JOURNAL OF BACTERIOLOGY, Oct. 2011, p. 5300–5313 Vol. 193, No. 19
0021-9193/11/$12.00 doi:10.1128/JB.05287-11

Genomes and Characterization of Phages Bcep22 and BcepIL02, Founders of a Novel Phage Type in *Burkholderia cenocepacia*▽†

Jason J. Gill,[1,2] Elizabeth J. Summer,[1,2] William K. Russell,[3] Stephanie M. Cologna,[3] Thomas M. Carlile,[1] Alicia C. Fuller,[1] Kate Kitsopoulos,[1] Leslie M. Mebane,[1] Brandi N. Parkinson,[1] David Sullivan,[1] Lisa A. Carmody,[4] Carlos F. Gonzalez,[2,5] John J. LiPuma,[4] and Ry Young[1,2]*

Department of Biochemistry and Biophysics, Texas A&M University, College Station, Texas 77843-2128[1]; Center for Phage Technology, Texas A&M University, College Station, Texas 77843-2128[2]; Department of Chemistry, Texas A&M University, College Station, Texas 77843-3012[3]; Department of Pediatrics and Communicable Disease, University of Michigan Medical School, Ann Arbor, Michigan 48109[4]; and Department of Plant Pathology and Microbiology, Texas A&M University, College Station, Texas 77843-2132[5]

Received 19 May 2011/Accepted 18 July 2011

Within the *Burkholderia cepacia* complex, *B. cenocepacia* is the most common species associated with aggressive infections in the lungs of cystic fibrosis patients, causing disease that is often refractive to treatment by antibiotics. Phage therapy may be a potential alternative form of treatment for these infections. Here we describe the genome of the previously described therapeutic *B. cenocepacia* podophage BcepIL02 and its close relative, Bcep22. Phage Bcep22 was found to contain a circularly permuted genome of 63,882 bp containing 77 genes; BcepIL02 was found to be 62,714 bp and contains 76 predicted genes. Major virion-associated proteins were identified by proteomic analysis. We propose that these phages comprise the founding members of a novel podophage lineage, the Bcep22-like phages. Among the interesting features of these phages are a series of tandemly repeated putative tail fiber genes that are similar to each other and also to one or more such genes in the other phages. Both phages also contain an extremely large (ca. 4,600-amino-acid), virion-associated, multidomain protein that accounts for over 20% of the phages' coding capacity, is widely distributed among other bacterial and phage genomes, and may be involved in facilitating DNA entry in both phage and other mobile DNA elements. The phages, which were previously presumed to be virulent, show evidence of a temperate lifestyle but are apparently unable to form stable lysogens in their hosts. This ambiguity complicates determination of a phage lifestyle, a key consideration in the selection of therapeutic phages.

Bacteria of the genus *Burkholderia* constitute an ecologically and metabolically diverse group of organisms, which includes soil saprophytes, nitrogen-fixing rhizosphere inhabitants, biological control agents, and species that are pathogenic to plants, animals, and humans (14). Within this genus, the *Burkholderia cepacia* complex is currently comprised of at least 17 bacterial species, including several species recognized as opportunistic human pathogens capable of infecting immunocompromised individuals (50, 85). Especially at risk from *B. cepacia* complex infections are individuals with cystic fibrosis (CF) or chronic granulomatous disease. While all members of the *B. cepacia* complex have been associated with infections in the lungs of CF patients, *B. cenocepacia* is the most common member associated with persistent and aggressive infections in these patients, causing necrotizing pneumonia, sepsis, and death. Clinical outcomes of *B. cenocepacia* infections in CF patients are typically poor, and this bacterium is often innately resistant to many chemical antibiotics. The severity of these infections, coupled with the current lack of available effective treatments, has ignited interest in the development of alternative treatment options.

The use of bacteriophages for the treatment of bacterial infections, a practice commonly termed phage therapy, is an old idea, first proposed by the phage codiscoverer Felix d'Herelle (78). While phage therapy was largely abandoned in the West with the discovery of chemical antibiotics, the practice continued in the former Soviet Union and its satellite states and is still used routinely today in these regions as an adjunct to antibiotics in human medicine (72). Phage therapy has been a topic of increasing research in the West over the past several years, as new bacterial pathogens have emerged, antimicrobial resistance in important human pathogens has become more widespread, and the development pipelines for novel small-molecule antimicrobials have largely run dry (63). Phages that have utility as potential therapeutic agents should meet at least two minimal criteria: they should be virulent (unable to form lysogens that would then be insensitive to further phage infections) and they must target the specific strains of the pathogen that they are intended to treat. While determination of the latter is relatively straightforward, the issue of phage virulence is not always clear. By its nature, a temperate lifestyle is easy to prove (by isolation of stable lysogens) but not as easily disproven, as failure to establish lysogeny is not conclusive proof of virulence.

A number of phages capable of infecting clinically relevant

* Corresponding author. Mailing address: Department of Biochemistry and Biophysics, Texas A&M University, 2128 TAMU, College Station, TX 77843-2128. Phone: (975) 845-2087. Fax: (979) 862-9718. E-mail: ryland@tamu.edu.

† Supplemental material for this article may be found at http://jb.asm.org/.

▽ Published ahead of print on 29 July 2011.

members of the *B. cepacia* complex have been characterized, and some have also been sequenced. Several of these phages, such as NS1 and NS2 (57), DK1 and DK3 (41), KS4 and KS9 (71), BcepMu (77), and KS10 (22), were isolated from supernatants of lysogenic bacteria and are therefore temperate phages. Based on their genome organization, the environmentally isolated phages Bcep781, Bcep1, Bcep43, and BcepB1A have been reported to be virulent phages (76). However, while these phages appear to possess the desired characteristic of virulence, they were not isolated on Bcc strains that are clinically relevant for purposes of phage therapy. Recent work has shown that a lysogeny-defective mutant of the BcepMu-like KS4, the temperate phage KS14, and the less-well-characterized phage KS12 reduced *B. cepacia* complex-induced mortality in a wax worm model (70). A repressorless virulent mutant of phage KS9 has also been shown to reduce mortality in this model (49). Here we describe two related *B. cepacia* complex phages, Bcep22 and BcepIL02. Both phages are able to infect pathogenic clinical isolates of *B. cenocepacia*, and one of them, BcepIL02, has been shown to reduce the colonization of pathogenic *B. cenocepacia* in a mouse lung model (10).

MATERIALS AND METHODS

Strains and culture conditions. Phage Bcep22 was isolated from soil collected in the summer of 2000 from Orange County, NY, by enrichment against a *B. cenocepacia* strain (S198C28A) isolated from the same soil sample. Phage BcepIL02 was isolated from a corn field soil sample in Champaign County, IL, in the summer of 2006, as described previously (10). *B. cenocepacia* strains were routinely cultured aerobically at 37°C on tryptone nutrient broth (TNB; 8.5 g/liter NaCl, 1 g/liter glucose, 5 g/liter tryptone, 2.5 g/liter yeast extract) or tryptone nutrient agar (TNA; TNB amended with 15 g/liter Bacto agar). Phages Bcep22 and BcepIL02 were routinely cultured at 37°C on *B. cenocepacia* strains AU1054 and PC184, respectively, using TNA plates and overlays of TNA soft agar (TNB amended with 5 g/liter Bacto agar) and standard culture techniques (1). Small-volume phage lysates were harvested from TNA soft agar overlays displaying nearly confluent lysis by flooding plates with SM buffer (50 mM Tris-HCl [pH 7.5], 100 mM NaCl, 8 mM $MgSO_4$, 0.01% [wt/vol] gelatin), aspirating the resulting soft agar slurry, centrifuging (10,000 × *g*, 10 min, 4°C), and sterilizing by passage through a 0.22-μm syringe filter (Millipore, Billerica, MA). For lysogen formation assays (see below), plate lysates were harvested in LB broth (10 g/liter tryptone, 5 g/liter yeast extract, 10 g/liter NaCl) instead of SM buffer. Large-scale phage lysates for proteomic analysis were prepared by inoculating 2 liters of TNB with the appropriate bacterial host to a final optical density at 550 nm (OD_{550}) of 0.05, adding phage to a final multiplicity of infection (MOI) of 0.025, and incubating at 37°C with vigorous aeration. When the culture was determined to have mostly lysed (as determined by monitoring the OD_{550}) after 10 to 12 h, the lysate was harvested by shaking for 10 min in the presence of 0.5% (vol/vol) $CHCl_3$, centrifugation (12,000 × *g*, 15 min, 4°C), and retention of the clarified supernatant.

Genome sequencing and annotation. Phage genomic DNA was prepared from filter-sterilized plate lysates with the Wizard DNA cleanup kit (Promega, Madison, WI) using a modified protocol (73). Phage Bcep22 was sequenced to ~8-fold coverage by dideoxy termination sequencing as described previously (77). Phage BcepIL02 was sequenced to 23-fold coverage by 454 pyrosequencing at a commercial facility (Roche/454 Life Sciences, Branford, CT). Gap closure was completed by amplification of gap regions by PCR followed by dideoxy termination sequencing of the products. Completed DNA sequences were analyzed for potential protein-coding genes with Genemark.hmm (48), and the predicted coding regions were manually refined in Artemis (66). Protein sequences were compared to the NCBI protein database by using BLASTp (9); tRNA genes were detected with tRNAscan (68); rho-independent terminators were detected with TransTerm HP (37); promoters were detected using BPROM (Softberry). Potential functional domains within proteins were determined by analysis with the stand-alone InterProScan suite version 4.7 run locally (29). Protein localization was predicted by use of SignalP 3.0 (6), LipoP 1.0 (33), and TMHMM (http://www.cbs.dtu.dk/services/TMHMM). Phage genome maps were rendered with DNA Master (http://cobamide2.bio.pitt.edu/computer.htm). Pairwise protein sequence similarities were calculated by the Dice method, using BLASTp determinations of amino acid identity.

