure 5B** and **Supplementary Figure 5**) (Gerdes et al., 1986; Yarmolinsky, 1995). PBf2 consisted of 43,923 bp, was circular and contained 57 coding regions (5 domains of unknown function, 30 hypothetical proteins and 22 of putative function) (**Figure 5C**). A PLSDB search revealed 2 hits: *B. ovatus* strain 3725 D1 iv plasmid unnamed2 (NZ\_CP041397.1) and *B. thetaiotaomicron* F9-2 plasmid p1-F9 DNA (AP022661.1). Interestingly, an additional blastn search reported a 97% query cover and 98.2% percentage identify match to *B. xylanisolvans* strain H207 plasmid unnamed1. No virulence or antibiotic resistance genes were identified on either plasmid.

The remaining five contigs were identified as belonging to the GB124 genome of 5,093,249 bp with a GC content of 43.87%. A total of 4266 ORFs were predicted of which 72 were tRNA, 1 tmRNA and 2 CRISPR repeat regions (**Supplementary Table 6**). Two incomplete prophage regions were identified using PHASTER (Arndt et al., 2016). They were 42.5 and 13.2 kbp in size, all features are depicted in **Figure 5A**. A total of 5 IS elements were identified. However, investigation of the GenBank file and



protein sequences revealed only 1 IS element, IS1182 family ISBf3. Genes flanking the IS element had no known function. No *B. fragilis* virulence factors were identified. Resfinder, NCBI and AMRFinderPlus databases revealed two antimicrobial resistance genes (*cepA* and *tetQ*). The *cepA* gene, observed in >90% of *B. fragilis* isolates, encodes the  $\beta$ -lactamase protein and confers resistance to cephalosporins (except cefoxitin) and penicillin (Rogers et al., 1993; Mastrantonio et al., 1996). *tetQ* gene-related resistance is common among *B. fragilis* isolates and encodes a protein that protects the bacterial ribosomes from tetracycline (Rasmussen et al., 1993; Roberts, 1996). Bacterial contigs >200 bp were submitted to GenBank thus omitting 162bp contig from the assembly.