Proteomic analysis. Phages were purified for proteomic analysis by two rounds of banding in a CsCl gradient, based on previously described methods (42, 67). Phage was then dialyzed thoroughly against gelatin-free SM buffer and concentrated further using a 100-kDa molecular weight cutoff Amicon ultrafiltration unit (Millipore). Phage proteins were analyzed by SDS-PAGE by boiling purified phage for 5 min in Laemmli sample buffer (62.5 mM Tris-HCl [pH 6.8], 2% SDS, 10% glycerol, 5% β-mercaptoethanol, 0.001% bromophenol blue), and the equivalent of ~3×10^{10} PFU was loaded per lane on a 10% SDS-PAGE gel by standard methods (40). Coomassie-stained bands were excised and subjected to reduction using dithiothreitol and alkylation with iodoacetamide followed by tryptic digestion carried out robotically using a ProGest robotic device (Genomic Solutions, Ann Arbor, MI) suited for automated digestion protocols and temperature-controlled reactions. The resulting tryptic peptides were concentrated and desalted using C_{18} ZipTips (Millipore) and spotted robotically directly onto a matrix-assisted laser desorption ionization (MALDI) target (Genomic Solutions) using a 7.5-mg/ml solution of α-cyano 4-hydroxycinnamic acid (CHCA). Additionally, phage proteins were subjected to liquid chromatography-MALDI-tandem mass spectrometry (LC-MALDI-MS/MS) analysis. Phage samples were denatured in 6 M urea and 1 mM dithiothreitol (DTT; Sigma) and subjected to 5 freeze-thaw cycles between a dry ice-ethanol bath and a 37°C heat block. The samples were centrifuged at 14,000 rpm for 10 min, and the supernatant was subjected to solution-phase tryptic digestion. Briefly, 10 mM iodoacetamide was added to the protein solution followed by addition of trypsin at a 1:50 (wt/wt) ratio and incubated at 37°C overnight. Approximately 1 μg of the tryptic digestion solution was further separated by loading onto a trap column (300 μm by 1 cm; C_{18} PepMap) coupled with a Vydac (150 μm by 15 cm; C_{18}) analytical column, and chromatography was carried out using an LC Packings Ultimate system (Dionex). The eluent from the analytical column was spotted robotically onto the MALDI target via a ProBot (LC Packings) as described previously (65).

MALDI-MS and MS/MS analysis were carried out on a model 4700 Proteomics Analyzer (Applied Biosystems). For both the gel-based method and the LC-separated peptides, MS data were acquired from *m/z* 700 to 4,500, and the top 10 peaks in each spot were subjected to tandem MS analysis. Spectra were externally calibrated using bradykinin fragment 2–9 and adrenocorticotropic hormone fragment 18–39. The resulting spectra were subjected to database searching using GPS Explorer (Applied Biosystems) with an in-house license of mascot (v 2.1). MS and MS/MS spectra were used for the in-gel-digested proteins, and only the MS/MS spectra were used for database searching of the LC-MS/MS analysis results. Spectra were searched against a FASTA file containing the predicted protein products of Bcep22 and BcepIL02.

Proteins were considered to be virion associated if they were identified via either the whole-virion or the in-gel digestion methods by two or more unique peptides with ion scores greater than 40, or were identified by 5 or more unique peptides with deviations from the predicted ion mass of less than 20 ppm. Identities of low-scoring peptides were manually confirmed by examination of the raw MS peak output. All identified proteins had total ion score confidence intervals of greater than 99.5% as calculated with GPS Explorer, taking into account the mascot score.

Identification of proteins with homotypic interactions. A lambda repressor fusion library was constructed by fragmentation of Bcep22 DNA by hydroshear, end repair, and ligation into the vector pLM100 and screened as described previously (77). In principle, only DNA inserts encoding a protein sequence capable of complementing the missing C-terminal dimerization domain of the phage lambda *c*I repressor in pLM100 would survive subsequent challenge by phage lambda lacking the *c*I gene (52). Of the 96 surviving colonies screened, 71 contained inserts encoding an in-frame portion of a predicted Bcep22 gene. Only Bcep22 genes containing two or more such hits were retained for further analysis.

Isolation of bacteriophage-insensitive mutants. Bacterial cells were harvested from the surface of four independent overnight plate cultures of strains AU1054 and PC184 and used to inoculate TNB at A_{550} of ~0.1. Phages Bcep22 and BcepIL02 were added to these cultures, respectively, to a final input MOI of ~1, and these cultures were incubated with aeration for 16 to 18 h at 37°C. Cells surviving phage coculture were streaked onto TNB plates, and two surviving colonies showing robust growth in the heaviest, phage-containing part of the streak were subcultured twice to remove contaminating phage. A total of 16 phage-insensitive strains (8 from strain AU1054 challenged with Bcep22 and 8 from PC184 challenged with BcepIL02) were isolated. Additionally, three strains of *B. cenocepacia* AU0728 insensitive to phage BcepIL02 were isolated from homogenized mouse tissue following treatment of infected mice with this phage (10).

Phage sensitivity was assessed by applying 10-μl spots of serially diluted phage

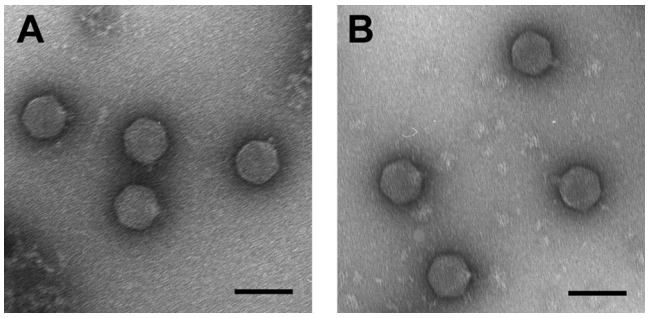

FIG. 1. Transmission electron micrographs of phages Bcep22 (A) and BcepIL02 (B). Samples were stained with 2% (wt/vol) uranyl acetate and observed at 100 kV. Bcep22 was observed to have a capsid diameter of 71 nm (±1.5 nm) and a short, noncontractile tail measuring 15 nm (±2.3 nm) in length and 14 nm (±1.7 nm) in width. The capsid of BcepIL02 was determined to be 72 nm (±2.6 nm) in diameter with a tail of 18 nm (±2.0 nm) in length and 15 nm (±1.0 nm) in width. The differences in dimensions between the two phages were determined to be not statistically significant based on the *t* test ($P < 0.05$). Bars, 100 nm.

(Bcep22 or BcepIL02) to soft agar lawns inoculated with the mutant strains and comparing the efficiency of plating (EOP) to the plating efficiency of the phage on its respective wild-type host. Production of phage by the mutant strains was assayed by removing 0.5-ml aliquots of 16- to 18-h liquid cultures of each strain, mixing vigorously with 20 μl chloroform, and centrifuging at 12,000 × *g* for 1 min at room temperature (RT). Ten-microliter aliquots of each supernatant were spotted to lawns of the mutant strain from which it was produced and also to the respective wild-type host, and lawns were observed for plaque formation.

Assays for lysogen formation. To examine lysogen formation under high-MOI conditions, overnight cultures of strain PC184 were subcultured 1:50 into fresh LB broth and cultured with aeration at 37°C to an OD_{550} of 0.45 to 0.5. Cells were pelleted by centrifugation (12,000 × *g*, 1 min, RT) and resuspended to a final OD_{550} of 1.0 in 0.25 ml of a fresh (<2-week-old) phage BcepIL02 lysate (harvested in LB) or sterile LB. These cell-phage mixtures were incubated for 30 min at RT and then centrifuged as described above. An aliquot of the supernatants was removed, serially diluted, and plated to lawns of PC184 to determine the phage titer remaining following incubation. The remainder of the supernatants were removed, and the cell pellets were resuspended in 0.25 ml of sterile LB. These cells were serially diluted and plated to enumerate the bacterial survivors remaining following phage exposure. This number of surviving CFU was compared to the CFU contained in the LB-only control reaction mixtures to calculate the percentage of bacterial survivors. The postincubation supernatant titers were compared to the titer of the phage stock to calculate the percentage of phage adsorbed to cells during the incubation period. From these data, the actual MOI (MOI_{actual}, the ratio of the number of adsorbed phage to the number of CFU in the phage-free controls [35]) was calculated. These MOI_{actual} values were used to calculate the proportions of cells adsorbing no phage, a single phage, and multiple phage based on a Poisson distribution (1). This experiment was replicated three times.

Transmission electron microscopy. Phage were applied to thin carbon films and stained with 2% (wt/vol) uranyl acetate according to the Valentine method (84). Carbon films were applied to 300-mesh copper grids, air dried, and viewed in a JEOL 1200 EX transmission electron microscope under 100 kV accelerating voltage. Five virions of each phage were measured, and these data were used to calculate mean values and standard deviations.

RESULTS AND DISCUSSION

Overview of Bcep22 and BcepIL02. Bacteriophages Bcep22 and BcepIL02 are both *Podoviridae* and are nearly identical morphologically (Fig. 1). Both phages have ~70-nm isometric heads and short, noncontractile tails. No long tail fibers were observed, but their presence cannot be ruled out, given the difficulty of visualization of phage tail fibers by negative stain. Both phages produce clear, 1- to 3-mm-diameter plaques when plated on their respective hosts on either TNA or LB medium. Phage BcepIL02 is able to form plaques on both its own host (*B. cenocepacia* PC184) and the host of Bcep22 (*B. cenocepacia* AU1054) with equivalent efficiency, while Bcep22 forms plaques only on strain AU1054.

Overview of the genomes of Bcep22 and BcepIL02. The genome of Bcep22 (GenBank accession number AY349011) was found to be 63,882 bases in length, with a 65.3% G+C content, a coding density of 94.6%, and containing 77 predicted protein-coding genes and one predicted serine tRNA gene. The genome of BcepIL02 (GenBank accession number FJ937737) was determined to be 62,714 bases in length, with a 66.2% G+C content, a coding density of 95.2%, and containing 76 predicted protein-coding genes and one predicted serine tRNA gene. The G+C contents of phages Bcep22 and BcepIL02 are both similar to that of *B. cenocepacia*, which possesses a 66.9% G+C content (28). Based on restriction maps of phage genomic DNA and the assembly of individual reads from the Bcep22 clone library, the genomes of both phages appeared to be circularly permuted. The genome sequences were therefore arbitrarily opened immediately after the predicted lysis cassettes.

The genomes of Bcep22 and BcepIL02 were found to be diverged while retaining overall synteny (Fig. 2). Pairwise DNA MegaBLAST analysis of the two genomes produced sequence alignments over 69.8% of the phage genomes, with 89.0% similarity between the two phages in these aligned regions at the DNA level. The phages contain 51 homologous genes that possess amino acid similarities ranging from 15% to 98% (Fig. 2; Table 1). In addition to these, each phage genome contains a set of redundant predicted tail fiber-encoding genes (three in Bcep22, four in BcepIL02), which are similar to each other and also to one or more such genes in the other phages (Fig. 3). Edited BLASTp results for all phage proteins are available in Tables S1 and S2 of the supplemental material.

Proteomic analysis and virion-associated proteins. Proteomic analysis was conducted on CsCl-purified virions of both Bcep22 and BcepIL02 by using dual approaches: whole-virion digestion followed by LC-MALDI analysis, and MALDI analysis of digested SDS-PAGE species. In BcepIL02, 16 proteins were identified as virion associated by one or both methods (Table 2). In Bcep22, eight proteins were directly detected and an additional five proteins were inferred to be virion associated based on the detection of a homologous protein in BcepIL02 (Table 2).

For both phages, the most abundant virion protein, as judged by SDS-PAGE, was an ~40-kDa species, in agreement with the predicted 40-kDa masses of the putative major capsid proteins, gp55 of Bcep22 and gp50 of BcepIL02 (data not shown). No other bands could be determined by MALDI analysis to be cleavage products of these proteins. Peptides corresponding to residues 2 to 12 (SSTVIPFGDSK) and 350 to 361 (DFGVISIDTAAK) of the 364-amino-acid BcepIL02 putative major capsid protein gp50 were detected by the LC-MALDI method, indicating that a full-length version of this protein is assembled into the virion structure following cleavage of the N-terminal methionine by the general methionine aminopeptidase.

Identification of proteins with homotypic interactions. Seventy clones containing homotypically interacting lambda re-

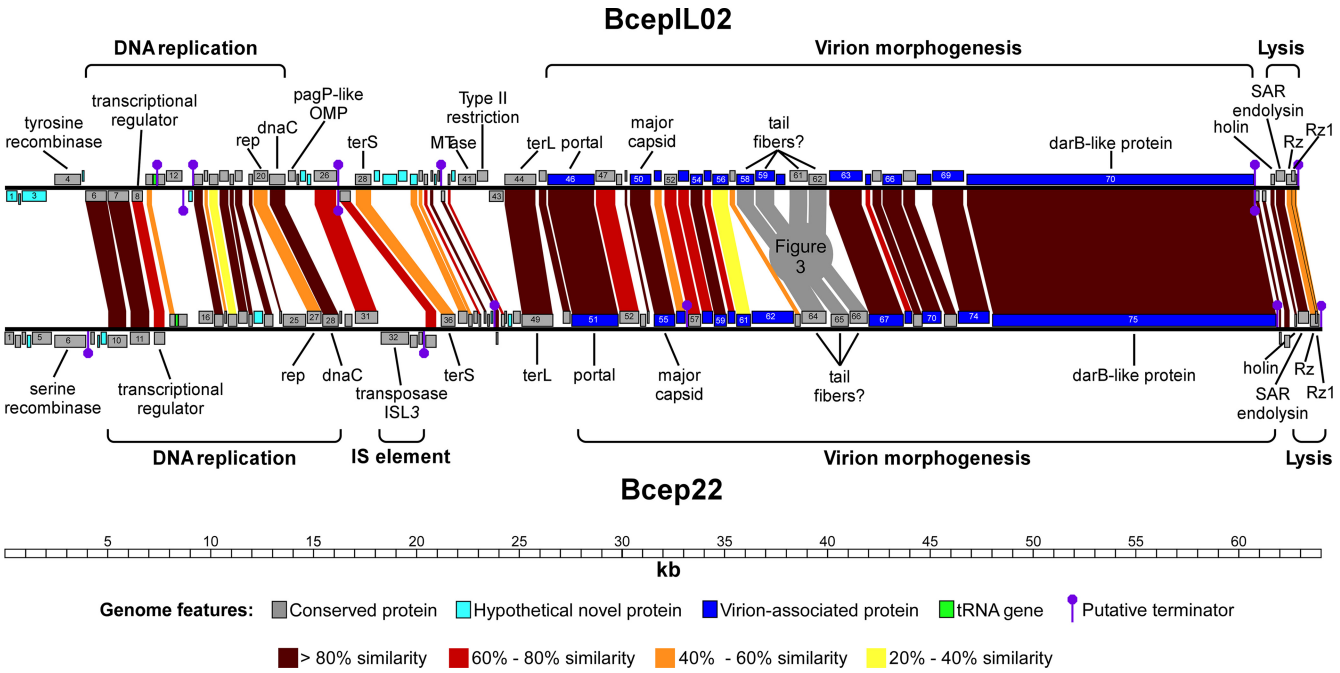

FIG. 2. Comparative genomic maps of phages Bcep22 and BcepIL02. Predicted genes are represented by boxes above and below the heavy black line for each phage genome; boxes above the lines are genes encoded on the forward strand, and those below the lines are on the reverse strand. The ruler below the genomes indicates the scale (in kb). Genome features (conserved, hypothetical novel, and virion-associated proteins; tRNA genes; rho-independent terminators) are color coded according to the legend below the ruler. Protein similarity is denoted by colored bars between genome maps, with colors indicating the percent relatedness according to the legend at the bottom of the figure; similarity of less than 20% is not shown on this map. Selected genes and gene modules are annotated based on predicted function.

pressor fusions were identified, matching various regions of 15 Bcep22 proteins ranging from 12 matches to Bcep22 gp39 to two matches against gp10, gp13, gp19, gp27, gp37, and gp44 (Fig. 4). Of these 15 proteins, 10 are either conserved or hypothetical novel proteins with no known function, although one of these, gp68, was found to be virion associated. Of the remaining five proteins, all have predicted functions associated with a multimeric state, thereby strengthening their functional predictions, and also have homologs present in BcepIL02 (Table 1). Bcep22 gp10 is predicted to function as a RecT-like recombination strand invasion protein, working in concert with the gp11 RecB-like nuclease as the functional equivalent of the *Escherichia coli* RecE/RecT or coliphage lambda Redα/Redβ systems (25, 56) and expected to form multimeric complexes on single-stranded DNA. Bcep22 gp13, a putative transcriptional regulator, appears to be capable of oligomerization via its C terminus (Fig. 4), suggesting it has similar architecture to the lambda *c*I repressor, which contains an N-terminal DNA-binding domain and a C-terminal dimerization domain (3). This finding further supports the annotation of gp13 (and by extension, BcepIL02 gp8) as *c*I-like dimerizing repressors, as the Bcep22 gp13 C terminus is effectively able to complement the lambda *c*I C-terminal domain. The Bcep22 gp13 putative transcriptional regulator could possibly function as a lysogenic repressor.