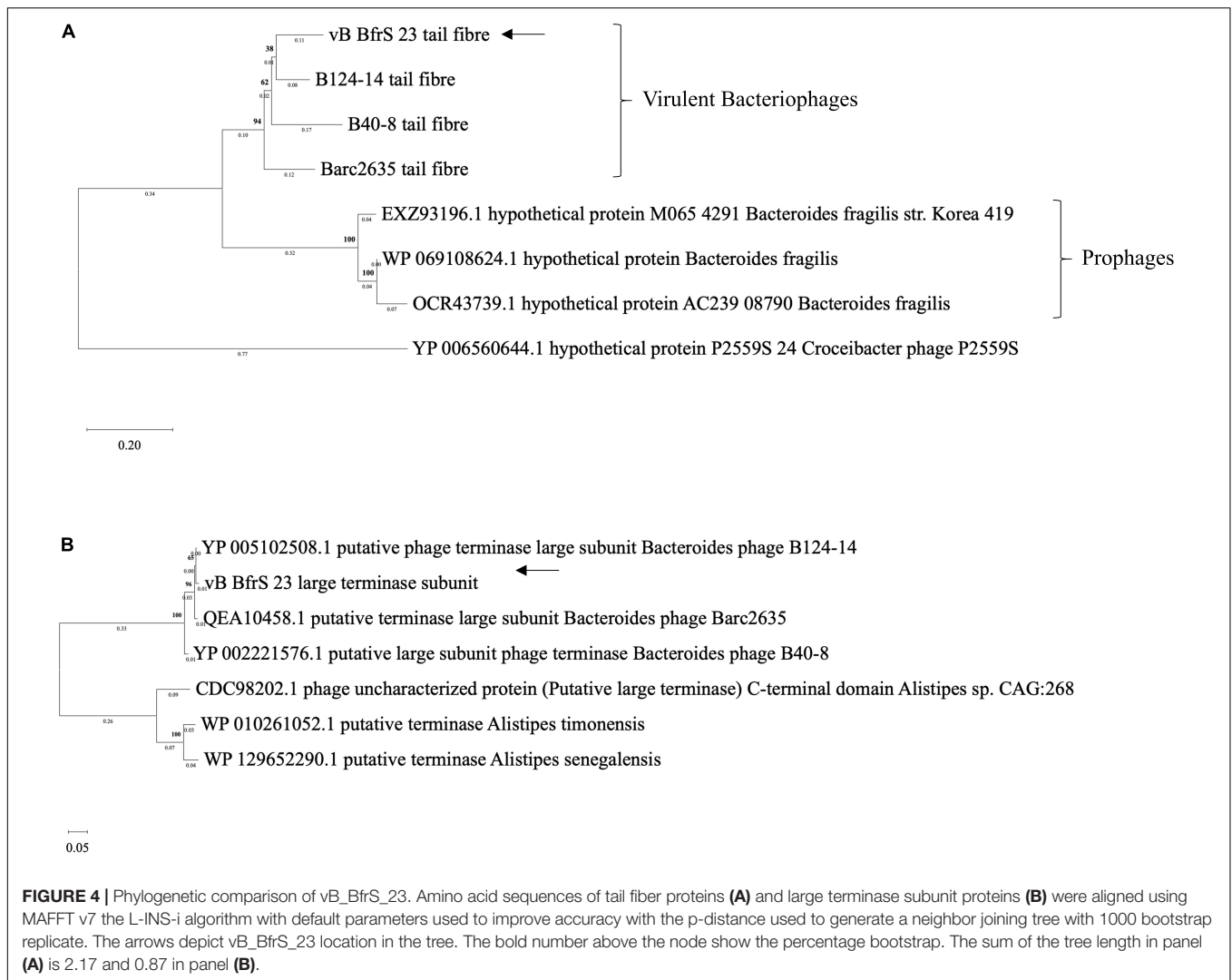
## DISCUSSION

The escape of *B. fragilis* from the gut environment into other parts of the body can result in major pathology, including bacteremia and abscess formation in various tissues. Although phages are a therapeutic option to treat and kill pathogenic *B. fragilis* strains, only three to date had been described and their genomes sequenced. Here, we identify a new highly specific virulent phage that is only able to infect a single host among a panel

of *B. fragilis* strains tested. This supports similar finding of phi B124-14 which infected 5 out of 15 *B. fragilis* spp. tested (Ogilvie et al., 2012). This narrow host range may reflect extreme niche specialization exhibited by close phylogenetic and evolutionary relationships to gut bacteria (Zaneveld et al., 2010; Ogilvie et al., 2012). The morphological features of the phage identify it as *Siphoviridae*. The absence of any identifiable genes essential for the lysogenic life cycle is consistent with vB\_BfrS\_23 being a virulent phage. Despite identifying recombinase (CD7) and anti-repressor genes (CD70) which are associated with temperate life cycles, no integrase or excision genes that are essential for lysogenic life cycle were identified. Recombinase and anti-repressor genes have also been identified in phiB124-14, B40-8 and Barc2635 (Figure 3). The investigators that initially described phiB124-14 concluded that it was a virulent phage based upon a deviation in GC content between the phage and host, as we have seen between vB\_BfrS\_23 (38.6% GC content) and *B. fragilis* GB124 (43.87% GC content) (Deschavanne et al., 2010; Ogilvie et al., 2012). Thus, we assume that vB\_BfrS23 is a virulent phage that may be a model candidate for human-specific microbial source tracking in contaminated surface and groundwater.

The phage vB\_BfrS\_23 contains a putative peptidase protein, but lacks any homology to known holin proteins (small