Bcep22 gp27, predicted to be a replication initiation protein, was found to possess a self-interacting domain in its N terminus (Fig. 4). This is consistent with the behavior of other initiation proteins, such as *E. coli* DnaA, which oligomerizes via its N-terminal domain (34). Bcep22 gp44 is also predicted to function as a transcriptional regulator, due to the presence of a helix-turn-helix (HTH) DNA-binding domain (IPR001387). Unlike gp13, gp44 appears to lack a *c*I-like C-terminal dimerization domain, but it may be able to form dimers on its own in a manner similar to other small, HTH domain repressors, such as lambda Cro (80). Finally, the Bcep22 gp64 putative tail fiber was found to homotypically interact across a broad section of its C terminus. This finding is not unexpected for a putative tail fiber protein, as phage tail fibers commonly form trimers prior to assembly into the virion (45). It is not clear, however, why only one of the three putative tail fiber orthologs found in Bcep22 was detected in this assay. Bcep22 gp64 is highly diverged in its C terminus from the other tail fibers in both Bcep22 and BcepIL02 (Fig. 3), suggesting that this more-atypical sequence possesses self-assembly properties lacking in the other tail fiber orthologs.

Relationships to other phages. While Bcep22 and BcepIL02 were found to be highly related to one another, these phages are not clearly members of any of the currently designated phage lineages, nor are they closely related to other previously sequenced phages. A published analysis of *Podoviridae* taxonomy based on protein sequence relatedness was unable to reliably group Bcep22 with any other sequenced podophages (44).

Phage Bcep22 will be used for presentation of broader interorganism analyses, as it is the founder of the Bcep22-like phage type. The genomes with the closest relationship to the Bcep22-like phages described here are those of an unnamed draft plasmid in *Ochrobactrum intermedium* LMG 3301 (ACQA01000004), and *Sinorhizobium meliloti* phage PBC5 (AF448724), a phage of undetermined morphology and life-

TABLE 1. Comparative annotations of *Burkholderia* phages Bcep22 and BcepIL02[a]

Phage Bcep22			% protein identity	Phage BcepIL02		
Putative function	Calculated mass (kDa)	Gene name		Gene name	Calculated mass (kDa)	Putative function
Conserved protein	18.3	*1*		*1*	22.7	Hypothetical novel
Conserved protein	12.6	*2*		*2*	7.1	Hypothetical novel
Conserved protein	9.8	*3*		*3*	43.5	Conserved protein
Hypothetical novel	7.5	*4*		*4*	47.7	Tyrosine recombinase
Conserved protein	36.0	*5*		*5*	4.3	Hypothetical novel
Serine recombinase	61.1	*6*				
Conserved protein	8.8	*7*				
Hypothetical novel	5.7	*8*				
Hypothetical novel	14.0	*9*				
RecT-like protein	37.4	*10*	90.7	*6*	37.8	RecT-like protein
Nuclease/RecB-like protein	38.4	*11*	96.8	*7*	38.5	Nuclease/RecB-like protein
Transcriptional regulator	21.4	*13*	76.7	*8*	19.3	Transcriptional regulator
Conserved protein	12.8	*14*	55.9	*9*	14.2	Conserved protein
HNH endonuclease	18.7	*15*		*10*	13.5	Conserved protein
HNH endonuclease	28.5	*16*		*11*	31.0	Conserved DUF2303 protein
				12	9.5	Hypothetical novel
SSB protein	18.7	*17*	83.9	*13*	19.3	SSB protein
Conserved protein	7.9	*18*	76.4	*14*	8.1	Conserved protein
Conserved protein	16.8	*19*	38.9	*15*	18.0	Conserved protein
Conserved protein	18.1	*20*	86.1	*16*	18.0	Conserved protein
Conserved protein	9.7	*21*	96.7	*17*	9.7	Conserved protein
Hypothetical novel	19.2	*22*				
Conserved protein	17.0	*23*	85.0	*18*	16.1	Conserved protein
Conserved protein	8.6	*24*	89.7	*19*	8.6	Conserved protein
Methyltransferase	39.7	*25*				
Replication protein	28.7	*27*	51.0	*20*	30.6	Replication protein
DnaC-like helicase loader	30.4	*28*	87.9	*21*	30.4	DnaC-like helicase loader
				22	15.2	PagP-like OMP
Conserved protein	6.3	*29*	44.0	*23*	4.6	Conserved protein
Conserved protein	15.2	*30*		*24*	10.8	Hypothetical novel
				25	11.4	Hypothetical novel
Conserved protein	44.8	*31*	64.5	*26*	45.5	Conserved protein
Transposase, ISL3 family	55.3	*32*				
Conserved protein	14.1	*33*				
Conserved protein	7.4	*34*				
Conserved protein	20.6	*35*	76.4	*27*	20.3	Conserved protein
Terminase small subunit	28.7	*36*	52.4	*28*	30.3	Terminase small subunit
				29	13.2	Hypothetical novel
				30	28.9	Hypothetical novel
				31	19.0	Hypothetical novel
Conserved protein	16.6	*37*	43.9	*32*	15.7	Conserved protein
Conserved protein	11.5	*38*	56.0	*33*	11.2	Conserved protein
Conserved protein	7.5	*39*				
Conserved protein	7.7	*40*	64.7	*34*	7.8	Conserved protein
Conserved protein	6.9	*41*	85.3	*35*	7.3	Conserved protein
				36	6.6	Hypothetical novel
Conserved protein	5.2	*42*	59.1	*37*	5.5	Conserved protein
Hypothetical novel	5.8	*43*				
Transcriptional regulator	7.9	*44*	90.3	*38*	7.9	Transcriptional regulator
Conserved protein	7.2	*45*	72.6	*39*	7.5	Conserved protein
Conserved protein	5.4	*46*	44.1	*40*	6.8	Conserved protein, Sec signal
Hypothetical novel	7.2	*47*		*41*	35.1	Methyltransferase
Conserved protein	14.6	*48*		*42*	23.0	Type II restriction endonuclease
				43	29.5	DNA ligase
Terminase large subunit	60.9	*49*	92.4	*44*	61.2	Terminase large subunit
Conserved protein	16.2	*50*	93.9	*45*	16.3	Conserved protein
Portal protein	86.8	*51*	91.2	*46*	86.8	Portal protein
Conserved protein	34.0	*52*	77.6	*47*	34.6	Conserved protein
Conserved protein	11.8	*53*	15.1	*48*	12.1	Conserved protein
CsrA-like regulator	6.6	*54*	90.6	*49*	6.6	CsrA-like regulator
Major capsid protein	40.2	*55*	92.9	*50*	40.3	Major capsid protein
Conserved, virion associated	16.2	*56*	59.7	*51*	15.2	Conserved, virion associated
Conserved protein	22.7	*57*	78.9	*52*	22.1	Conserved protein
Conserved, virion associated	21.7	*58*	71.4	*53*	23.8	Conserved, virion associated
Conserved protein	23.7	*59*	87.8	*54*	23.9	Conserved protein

Continued on following page

TABLE 1—*Continued*

Phage Bcep22			% protein identity	Phage BcepIL02		
Putative function	Calculated mass (kDa)	Gene name		Gene name	Calculated mass (kDa)	Putative function
Conserved, virion associated	14.3	*60*	66.7	*55*	14.8	Conserved, virion associated
Conserved, virion associated	26.0	*61*	37.3	*56*	30.1	Conserved, virion associated
Pectin lyase-like protein	72.5	*62*				
Conserved, virion associated	9.9	*63*	45.4	*57*	10.1	Conserved, virion associated
Tail fiber	45.2	*64*	NA[b]	*58*	33.0	Tail fiber
Tail fiber	29.5	*65*	NA[b]	*59*	36.1	Tail fiber
				60	18.9	Conserved protein
Tail fiber	30.8	*66*	NA[b]	*61*	29.5	Tail fiber
				62	30.7	Tail fiber
Conserved, virion associated	62.5	*67*	92.2	*63*	62.9	Conserved, virion associated
Conserved, virion associated	12.5	*68*	69.5	*64*	12.7	Conserved, virion associated
Acyl-CoA N-acyltransferase	16.3	*69*	93.8	*65*	16.4	Acyl-CoA N-acyltransferase
Conserved, virion associated	35.2	*70*	88.8	*66*	35.1	Conserved, virion associated
PAPS reductase-like	27.6	*73*	93.6	*67*	27.3	PAPS reductase-like
				68	23.1	Novel, virion associated
Conserved, virion associated	56.1	*74*	81.0	*69*	55.9	Conserved, virion associated
DarB-like antirestriction protein	496.4	*75*	87.9	*70*	503.9	DarB-like antirestriction protein
Conserved protein	8.7	*76*	97.5	*71*	8.6	Conserved protein
Possible antiholin	11.3	*77*	98.0	*72*	11.3	Possible antiholin
Holin	7.0	*78*	81.2	*73*	8.3	Holin
SAR endolysin	19.3	*79*	93.1	*74*	19.3	SAR endolysin
Rz-like lysis protein	18.4	*80*	47.8	*75*	18.5	Rz-like lysis protein
Rz1-like lysis protein	7.7	*81*		*76*	7.2	Rz1-like lysis protein

[a] Putative assignable functions and protein molecular masses are shown. Protein identities were calculated by the Dice method using BLASTp alignment identity scores. OMP, outer membrane protein; SSB, single-stranded DNA-binding protein; CoA, coenzyme A; PAPS, phosphoadenylyl sulfate.
[b] NA, not applicable; the protein matched more than one protein in the other phage. See Fig. 3 and text for details.

style. Of phages that have been characterized more extensively, Bcep22 shares homology in eight of its proteins with temperate coliphage 933W (AF125520) and *Pseudomonas aeruginosa* phage F116 (AY625898). Phage 933W is a member of a closely related group of temperate phages associated with enterohemorrhagic *E. coli* strains, including phages VT2-Sakai (AP000363) and Stx2-converting phage 86 (AB255436). These 933W-like phages are themselves considered to be a subtype of lambdoid phage (11). Phage F116, like Bcep22, is somewhat of an "orphan" phage, bearing no clear relationship to other sequenced phages (44). Gene arrangements between Bcep22 and these related organisms are mostly syntenic, with genes encoding the terminase large subunit, portal, and major capsid proteins remaining conserved in all four, which is not entirely unexpected considering that these proteins must interact closely during phage assembly.

Tail fiber genes. A striking feature of the Bcep22 and BcepIL02 genomes is the presence of several tandemly repeated copies of putative tail fiber genes. The genome of Bcep22 contains three copies of these genes (genes *64*, *65*, and *66*), whereas BcepIL02 contains four copies (genes *58*, *59*, *61*, and *62*) (Fig. 2). As shown in Fig. 3, all of these genes show peptide

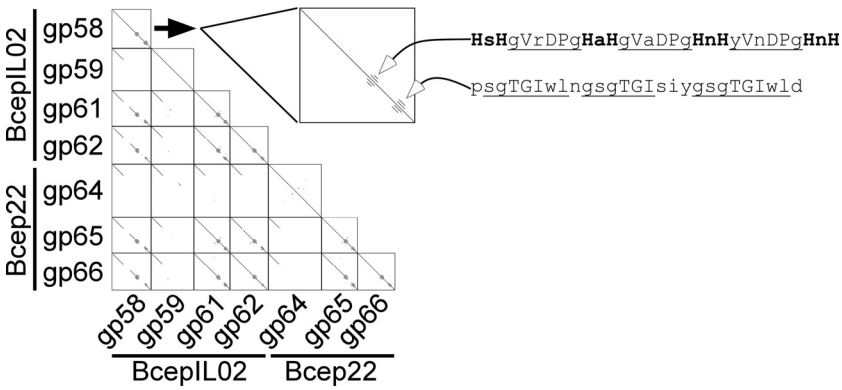

FIG. 3. Dot plot comparisons of the predicted redundant tail fiber proteins from phages Bcep22 and BcepIL02. Note the conservation of the N-terminal portions of these proteins and more-divergent C-terminal domains. The fine structures of the two repeat motifs present in five of the seven described proteins are shown, using an enlargement of BcepIL02 gp58 as an example. In the first motif (top), four histidine pairs (HxH; bold) are separated by partially conserved spacers (underlined); uppercase residues are conserved across all repeats in all five proteins. The second repeat motif (bottom) is less well conserved; repeats conserved in BcepIL02 gp58 are underlined, residues conserved across all repeats are shown in uppercase letters.

TABLE 2. Proteins determined to be virion-associated by proteomic analysis of CsCl-purified Bcep22 and BcepIL02 virions[a]

Bcep22		% identity[c]	BcepIL02		Putative function
Method(s)[b]	Protein		Protein	Method(s)	
G, W	**gp51**	91.2	**gp46**	W	Portal protein
G, W	**gp55**	92.9	**gp50**	G, W	Major capsid protein
W	**gp56**	59.7	**gp51**	G, W	Conserved phage protein
W	**gp58**	71.4	**gp53**	G, W	Conserved phage protein
	gp59	87.8	**gp54**	G, W	Conserved phage protein
	gp60	66.7	**gp55**	G, W	Conserved phage protein
	gp61	37.3	**gp56**	W	Conserved phage protein
G	**gp62**				Pectin lyase-like protein
			gp58	G, W	Tail fiber
			gp59	G, W	Tail fiber
			gp60	G	Conserved lectin-like protein
G	**gp67**	92.2	**gp63**	W	Conserved phage protein
	gp68	69.5	**gp64**	W	Conserved phage protein
	gp70	88.8	**gp66**	G, W	Conserved phage protein
			gp68	W	
G	**gp74**	81.0	**gp69**	W	Conserved phage protein
G, W	**gp75**	87.9	**gp70**	W	DarB-like antirestriction protein

[a] Proteins listed in boldface were directly detected in the proteomic analysis, while italicized protein names were inferred to be virion associated based on their similarity to a protein directly detected in the other phage.
[b] The method used for detection: W, whole phage digest; G, gel slice digest.
[c] Dice similarity between pairs of BcepIL02 and Bcep22 proteins.

sequence similarity to one another in their N-terminal domains and various levels of similarity in their C-terminal domains, suggesting that they arose from tandem gene duplications within each phage, followed by divergence of their C-terminal domains. Each phage appears to contain one "outlier" tail protein (Bcep22 gp64 and BcepIL02 gp59) that is not related in its C-terminal domain to any of the other tail fiber proteins in either phage. These proteins, being the only significantly diverged proteins among these putative tail fibers, may also be the determinants that allow phage BcepIL02 to infect both strain PC184 and AU1054 while limiting phage Bcep22 to its own host, strain AU1054. The presence of the divergent tail fibers may represent horizontal gene transfer events rather than simple sequence divergence. All of these BcepIL02 and Bcep22 putative tail fiber gene products show similarity in their N-terminal domains to gp14 of *Sinorhizobium meliloti* phage PBC5, which contains only a single copy of this gene. All of these proteins except for Bcep22 gp64 and BcepIL02 gp59 also have similarity, primarily in their C-terminal domains, to a putative tail protein in *Ralstonia* phage RSL1 (YP_001950074), a protein which also contains a phage tail collar domain (IPR011083). Three of these proteins (Bcep22 gp66 and BcepIL02 gp58 and gp62) contain detectable phage T4 gp12-like receptor-binding domains in their C termini (SSF88874). The T4 gp12 short tail fibers are responsible for the second, irreversible step of phage adsorption to its *E. coli* host, recognizing the core domain of bacterial lipopolysaccharide (64).

Of the five tail fiber proteins with C-terminal similarity (Bcep22 gp65 and gp66 and BcepIL02 gp58, gp61, and gp62), all contain two sets of conserved, tandemly repeating motifs in their C-terminal domains, representatives of which are shown in Fig. 3. One of the motifs present in each of these proteins is a set of four histidine pairs (HxH), each separated by a 6-residue spacer with the consensus sequence xVxDPx. Such HxH pairs are known to be involved in the coordination of metal ions in other phage tail fibers, such as T4 gp37 (2) and T4 gp12 (83). Their presence further supports the annotation of these proteins as tail fibers and suggests that they may form rod-like trimers, each trimer coordinating four metal ions. The other repeating motif is a more degenerate, glycine- and serine-rich region containing two to three 6- to 9-residue repeats. These repeats are more highly conserved intragenically than they are intergenically, with only the residues TGI conserved across all

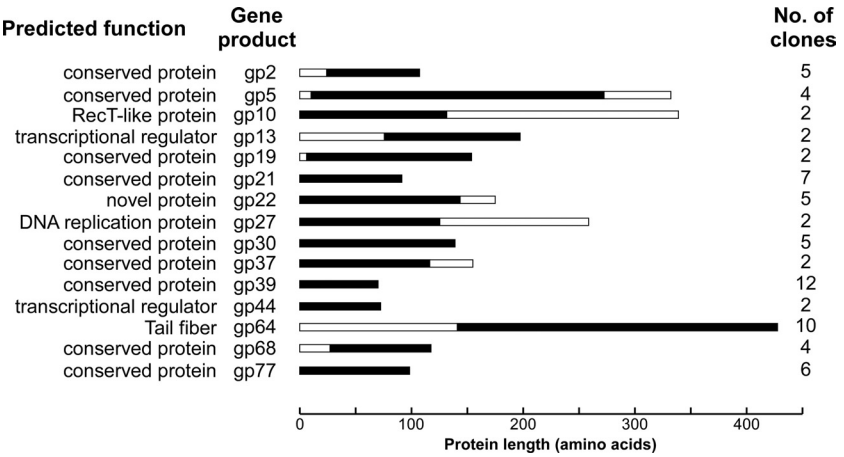

FIG. 4. Bcep22 proteins with detected homotypic interactions. Proteins are represented by white bars, which are shaded black in the regions where one or more interacting protein domains were determined. Corresponding Bcep22 protein names and predicted functions are listed on the left, and the number of clones recovered with inserts corresponding to each protein are listed on the right. All proteins are drawn to scale.

repeats. The positions of these repeats near the C terminus of each protein suggest that they may be involved in receptor interaction, but no clear function for these tandem repeats can be assigned currently.

While phages Bcep22 and BcepIL02 appear to contain multiple putative tail fiber protein orthologs, it is not clear how these multiple genes may function as tail fibers in the mature virion. The similarity of these proteins at the N terminus, which is commonly the domain by which the phage tail fiber attaches to the virion (45), suggests that all of these proteins are able to associate with a common capsid domain. Assembly of two to four different tail fiber proteins onto a single virion seems highly improbable, given the asymmetries this would impose on the virion's structure and, presumably, its ability to adsorb to the host cell. These genes are not arranged into any obvious transcriptional unit that would allow selective expression of single proteins in either phage, such as bounding by promoter/terminator pairs or an invertible DNA segment like that found in coliphage Mu (55). In the proteomic analysis of BcepIL02, two of these putative fibers, gp58 and gp59, were detected associated with the virion (Table 2), suggesting that at least two of the four orthologous proteins in this phage are functional and can participate in virion assembly. It is possible that despite the expression of several tail fiber proteins, the phages are able to exclusively assemble a single molecular species per virion, giving rise perhaps to a heterogeneous population of phages that are genetically identical but contain different tail fiber species assembled onto the completed virion. Alternatively, all proteins may be assembled together into a larger, multiprotein tail fiber analogous to those produced by the T-even phages, although this model seems less likely due to the presence of the conserved N termini and diverged C termini of these proteins and the lack of protein commonality between the two phages. If these proteins were capable of forming linearly ordered fiber structures, they would be expected to measure on the order of 20 to 40 nm in length, based on comparisons with the lengths of other phage tail fibers. However, no extended tail fibers were observed by transmission electron microscopy, although their presence cannot be ruled out definitively based solely on negative staining.