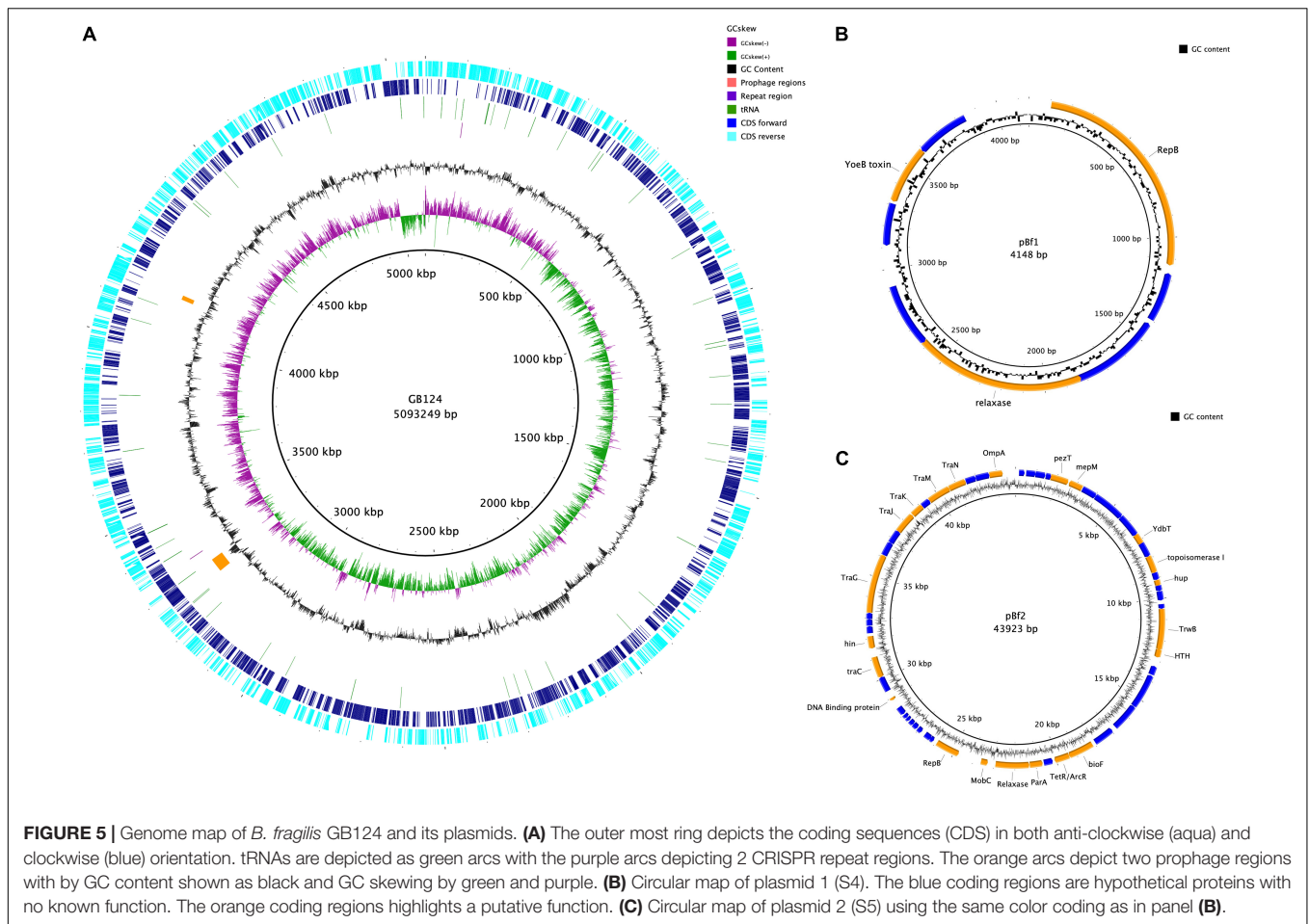


membrane proteins) which is not unusual for phages belonging to the *Siphoviridae* family (Hawkins et al., 2008; Duhaime et al., 2011). Double-stranded DNA phages typically lyse host cells using a holin-endolysin system. Active degradation of bacterial peptidoglycan is achieved with a muralytic enzyme or endolysin (Young, 1992; Young and Blasi, 1995). Endolysins accumulate in an active state in the cytosol, the holin proteins bind to the membrane, and the membrane is permeabilized to the endolysin. This leads to the breakdown of murein's resulting in the cell bursting. All this is time dependent and is programmed into the holin gene (Wang et al., 2000). It appears that the putative peptidase protein resides within an undefined lytic life cycle module in which there may be a holin-endolysin system. Interestingly, a putative thymidylate synthase was identified (CDS11) within the replication and regulation gene module. It is highly conserved across bacterial and mammalian species and shares remarkable structural and functional similarities (Carreras and Santi, 1995; Escartin et al., 2008). The exact function of ThyA within the phage genome is unknown but its additional copies may be of importance for survival of its host by enhancing

growth (Stern et al., 2010). No tRNA genes were identified. The genome map highlights only 13 of the 73 predicted coding regions with a putative function, emphasizing the fact that phages are under-characterized.

In comparing the genome of vB\_BfrS\_23 with that of  $\phi$ B124-14, the former is 852 bp larger. Both genomes have genes unique to them that are primarily located around the same gene module and near the cos site, possibly due to recombination events. The Barc2635 genome is 2,021 bp smaller than vB\_BfrS\_23. There are distinct putative genes present in vB\_BfrS\_23 that are missing in Barc2635 including CDS 6, 31-34 and 71 hypothetical proteins. The tail fiber protein (CDS 21) is also smaller in Barc2635, B40-8 and  $\phi$ B124-14 compared to vB\_BfrS\_23 by 40 – 483bp (Figure 3). The relatively large tail fiber is consistent with that is seen for *Bacteroides fragilis* phages.

The large terminase subunit and tail fiber phylogenetic comparison shows these genes share homology with other known *B. fragilis* phages. Furthermore, the large terminase subunit is smaller in vB\_BfrS\_23 (CDS 51) and Barc2635 compared to  $\phi$ B124-14. Interestingly the lytic tail fiber genes share higher



nucleotide identity with other virulent *B. fragilis* phages compared to the prophage tail fibers, as they appear on another clade. This may be due to the life cycle of the phages and may be an indicator of differences in host range of virulent versus temperate phages. The large terminase subunit shares more identity to  $\phi$ B124-14, Barc2635 and B40-8 *B. fragilis* phages (Figure 4A).

The thermal assays provided useful information about vB\_BfrS\_23. The phage was resilient to the lower temperatures tested (4, 24, 30, 37°C) and at 40 and 45°C an increase in the number of plaques was observed. A key observation is that the phage forms aggregates (Figure 1B) and at higher temperatures ( $\leq 40$  and 45°C) the plaque size was consistently 0.5 mm (Supplementary Figure 7). This characteristic of *B. fragilis* phages showing different plaques sizes and the phage agglomeration or clumping has been reported previously (Keller and Traub, 1974) for *B. fragilis* phages isolated from animal sera (Keller and Traub, 1974). It is possible therefore that the temperature kinetics impacts phage clustering by influencing the binding between phages, potentially mediated by Ig-like or carbohydrate adherence domains (CAD; Sathaliyawa et al., 2010; Barr et al., 2013; Tariq et al., 2015) enabling a more accurate determination of phage numbers in the sample. Alternatively, this could be an artifact of the experimental procedure used to generate samples for EM imaging including the duration

of vortexing, the media and need for high phage titers. The vB\_BfrS\_23 thermal stability characteristics supports work done previously with naturally occurring GB124 phages (Bertrand et al., 2012; McMinn et al., 2014). The first of these studies showed the phages persisted longer at 5°C compared to higher temperatures (20 and 35°C), where few plaques were observed after 7 days (McMinn et al., 2014). Similarly, a meta-analysis of virus inactivation found inactivation occurred faster at temperatures  $\geq 50^\circ\text{C}$  than at  $< 50^\circ\text{C}$ . Adaptation to temperature is important in viral ecology as it influences phage infection, propagation and importantly, viability (Meschke and Sobsey, 2003; Jacquet et al., 2005; Ebdon et al., 2007; Jonczyk et al., 2011).