Mobile DNA and morons. The genome of Bcep22 contains two free-standing HNH-type homing endonucleases, genes *15* and *16*. These genes do not appear to be associated with introns or inteins, as they do not interrupt any apparent coding sequence, and the region containing these genes is flanked by two complete and highly conserved genes: a serine tRNA gene located 11 bp upstream of the start of Bcep22 gene *15*, and Bcep22 gene *17*, which encodes a single-stranded DNA-binding protein and overlaps Bcep22 gene *16* by 1 base. The genome of Bcep22 also contains one putative IS element, which contains Bcep22 genes *32*, *33*, and *34*. The putative transposase encoded by Bcep22 gene *32* is a member of the ISL*3* family (51). The IS element is bounded by imperfect inverted repeats (IR) 24 bp in length with 5 mismatches; these repeats are flanked by 8-bp direct repeats (see Fig. S1 in the supplemental material). IR-L is located between genes *31* and *32* and overlaps the C terminus of the predicted transposase by 10 bp. IR-R is located between genes *34* and *35* in a predicted noncoding region. The two other genes contained in this IS element, *33* and *34*, are conserved hypothetical genes. gp34 is closely related to gp55 of Bcep781 and gp54 of Bcep43, both of which are virulent phages that infect the related bacterium *Burkholderia cepacia* (76).

BcepIL02 contains one putative moron, gene *22*. This gene is not present in Bcep22, whereas the genes immediately preceding and following it have homologs in Bcep22 (genes *28* and *29*). Morons (or "more DNA") are horizontally acquired genes that are usually expressed while in the prophage state, presumably providing some fitness advantage to the lysogenized host (32). gp22 possesses weak (E, ~0.02) similarity to the PagP homologs of *Bordetella parapertussis* (NP_885888), *Bordetella pertussis* (NP_881581), and *Bordetella bronchiseptica* (NP_890716) and has an outer membrane β-barrel transmembrane domain (IPR011250) and a PagP-like lipid A acylation domain (IPR009746). In *Salmonella*, PagP adds a palmitate group to the lipid A component of lipopolysaccharide, resulting in increased resistance to cationic peptides (24). Such heptaacylated lipid A has also been shown to reduce endotoxin-associated inflammation (81) and therefore could be considered a virulence factor. The outer membrane localization of BcepIL02 gp22 is further supported by the presence of an N-terminal secretion signal and signal peptidase I cleavage site between residues 19 and 20. The DNA sequence immediately upstream of BcepIL02 *22* was found to contain a potential promoter signal [TTGCAA(N_{17})GATAAT(N_8)G], with the predicted transcript start located 27 bases upstream of the translational start. The nearest rho-independent terminator was detected immediately downstream of BcepIL02 gene *26* (Fig. 2), suggesting the *22* promoter initiates a transcript containing genes *22* through *26*. The presence of this gene in BcepIL02 is remarkable, as the currently sequenced phages and prophages of *B. cenocepacia* have been noted for their lack of virulence factors (75). If BcepIL02 is actually a temperate phage, the PagP-like gp22 would be one of the rare *Burkholderia* phage-encoded proteins, along with BcepMu gp53 (77), with a putative role in bacterial virulence.

DNA replication and packaging. In a comprehensive study of phage DNA replication modules, Weigel and Seitz (88) placed the DNA replication module of phage Bcep22 (putatively spanning genes *10* to *28*) in the "initiator-helicase loader" type in their classification system. Phage Bcep22 gp27 is the putative replication initiator protein, exhibiting similarity to several other annotated phage replication proteins, such as that of *Salmonella* phage Gifsy-1 (AAX65107) and containing a DNA-binding Cdc6 domain (cd08768) associated with DNA replication initiation. Bcep22 gp28 is a predicted DnaC-like helicase loader, based on its similarity to other DnaC-like proteins, such as that of the *E. coli* Rac prophage (NP_415878). Phage BcepIL02 gp20 and gp21 are similar to these Bcep22 proteins. Both Bcep22 and BcepIL02 also contain RecT/RecB-like gene pairs (Bcep22 gp10/gp11 and BcepIL02 gp6/gp7) that may possess recombination functions similar to exo and bet in coliphage lambda.

Bcep22 and BcepIL02 possess a similar (92.4%) terminase large subunit (TerL) homolog, gp49 in Bcep22 and gp44 in BcepIL02. These proteins are most closely related to the *terL* homologs found in *Pseudomonas aeruginosa* phage F116 (YP_164303) and *Sinorhizobium meliloti* phage PBC5 (NP_542306). Casjens et al. (12) found that although the Bcep22 and PBC5 TerL homologs could not be grouped conclusively with

other phage terminases with well-studied packaging mechanisms, their similarity to the terminase of phage T4 was suggestive of a headful mechanism. The closest-studied TerL homolog to this group is that of *E. coli* phage 933W (NP_049511), which appears to package its genome by a headful mechanism initiated from a diffuse packaging start site, producing terminally redundant, circularly permuted genomic DNA (62). Restriction digests of BcepIL02 genomic DNA yield a circular restriction map (data not shown), suggesting that DNA packaging in the phage is also initiated from a diffuse site and is circularly permuted, as expected for a headful mechanism.

The terminase small-subunit (TerS) homologs of Bcep22 and BcepIL02 were identified based on their homology to other annotated small-subunit terminases, such as that of phage F116 (YP_164302), and by the presence of a conserved terminase small-subunit domain (IPR005335). The genes encoding the putative TerS proteins of phages Bcep22 and BcepIL02 are located some distance away from their corresponding *terL* genes (Fig. 2). This arrangement is somewhat unusual, as *terS* genes are typically found adjacent to their *terL* counterparts, as described for numerous lambdoid phage types (11).

Lysis. Both phages Bcep22 and BcepIL02 contain closely related lysis cassettes with the canonical gene order of holin-endolysin-Rz/Rz1 (genes *78* to *81* and genes *73* to *76*, respectively). In both phages, the lysis genes appear to be arranged as part of a divergent transcriptional unit, with two conserved hypothetical genes upstream on the opposite strand (Bcep22 genes *76* and *77*; BcepIL02 genes *71* and *72*), bounded by rho-independent terminators (Fig. 2). The holins exhibit class II topology (two transmembrane domains, N-in, C-in) (87). Although the holin genes do not have a dual start motif, which in many phages allows the production of both holin and antiholin proteins from the same cistron, the smaller of the divergently transcribed genes in this cluster has a single transmembrane domain and thus might serve as an antiholin.

Phages Bcep22 and BcepIL02 encode endolysins, detected as *N*-acetylmuramoylhydrolase lysozymes (EC 3.2.1.17) of the same class as the soluble lysozyme found in phage T4, which contains a conserved EX_8D/CX_5T motif. However, the Bcep22 and BcepIL02 endolysins contain a motif of $EX_6EX_5DX_2T$, with two potentially catalytic Glu residues, neither of which are located at the −8 position from the Asp, and an atypically short spacing between the Asp and Thr residues. It is therefore difficult to determine which of these Glu residues might participate in the catalytic cleavage of peptidoglycan, but based on sequence alignments of the proteins with other known lysozymes, the first of the two residues appears to be the most likely possibility. Both the Bcep22 and BcepIL02 endolysins also contain unambiguous SAR (signal anchor release) domains. Endolysins with SAR domains are exported by the *sec* system and accumulate as enzymatically inactive periplasmic proteins tethered to the bilayer by the SAR domain (90). The presence of a SAR endolysin makes it possible that the Bcep22/BcepIL02 holins are pinholins, a class of holins that act by depolarizing the membrane rather than forming holes large enough to allow release of the endolysin, as in the case of canonical holins (59, 60). Unlike the SAR endolysin of phage P1, which utilizes cysteine isomerization for activation of the enzyme (89), the SAR endolysins of Bcep22 and BcepIL02 contain no cysteine residues in their SAR domains, suggesting that they are instead regulated by conformational changes associated with protein release from the bilayer (79). The final step in phage lysis, disruption of the outer membrane, is accomplished by the spanin protein (7, 74). Most phages encode two spanin subunits, the functional equivalents of the coliphage lambda Rz and Rz1 proteins. The Rz1 homologs of phages Bcep22 and BcepIL02, gp81 and gp76, respectively, belong to the embedded class of *Rz1* genes, in that the entire *Rz1*-equivalent reading frame is embedded in the +1 reading frame of the *Rz* equivalent.