The exact origin of *B. fragilis* strain GB124 is unknown. It was isolated from untreated wastewater from a treatment plant in South East England and wastewater and impacted surface waters in the United Kingdom, United States, Brazil, and India (Payan et al., 2005; Prado et al., 2018; Wadhwa et al., 2018). Coupled with previously reported existence of a clear human gut-associated eco-genomic signature within the *Bacteroides* phage genomes (Ogilvie et al., 2012, 2013, 2018), it is assumed to be a human gut commensal.

*Bacteroides fragilis* phages when compared with other *Bacteroides* spp. phages have been shown to share little homology (Gilbert et al., 2017). A study looking at  $\phi$ Brb01

and  $\phi$ Brb02 phages, that are capable of infecting a *Bacteroides* isolate, has shown them to be phylogenetically distant to both  $\phi$ B124-14 and  $\phi$ B40-8 based on the comparative analysis of the large terminase subunit gene (Gilbert et al., 2017). Similarly, vB\_BfrS\_23 shares little or no sequence identity to 27 recently published *B. thetaiotaomicron* phages (Hryckowian et al., 2020). Although some CrAssphages such as  $\phi$ CrAss001 infect *Bacteroides intestinalis* (Shkoporov et al., 2018), they share little sequence identity to  $\phi$ B124-14 and  $\phi$ B40-8 (Garcia-Aljaro et al., 2017). Considering that the *B. fragilis* phage presented in this study shares a high sequence homology to other *B. fragilis* phages including  $\phi$ B124-14 and  $\phi$ B40-8 and thus we can infer that vB\_BfrS\_23 is also unrelated to other *Bacteroides* spp phages and CrAssphages.

## CONCLUSION

The isolation and characterization of phage vB\_BfrS\_23 not only adds to and builds on fledgling phage databases, but it should also facilitate the detection other *Bacteroides* phages in human fecal metagenomes. As both the bacterial host and the new phage reported here are sourced from municipal wastewater they have considerable potential as, (1) highly specific novel therapeutic agents, (2) as tools for testing the efficacy of water and wastewater reuse technologies (spiking studies), and (3) as molecular or metagenome-based Microbial Source Tracking genetic marker for identifying human fecal transmission pathways in contaminated water and food (McMinn et al., 2014; Dias et al., 2018).

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://www.ncbi.nlm.nih.gov/genbank/>, SAMN14 843706 SRA; <https://www.ncbi.nlm.nih.gov/genbank/>, SRX8283257; <https://www.ncbi.nlm.nih.gov/genbank/>, SRX SRX828326; <https://www.ncbi.nlm.nih.gov/genbank/>, SRX8275163; <https://www.ncbi.nlm.nih.gov/genbank/>, SRX8275162.

## REFERENCES

- Alikhan, N. F., Petty, N. K., Ben Zakour, N. L., and Beatson, S. A. (2011). BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* 12:402. doi: 10.1186/1471-2164-12-402
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410.
- Arndt, D., Grant, J. R., Marcu, A., Sajed, T., Pon, A., Liang, Y. J., et al. (2016). PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res.* 44, W16–W21. doi: 10.1093/nar/gkw387
- Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., et al. (2008). The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. doi: 10.1186/1471-2164-9-75
- Bar-On, Y. M., Phillips, R., and Milo, R. (2018). The biomass distribution on Earth. *Proc. Natl. Acad. Sci. U.S.A.* 115, 6506–6511. doi: 10.1073/pnas.1711842115

## AUTHOR CONTRIBUTIONS

SC and MT conceived and designed the experiments. MT, FN, JE, LH, and SC wrote the manuscript. SC supervised the research. MT, FN, RH, CB, and LH executed the experimental work. MT, FN, LH, MC, TW, JE, and SC carried out the data interpretation. All authors revised, read, and approved the final manuscript.

## FUNDING

This work was supported in part by the UK Biotechnology and Biological Sciences Research Council (BBSRC) and BBSRC Institute Strategic Programme grant BB/R012490/1 to the Gut Microbes and Health programme and its constituent project(s) BBS/E/F/000PR10353 and BBS/E/F/000PR10356 (SRC), Invest in ME Research and UEA-Faculty of Health co-funded Ph.D. studentships (FN and RH), and award R42894 from The Solve ME/CFS Initiative (FN).

## ACKNOWLEDGMENTS

We thank Dr. Lesley Ogilvie (Max Planck Institute for Molecular Genetics, Berlin, Germany) and Dr. Brian Jones (Dept of Biology and Biochemistry, University of Bath, BA2 7AY, UK) for useful discussions during design and analysis of this study. Additionally, we thank the late Dr. Ella Bond (Norwich, UK) for providing *Bacteroides fragilis* strains. We would like to thank Shen-Yuan Hsieh (Gut Microbes and Health Research Programme, Quadram Institute Biosciences, Norwich Research Park, Norwich, UK and Norwich medical School, University of East Anglia, Norwich NR4 7TJ, UK) for useful laboratory support. We thank the JIC Bioimaging facility and staff for enabling TEM phage analysis.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2020.583378/full#supplementary-material>

- Barr, J. J., Auro, R., Furlan, M., Whiteson, K. L., Erb, M. L., Pogliano, J., et al. (2013). Bacteriophage adhering to mucus provide a non-host-derived immunity. *Proc. Natl. Acad. Sci. U.S.A.* 110, 10771–10776. doi: 10.1073/pnas.1305923110
- Bertrand, I., Schijven, J. F., Sanchez, G., Wyn-Jones, P., Ottoson, J., Morin, T., et al. (2012). The impact of temperature on the inactivation of enteric viruses in food and water: a review. *J. Appl. Microbiol.* 112, 1059–1074. doi: 10.1111/j.1365-2672.2012.05267.x
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170
- Booth, S. J., Vantassell, R. L., Johnson, J. L., and Wilkins, T. D. (1979). Bacteriophages of *Bacteroides*. *Rev. Infect. Dis.* 1, 325–336.
- Brettin, T., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Olsen, G. J., et al. (2015). RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci. Rep.* 5:8365. doi: 10.1038/srep08365