P1 DarB homologs. Bcep22 and BcepIL02 encode strikingly large virion-associated proteins, gp75 (4,602 amino acids [aa]) and gp70 (4,667 aa), respectively (Tables 1 and 2), encoded by genes that occupy over 20% of the phage genomes. Based on sequence similarity and the presence and order of protein functional domains, these proteins appear to be functional equivalents of the 2,255-aa phage P1 DarB protein (YP_006479). In *E. coli*, P1 DarB is required for protecting P1 DNA from restriction by the host type I restriction system. Complementation analysis indicated that DarB operates only in *cis*, suggesting that it may bind to or modify the cognate phage DNA during or shortly after its ejection into the host cell (31). Like Bcep22 gp75 and BcepIL02 gp70, P1 DarB is virion associated (31) and contains a methyltransferase domain followed by a DExH helicase domain (see Fig. S2 in the supplemental material). It is not clear based on our analysis whether these virion-associated DarB-like proteins interact with DNA or RNA, but the association of these proteins with protection of DNA from host restriction would imply a protein-DNA interaction. In P1, *darB* function is dependent on *darA*, which is believed to possess a chaperone function for assembly of DarB into the virion (30). No *darA* homolog was detected in either BcepIL02 or Bcep22.

In addition to being considerably larger than P1 DarB, both gp75 and gp70 contain an N-terminal soluble lytic transglycosylase (SLT) domain that is absent from all other DarB-like proteins identified. A number of podophages, including P22, SP6, and K1-5 (53), and phage T7 (54) have been determined to carry virion-associated proteins that possess muralytic activity.

During analysis of Bcep22 gp75, BcepIL02 gp70, and P1 DarB, it became apparent that this type of large, multidomain helicase/methyltransferase protein is widely distributed across a variety of bacteria. Because the P1 *darB* and its Bcep22 and BcepIL02 homologs are located in phage genomes, the genomic contexts of other DarB-like genes were of interest. Using P1 DarB as the query sequence, 312 proteins greater than 1,500 amino acids in length were identified in the NCBI nr database at E values of $<10^{-5}$. Of these, 206 were located in the chromosomal sequences of incomplete or draft bacterial genomes and were excluded from further analysis, except for proteins in draft genomes unambiguously annotated as plasmids, which were retained. The resulting final collection contained 106 P1 DarB-like proteins. All of these were found to contain at least one methyltransferase domain (IPR002296, IPR006935, IPR013216, IPR002052, IPR011639, or SSF53335) and at least one helicase domain (IPR000330, IPR001650, IPR014021, or IPR014001), suggesting that these proteins provide some function similar to that provided by P1 DarB, possibly protection from host restriction by DNA methylation

TABLE 3. Selected homologs of the *E. coli* phage P1 DarB protein (YP_006479) and their genomic contexts, illustrating the association of these proteins with mobile DNA elements[a]

Source organism	E-value (vs. P1 darB)	Accession no.	Protein length (aa)	Taxonomic class	Gene location	Reference
Agrobacterium tumefaciens MAFF301001	1.9E−14	NP_053287	1,693	Alphaproteobacteria	Plasmid pTi-SAKURA	
Agrobacterium tumefaciens C58	2.0E−15	NP_396621	1,693	Alphaproteobacteria	Plasmid Ti	20
Bacillus cereus 03BB108	3.2E−20	ZP_03115564	2,366	Bacilli	Plasmid p03BB108_86	
Bacteroides fragilis YCH46	2.5E−12	YP_097425	1,938	Bacteroidetes	Conjugative transposon CTnYCH46-1	39
Burkholderia phage Bcep22	1.6E−14	NP_944303	4,602	Betaproteobacteria	Bacteriophage genome	
Burkholderia phage BcepIL02	1.2E−14	YP_002922742	4,667	Betaproteobacteria	Bacteriophage genome	
Burkholderia vietnamiensis G4	1.9E−132	YP_001109674	2,273	Betaproteobacteria	Plasmid pBVIE05	
Burkholderia vietnamiensis G4	1.5E−15	YP_001109960	1,853	Betaproteobacteria	Plasmid pBVIE04	
Campylobacter coli	1.2E−11	YP_063409	1,932	Epsilonproteobacteria	Plasmid pCC31	
Campylobacter coli RM2228	9.2E−11	ZP_00371038	1,854	Epsilonproteobacteria	Plasmid pCC178	
Campylobacter jejuni subsp. *jejuni* 81-176	6.6E−11	YP_247529	1,932	Epsilonproteobacteria	Plasmid pTet	
Clostridium difficile 630	1.6E−23	YP_001086891	2,907	Clostridia	CTn*2* conjugative transposon	69
Clostridium difficile 630	1.0E−22	YP_001088370	3,011	Clostridia	CTn*5* conjugative transposon	69
Enterococcus faecalis N03-0233	1.9E−24	ACQ89838	2,727	Bacilli	VanG2 mobile element	8
Enterococcus faecalis T1	2.8E−17	ZP_05423918	2,266	Bacilli	Plasmid unnamed	
Enterococcus faecalis V583	4.2E−17	NP_815963	3,173	Bacilli	*vanB* cassette (Tn*1549*-like)	61
Escherichia coli O111:H$^-$ 11128	0.0E +00	YP_003237917	2,255	Gammaproteobacteria	Plasmid pO111_2 (P1-like prophage)	58
Gordonia westfalica DSM44215	2.3E−11	NP_954816	1,630	Actinobacteria	Plasmid pKB1	
Helicobacter pylori Shi170	1.8E−14	ACF17793	2,609	Epsilonproteobacteria	Putative conjugative transposon plasticity zone TnPZ	36
Leptospira phage LE1	4.8E−20	CAE14681	1,971	Spirochaetes	Bacteriophage genome	
Methylobacterium extorquens DM4	2.5E−14	YP_003068159	1,698	Alphaproteobacteria	*dcm* genomic island	86
Nocardia farcinica IFM 10152	7.3E−19	YP_121997	2,153	Actinobacteria	Plasmid pNF1	
Ochrobactrum intermedium LMG 3301	2.9E−15	ZP_04683497	3,526	Alphaproteobacteria	Plasmid VBI00030_1	
Rhodococcus equi 103	1.5E−14	NP_066779	3,229	Actinobacteria	Plasmid p103	
Rhodococcus equi PAM1126	1.1E−14	YP_002149612	2,949	Actinobacteria	Plasmid pVAPA1037	46
Rhodococcus equi PAM1593	1.2E−14	YP_002149544	2,949	Actinobacteria	Plasmid pVAPB1593	46
Rhodococcus erythropolis PR4	9.8E−16	YP_345073	2,936	Actinobacteria	Plasmid pREC1	
Rhodococcus opacus B4	1.5E−15	YP_002784492	2,889	Actinobacteria	Plasmid pKNR	
Sinorhizobium phage PBC5	3.4E−17	NP_542283	2,849	Alphaproteobacteria	Bacteriophage genome	
Streptococcus agalactiae 2603V/R	2.7E−16	NP_688280	2,274	Bacilli	Putative conjugative transposon	82
Streptococcus dysgalactiae subsp. *equisimilis* NS3396	3.8E−16	ABV55437	2,278	Bacilli	Conjugative element ICESde3396	16
Streptococcus equi subsp. *equi* 4047	5.3E−23	YP_002746095	2,913	Bacilli	ICESe1	27
Streptococcus equi subsp. *zooepidemicus*	1.4E−16	YP_002745281	2,281	Bacilli	ICESz2	27
Streptococcus pneumoniae ATCC 700669	6.4E−18	CAR69082	2,088	Bacilli	Tn*5252* transposon	15
Streptococcus pneumoniae CGSP14	2.8E−17	YP_001836022	2,077	Bacilli	Tn*2008*	17
Streptococcus pneumoniae G54	2.3E−16	YP_002037980	2,074	Bacilli	Tn*5253*	18
Streptococcus suis GZ1	2.5E−20	ADE30961	2,554	Bacilli	Acquired island region AI 33, antimicrobial resistance	91
Vibrio sp. 0908	0.0E +00	YP_001595675	2,349	Gammaproteobacteria	Plasmid p0908	
Vibrio vulnificus YJ016	7.7E−18	NP_936882	4,392	Gammaproteobacteria	933W-like prophage element	

[a] E values were obtained from BLASTp results for P1 DarB versus the NCBI nr database. References supporting the genomic context in chromosomally located DarB homologs are provided.