- Carreras, C. W., and Santi, D. V. (1995). The catalytic mechanism and structure of thymidylate synthase. *Ann. Rev. Biochem.* 64, 721–762. doi: 10.1146/annurev.bi.64.070195.003445
- Carver, T., Harris, S. R., Berriman, M., Parkhill, J., and McQuillan, J. A. (2012). Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics* 28, 464–469. doi: 10.1093/bioinformatics/btr703
- Casjens, S. R., and Hendrix, R. W. (2015). Bacteriophage lambda: early pioneer and still relevant. *Virology* 479, 310–330. doi: 10.1016/j.virol.2015.02.010
- Chen, L. H., Zheng, D. D., Liu, B., Yang, J., and Jin, Q. (2016). VFDB 2016: hierarchical and refined dataset for big data analysis-10 years on. *Nucleic Acids Res.* 44, D694–D697. doi: 10.1093/nar/gkv1239
- Chun, J., Oren, A., Ventosa, A., Christensen, H., Arahal, D. R., da Costa, M. S., et al. (2018). Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int. J. Syst. Evol. Microbiol.* 68, 461–466. doi: 10.1099/ijsem.0.002516
- Davies, E. V., Winstanley, C., Fothergill, J. L., and James, C. E. (2016). The role of temperate bacteriophages in bacterial infection. *FEMS Microbiol. Lett.* 363:fnw015. doi: 10.1093/femsle/fnw015
- Deschavanne, P., Dubow, M. S., and Regeard, C. (2010). The use of genomic signature distance between bacteriophages and their hosts displays evolutionary relationships and phage growth cycle determination. *Viol. J.* 7:163. doi: 10.1186/1743-422x-7-163
- Dias, E., Ebdon, J., and Taylor, H. (2018). The application of bacteriophages as novel indicators of viral pathogens in wastewater treatment systems. *Water Res.* 129, 172–179. doi: 10.1016/j.watres.2017.11.022
- Duhaime, M. B., Wichels, A., Waldmann, J., Teeling, H., and Glockner, F. O. (2011). Ecogenomics and genome landscapes of marine *Pseudoalteromonas* phage H105/1. *ISME J.* 5, 107–121. doi: 10.1038/ismej.2010.94
- Ebdon, J., Maite, M., and Taylor, H. (2007). The application of a recently isolated strain of *Bacteroides* (GB-124) to identify human sources of faecal pollution in a temperate river catchment. *Water Res.* 41, 3683–3690. doi: 10.1016/j.watres.2006.12.020
- Escartin, F., Skouloubri, S., Liebl, U., and Myllykallio, H. (2008). Flavin-dependent thymidylate synthase X limits chromosomal DNA replication. *Proc. Natl. Acad. Sci. U.S.A.* 105, 9948–9952. doi: 10.1073/pnas.0801356105
- Feldgarden, M., Brover, V., Haft, D. H., Prasad, A. B., Slotta, D. J., Tolstoy, I., et al. (2020). Validating the AMRFinder tool and resistance gene database by using antimicrobial resistance genotype-phenotype correlations in a collection of isolates. *Antimicrob. Agents Chemother.* 64:e00361-20. doi: 10.1128/AAC.00361-20
- Galata, V., Fehlmann, T., Backes, C., and Keller, A. (2019). PLSDB: a resource of complete bacterial plasmids. *Nucleic Acids Res.* 47, D195–D202. doi: 10.1093/nar/gky1050
- García-Aljaro, C., Balleste, E., Muniesa, M., and Jofre, J. (2017). Determination of crAssphage in water samples and applicability for tracking human faecal pollution. *Microb. Biotechnol.* 10, 1775–1780. doi: 10.1111/1751-7915.12841
- Gerdes, K., Rasmussen, P. B., and Molin, S. (1986). Unique type of plasmid maintenance function - postsegregational killing of plasmid-free cells. *Proc. Natl. Acad. Sci. U.S.A.* 83, 3116–3120.
- Gilbert, R. A., Kelly, W. J., Altermann, E., Leahy, S. C., Minchin, C., Ouwerkerk, D., et al. (2017). Toward understanding phage: host interactions in the rumen; complete genome sequences of lytic phages infecting rumen bacteria. *Front. Microbiol.* 8:2340. doi: 10.3389/fmicb.2017.02340
- Gorvitovskaia, A., Holmes, S. P., and Huse, S. M. (2016). Interpreting *Prevotella* and *Bacteroides* as biomarkers of diet and lifestyle. *Microbiome* 4:15. doi: 10.1186/s40168-016-0160-7
- Hawkins, S. A., Layton, A. C., Ripp, S., Williams, D., and Sayler, G. S. (2008). Genome sequence of the *Bacteroides fragilis* phage ATCC 51477-B1. *Viol. J.* 5:97. doi: 10.1186/1743-422x-5-97
- Hendrix, R. W., Smith, M. C. M., Burns, R. N., Ford, M. E., and Hatfull, G. F. (1999). Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. *Proc. Natl. Acad. Sci. U.S.A.* 96, 2192–2197. doi: 10.1073/pnas.96.5.2192
- Hooper, L. V., Stappenbeck, T. S., Hong, C. V., and Gordon, J. I. (2003). Angiogenins: a new class of microbicidal proteins involved in innate immunity. *Nat. Immunol.* 4, 269–273. doi: 10.1038/ni888
- Hryckowian, A. J., Merrill, B. D., Porter, N. T., Van Treuren, W., Nelson, E. J., Garlena, R. A., et al. (2020). *Bacteroides thetaiotaomicron*-infecting bacteriophage isolates inform sequence-based host range predictions. *Cell Host Microbe* 28, 371–379.e5. doi: 10.1016/j.chom.2020.06.011
- Jacquet, S., Domaizon, I., Personnic, S., Ram, A. S. P., Hedail, M., Duhamal, S., et al. (2005). Estimates of protozoan- and viral-mediated mortality of bacterioplankton in Lake Bourget (France). *Freshwater Biol.* 50, 627–645. doi: 10.1111/j.1365-2427.2005.01349.x
- Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T., and Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* 9:5114. doi: 10.1038/s41467-018-07641-9
- Jonczyk, E., Klak, M., Miedzybrodzki, R., and Gorski, A. (2011). The influence of external factors on bacteriophages-review. *Folia Microbiol.* 56, 191–200. doi: 10.1007/s12223-011-0039-8
- Joshi, N. A., and Fass, J. N. (2011). *Sickle: A Sliding-Window, Adaptive, Quality-Based Trimming Tool for FastQ Files [Online]*. Available online at: <https://github.com/najoshi/sickle> (accessed July 24, 2019).
- Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010
- Keller, R., and Traub, N. (1974). Characterization of *Bacteroides-fragilis* bacteriophage recovered from animal sera - observations on nature of *Bacteroides* phage carrier cultures. *J. Gen. Virol.* 24, 179–189. doi: 10.1099/0022-1317-24-1-179
- Kim, J. M., Lee, J. Y., Yoon, Y. M., Oh, Y. K., Kang, J. S., Kim, Y. J., et al. (2006). *Bacteroides fragilis* enterotoxin induces cyclooxygenase-2 and fluid secretion in intestinal epithelial cells through NF-kappa B activation. *Eur. J. Immunol.* 36, 2446–2456. doi: 10.1002/eji.200535808
- King, C. H., Desai, H., Sylvestry, A. C., LoTempio, J., Ayanyan, S., Carrie, J., et al. (2019). Baseline human gut microbiota profile in healthy people and standard reporting template. *PLoS One* 14:e0206484. doi: 10.1371/journal.pone.0206484
- Kropinski, A. M. (2018). Practical advice on the one-step growth curve. *Methods Mol. Biol.* 1681, 41–47. doi: 10.1007/978-1-4939-7343-9\_3
- Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054
- Lemire, S., Figueroa-Bossi, N., and Bossi, L. (2011). Bacteriophage crosstalk: coordination of prophage induction by trans-acting antirepressors. *PLoS Genet.* 7:e1002149. doi: 10.1371/journal.pgen.1002149
- Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S. N., Chitsaz, F., Geer, L. Y., et al. (2015). CDD: NCBI's conserved domain database. *Nucleic Acids Res.* 43, D222–D226. doi: 10.1093/nar/gku1221
- Mastrantonio, P., Cardines, R., and Spigaglia, P. (1996). Oligonucleotide probes for detection of cephalosporinases among *Bacteroides* strains. *Antimicrob. Agents Chemother.* 40, 1014–1016. doi: 10.1128/Aac.40.4.1014
- Mazmanian, S. K., Liu, C. H., Tzianabos, A. O., and Kasper, D. L. (2005). An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. *Cell* 122, 107–118. doi: 10.1016/j.cell.2005.05.007
- McMinn, B. R., Ashbolt, N. J., and Korajkic, A. (2017). Bacteriophages as indicators of faecal pollution and enteric virus removal. *Let. Appl. Microbiol.* 65, 11–26. doi: 10.1111/lam.12736
- McMinn, B. R., Korajkic, A., and Ashbolt, N. J. (2014). Evaluation of *Bacteroides fragilis* GB-124 bacteriophages as novel human-associated faecal indicators in the United States. *Let. Appl. Microbiol.* 59, 115–121. doi: 10.1111/lam.12252
- Meschke, J. S., and Sobsey, M. D. (2003). Comparative reduction of Norwalk virus, poliovirus type 1, F+ RNA coliphage MS2 and *Escherichia coli* in miniature soil columns. *Water Sci. Technol.* 47, 85–90.
- Ogilvie, L. A., Bowler, L. D., Caplin, J., Dedi, C., Diston, D., Cheek, E., et al. (2013). Genome signature-based dissection of human gut metagenomes to extract subliminal viral sequences. *Nat. Commun.* 4:2420. doi: 10.1038/ncomms3420
- Ogilvie, L. A., Caplin, J., Dedi, C., Diston, D., Cheek, E., Bowler, L., et al. (2012). Comparative (meta)genomic analysis and ecological profiling of human gut-specific bacteriophage phi B124-14. *PLoS One* 7:e35053. doi: 10.1371/journal.pone.0035053
- Ogilvie, L. A., Nzakizwanayo, J., Guppy, F. M., Dedi, C., Diston, D., Taylor, H., et al. (2018). Resolution of habitat-associated ecogenomic signatures in bacteriophage genomes and application to microbial source tracking. *ISME J.* 12, 942–958. doi: 10.1038/s41396-017-0015-7

- Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T., et al. (2014). The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res.* 42, D206–D214. doi: 10.1093/nar/gkt1226
- Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25, 1043–1055. doi: 10.1101/gr.186072.114
- Payan, A., Ebdon, J., Taylor, H., Gantzer, C., Ottoson, J., Papageorgiou, G. T., et al. (2005). Method for isolation of *Bacteroides* bacteriophage host strains suitable for tracking sources of fecal pollution in water. *Appl. Environ. Microbiol.* 71, 5659–5662. doi: 10.1128/Aem.71.9.5659-5662.2005
- Perez-Brocal, V., Garcia-Lopez, R., Nos, P., Beltran, B., Moret, I., and Moya, A. (2015). Metagenomic analysis of crohn's disease patients identifies changes in the virome and microbiome related to disease status and therapy, and detects potential interactions and biomarkers. *Inflamm. Bowel Dis.* 21, 2515–2532. doi: 10.1097/Mib.0000000000000549
- Prado, T., Bruni, A. D., Barbosa, M. R. F., Bonanno, V. M. S., Garcia, S. C., and Sato, M. I. Z. (2018). Distribution of human fecal marker GB-124 bacteriophages in urban sewage and reclaimed water of Sao Paulo city, Brazil. *J. Water Health* 16, 289–299. doi: 10.2166/wh.2017.011
- Puig, M., and Girones, R. (1999). Genomic structure of phage B40-8 of *Bacteroides fragilis*. *Microbiol Sgm* 145, 1661–1670. doi: 10.1099/13500872-145-7-1661
- Purnell, S., Ebdon, J., Buck, A., Tupper, M., and Taylor, H. (2015). Bacteriophage removal in a full-scale membrane bioreactor (MBR)—implications for wastewater reuse. *Water Res.* 73, 109–117. doi: 10.1016/j.watres.2015.01.019
- Purnell, S., Ebdon, J., Buck, A., Tupper, M., and Taylor, H. (2016). Removal of phages and viral pathogens in a full-scale MBR: implications for wastewater reuse and potable water. *Water Res.* 100, 20–27. doi: 10.1016/j.watres.2016.05.013
- Rasmussen, B. A., Bush, K., and Tally, F. P. (1993). Antimicrobial resistance in *Bacteroides*. *Clin. Infect. Dis.* 16, S390–S400. doi: 10.1093/clinids/16.Supplement\_4.S390
- Roberts, M. C. (1996). Tetracycline resistance determinants: mechanisms of action, regulation of expression, genetic mobility, and distribution. *FEMS Microbiol. Rev.* 19, 1–24. doi: 10.1111/j.1574-6976.1996.tb00251.x
- Rogers, M. B., Parker, A. C., and Smith, C. J. (1993). Cloning and characterization of the endogenous cephalosporinase gene, cepa, from *Bacteroides fragilis* reveals a new subgroup of ambler class-a beta-lactamases. *Antimicrob. Agents. Chemother.* 37, 2391–2400. doi: 10.1128/Aac.37.11.2391
- rrwrick/Porechop (2020). *Porechop [Online]*. Available online at: <https://github.com/rrwrick/Porechop> (accessed Jun 10, 2020).
- Saitou, N., and Nei, M. (1987). The neighbor-joining method—a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425.
- Sathaliyawa, T., Islam, M. Z., Li, Q., Fokine, A., Rossmann, M. G., and Rao, V. B. (2010). Functional analysis of the highly antigenic outer capsid protein, Hoc, a virus decoration protein from T4-like bacteriophages. *Mol. Microbiol.* 77, 444–455. doi: 10.1111/j.1365-2958.2010.07219.x
- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153
- Sender, R., Fuchs, S., and Milo, R. (2016). Revised estimates for the number of human and bacteria cells in the body. *PLoS Biol* 14:e1002533. doi: 10.1371/journal.pbio.1002533
- Shenoy, P. A., Vishwanath, S., Gawda, A., Shetty, S., Anegundi, R., Varma, M., et al. (2017). Anaerobic bacteria in clinical specimens—frequent, but a neglected lot: a five year experience at a tertiary care hospital. *J. Clin. Diagn. Res.* 11, DC44–DC48. doi: 10.7860/JCDR/2017/26009.10311
- Shkorporov, A. N., Khokhlova, E. V., Fitzgerald, C. B., Stockdale, S. R., Draper, L. A., Ross, R. P., et al. (2018). Phi CrAss001 represents the most abundant bacteriophage family in the human gut and infects *Bacteroides intestinalis*. *Nat. Commun.* 9:4781. doi: 10.1038/s41467-018-07225-7
- Siguier, P., Perochon, J., Lestrade, L., Mahillon, J., and Chandler, M. (2006). ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Res.* 34, D32–D36. doi: 10.1093/nar/gkj014
- Stern, A., Mayrose, I., Penn, O., Shaul, S., Gophna, U., and Pupko, T. (2010). An evolutionary analysis of lateral gene transfer in thymidylate synthase enzymes. *Syst. Biol.* 59, 212–225. doi: 10.1093/sysbio/syp104
- Sullivan, M. J., Petty, N. K., and Beatson, S. A. (2011). Easyfig: a genome comparison visualizer. *Bioinformatics* 27, 1009–1010. doi: 10.1093/bioinformatics/btr039
- Sun, Z. P., Levi, Y., Kiene, L., Dumoutier, N., and Lucena, F. (1997). Quantification of bacteriophages of *Bacteroides fragilis* in environmental water samples of Seine river. *Water Air Soil Poll.* 96, 175–183.
- Tariq, M. A., Everest, F. L. C., Cowley, L. A., De Soya, A., Holt, G. S., Bridge, S. H., et al. (2015). A metagenomic approach to characterize temperate bacteriophage populations from cystic fibrosis and non-cystic fibrosis bronchiectasis patients. *Front. Microbiol.* 6:97. doi: 10.3389/fmicb.2015.00097
- Tartera, C., and Jofre, J. (1987). Bacteriophages active against *Bacteroides fragilis* in sewage-polluted waters. *Appl. Environ. Microb.* 53, 1632–1637. doi: 10.1128/Aem.53.7.1632-1637.1987
- tseemann/abricate (2020). *Abricate [Online]*. Available online at: <https://github.com/tseemann/abricate> (accessed Jun 10, 2020).
- Tzianabos, A. O., Onderdonk, A. B., Rosner, B., Cisneros, R. L., and Kasper, D. L. (1993). Structural features of polysaccharides that induce intraabdominal abscesses. *Science* 262, 416–419. doi: 10.1126/science.8211161
- Wadhwa, A., Dutta, S., Ebdon, J., Chowdhary, G., Kapoor, R., Wang, Y. K., et al. (2018). Successful application of microbial source tracking using Gb-124 bacteriophage as an indicator of human fecal contamination in environmental samples in Kolkata, India. *Am. J. Trop. Med. Hyg.* 99, 639–639.
- Wang, I. N., Smith, D. L., and Young, R. (2000). Holins: the protein clocks of bacteriophage infections. *Annu. Rev. Microbiol.* 54, 799–825. doi: 10.1146/annurev.micro.54.1.799
- Wick, R. R., Judd, L. M., Gorrie, C. L., and Holt, K. E. (2017). Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput. Biol.* 13:e1005595. doi: 10.1371/journal.pcbi.1005595
- Wu, S. G., Powell, J., Mathioudakis, N., Kane, S., Fernandez, E., and Sears, C. L. (2004). *Bacteroides fragilis* enterotoxin induces intestinal epithelial cell secretion of interleukin-8 through mitogen-activated protein kinases and a tyrosine kinase-regulated nuclear factor-kappa B pathway. *Infect. Immun.* 72, 5832–5839. doi: 10.1128/Iai.72.10.5832-5839.2004
- Xu, J., Bjursell, M. K., Himrod, J., Deng, S., Carmichael, L. K., Chiang, H. C., et al. (2003). A genomic view of the human *Bacteroides thetaiotaomicron* symbiosis. *Science* 299, 2074–2076. doi: 10.1126/science.1080029
- Yarmolinsky, M. B. (1995). Programmed cell-death in bacterial populations. *Science* 267, 836–837. doi: 10.1126/science.7846528
- Young, R., and Blasi, U. (1995). Holins - form and function in bacteriophage lysis. *FEMS Microbiol. Rev.* 17, 191–205. doi: 10.1016/0168-6445(94)00079-4
- Young, R. Y. (1992). Bacteriophage lysis - mechanism and regulation. *Microbiol. Rev.* 56, 430–481. doi: 10.1128/Mmbr.56.3.430-481.1992
- Zaneveld, J. R., Lozupone, C., Gordon, J. I., and Knight, R. (2010). Ribosomal RNA diversity predicts genome diversity in gut bacteria and their relatives. *Nucleic Acids Res.* 38, 3869–3879. doi: 10.1093/nar/gkq066
- Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., et al. (2012). Identification of acquired antimicrobial resistance genes. *J. Antimicrob. Chemother.* 67, 2640–2644. doi: 10.1093/jac/dks261

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Tariq, Newberry, Haagmans, Booth, Wileman, Hoyles, Clokie, Ebdon and Carding. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.