(47). Interestingly, the large majority of these proteins, 84%, could be found associated with some form of mobile DNA element: 54 (51%) were located in plasmids, 29 (27%) were located within previously described chromosomal mobile DNA elements (variously documented as conjugative transposons, insertion sequences, integrative conjugative elements, or genomic islands), and 6 (6%) were located in bacteriophage genomes. A selected subset of these P1 DarB homologs is shown in Table 3, and a complete listing of all 106 proteins is presented in Table S3 of the supplemental

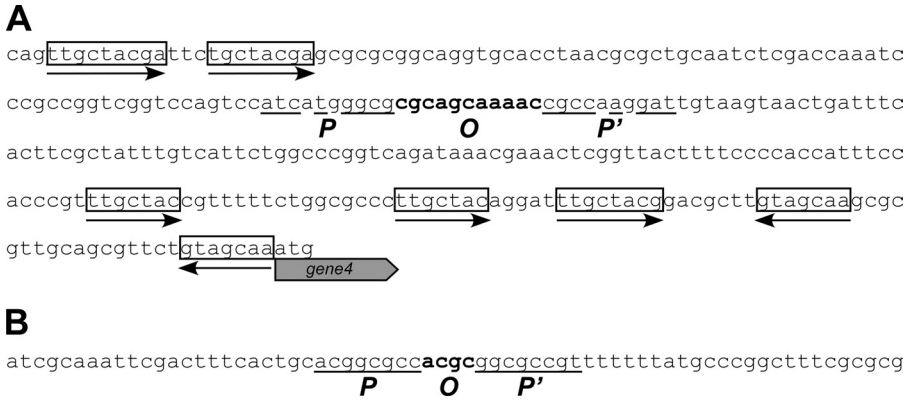

FIG. 5. The putative *attP* sites of phages BcepIL02 (A) and Bcep22 (B). (A) The 288 bp of DNA sequence upstream of BcepIL02 gene *4*, a predicted tyrosine recombinase. The 11-bp *attP* core site (labeled *O*) is written in bold and is flanked by the 10-bp imperfect complementary repeats of *P* and *P′*; complementary bases are underlined. This region was also found to contain seven 7- to 9-bp repeat motifs (shown boxed) with the core common sequence TGCTAC, and are putative *arm*-type recombinase-binding sites. Five sites were found on the coding DNA strand and two on the reverse strand; arrows below the boxes indicate arm site orientation. The final three bases of the displayed sequence are the start codon of the recombinase gene *4*. (B) The putative *attP* site of Bcep22. The 4-bp *O* core site (bold) is centered 126 bp upstream of the start of the predicted Bcep22 serine recombinase gene *6* (data not shown). *O* is flanked by 8-bp perfect complementary repeats *P* and *P′*; complementary bases are underlined.

material. Some of the mobile DNA regions containing DarB-like proteins are elements associated with important virulence factors (e.g., *Agrobacterium tumefaciens* plasmid Ti [20], *Rhodococcus equi* plasmid VapA1037 [46]), antibiotic resistance (such as the *Enterococcus faecalis vanB* cassette [61], or *Streptococcus suis* island AI 33 [91]), or phenotypic conversion (*Methylobacterium extorquens* dichloromethane utilization island *dcm* [86]).

Given the relatively large size of the proteins (ca. 1,500 to 4,600 amino acids) and their conserved domain structures, it would appear that these P1 *darB*-like genes are not merely hitchhiking DNA but provide some selective benefit to the mobile elements with which they are associated. The most likely mechanism of this benefit would be analogous to the proposed function of P1 DarB, wherein the mobile DNA is methylated during or shortly after transfer to a new cell in order to protect it from host restriction. If so, this would make these DarB-like proteins a novel class of widely distributed antirestriction proteins, similar in effect (if not in mechanism) to proteins such as ArdA, ArdB, and ArdC (4, 5, 13). The presence of these proteins in phages, plasmids, and IS elements also highlights the potential for horizontal genetic transfer, not only between phages (as has been well documented [26]) but also between phages and other bacterial genetic elements.

Lifestyle of phages Bcep22 and BcepIL02. Phage BcepIL02 was initially reported as a potentially therapeutic phage, showing efficacy against human pathogenic *B. cepacia* complex in a mouse lung model (10). Based on the common criteria for determining phage virulence, such as clear plaque morphology and the inability to form stable lysogens, BcepIL02 is to all outward appearances a virulent phage. Closer examination of the phage genome and the interaction between the phage and its host raised doubts about this classification. It has become clear that phage virulence is not a characteristic that can always be easily determined: a phage definitively can be proven as temperate by its ability to form stable lysogens, but the absence of stable lysogen formation is not necessarily diagnostic for virulence.

(i) Genomic evidence for a temperate lifestyle. The genomes of both Bcep22 and BcepIL02 encode predicted recombinases, Bcep22 gp6 and BcepIL02 gp4. The presence of such recombinases in phage genomes is typically cited as indicative of a temperate phage lifestyle, as proteins of this type often function as phage integrases. BcepIL02 encodes a predicted tyrosine recombinase, gp4. Based on BLASTp alignments to its closest *Caudovirales* neighbors, the catalytic tyrosine of BcepIL02 gp4 is predicted to be Y343 (corresponding to Y342 in lambda *int*), which is located 31 residues downstream from the conserved HxxR motif and 9 residues downstream from a conserved histidine residue, which correspond to lambda *int* H308, R311, and H333, respectively (23). The 402-bp region between genes *3* and *4* of the BcepIL02 genome was analyzed for structural features characteristic of a potential tyrosine recombinase-associated *attP* site. A 10-bp imperfect complementary repeat, separated by 11 bp, was found 172 bases upstream of the gene *4* start, colocalized with seven direct repeats that may serve as *arm* sites for phage integrase binding (Fig. 5A). The structure of this region is analogous to the *attP* region of phage lambda, which also utilizes a tyrosine recombinase for integration.

In Bcep22, the putative serine recombinase gp6 is highly related (60% identity) to a predicted serine recombinase gp53 (ABA60054; E = 0) of phage Bcep176 (NC_007497), a known functional temperate bacteriophage located within chromosome 2 of *Burkholderia multivorans* strain ATCC 17616 (AP009386; nucleotides 1620484 to 1665345). Based on the arrangement of the Bcep176 prophage in strain 17616, the gp53 recombinase is located 69 bp downstream from the predicted phage *attP* site and becomes the first gene at the extreme left end of the prophage. This location is typical of phage integrases in the prophage state and provides additional support that this serine recombinase functions as a phage integrase in Bcep176 and, due to its high sequence similarity to

Bcep22 gp6, is expected to provide the same function in Bcep22. In the Bcep22 genome, a potential *attP* site is located upstream of gene *6*, comprised of a pair of perfect 8-bp complementary repeats, spaced 4 bp apart (Fig. 5B). This pair of phages represents an unusual example of phage modular evolution, as Bcep22 has apparently substituted a serine recombinase for the tyrosine recombinase found in BcepIL02 that is typically associated with lambdoid phages.

Both phages Bcep22 and BcepIL02 also contain putative repressor proteins, gp13 in Bcep22 and gp8 in BcepIL02 (Fig. 2; Table 1), which may function as lysogenic repressors. Bcep22 gp13 contains an HTH, lambda-like DNA-binding domain at its N terminus (IPR001387, IPR010982, CATH1.10.260.40) and was also shown to dimerize via its C-terminal domain by homotypic interaction analysis. BcepIL02 gp8 is overall 77% identical to Bcep22 gp13 and 90% identical in its first 100 residues, the location of the HTH domain, and also contains a detectable DNA-binding domain (CATH1.10.260.40).

(ii) Phages Bcep22 and BcepIL02 are unable to form lysogens. To examine the potential for lysogeny, 16 phage-insensitive strains were isolated following challenge by phage Bcep22 or BcepIL02. Phage BcepIL02 exhibited EOPs ranging from 10^{-3} to 10^{-4} when plated on the BcepIL02-insensitive mutants, whereas Bcep22 exhibited EOPs of 10^{-3} to 10^{-4} ($n = 2$) or $<10^{-6}$ ($n = 6$) on the Bcep22-insensitive mutants. Additionally, the three AU0728 strains isolated during *in vivo* challenge (10) exhibited EOPs of ca. 10^{-4}, compared to plating on the wild-type stain. Culture supernatants of these insensitive strains were plated against lawns of the parental wild-type strains; none of the insensitive strains produced a phage capable of producing plaques on the wild-type parental strains, suggesting that these strains are not insensitive due to lysogeny but due to some other mechanism, such as resistance due to the alteration of a phage receptor. One of the PC184-derived insensitive strains was tested for its ability to adsorb phage BcepIL02, and the phage adsorption rate was found to be severely reduced compared to adsorption to the wild-type strain (data not shown), suggesting insensitivity due to some phage adsorption defect.

To further investigate the phage lifestyle, an alternative experimental approach was used to specifically examine the interaction of phage BcepIL02 and its host PC184. This system was examined because phage BcepIL02 already has been studied as a potentially therapeutic phage (10), and therefore its status as a virulent phage is of greater immediate interest. High MOI experiments were conducted by incubating cells of strain PC184 with a large excess of phage and enumerating the bacterial survivors. It previously has been demonstrated that the temperate coliphage lambda lysogenizes its host cell at a frequency that is strongly dependent on the number of virions adsorbed by a given cell, both in measurements of bulk culture (38) and at the single-cell level (92). In bulk systems, infections conducted at higher MOIs result in a greater frequency of lysogen formation, approaching 100% of cells lysogenized as the actual MOI reaches 5.

In the experiments reported here, after 30 min of incubation of cells with phage at an input MOI of ~10, actual MOIs of 5.7 to 6.4 were achieved (Table 4). If BcepIL02 were a virulent phage with no potential for becoming repressed during the infection cycle, a single round of infection at a real MOI of ~6 should result in a 1,000-fold reduction in CFU following challenge. As shown in Table 4, approximately 10 to 20% of the cells survived phage exposure, 36- to 133-fold greater than predicted. Twenty-four surviving colonies from these experiments were isolated, subcultured, and tested for sensitivity to the phage and phage production from culture supernatants as described previously. All 24 isolates exhibited EOPs equivalent to the wild-type strain, and none produced any detectable phage in their culture supernatants, indicating that these strains did not survive phage challenge due to either stable lysogeny or mutation to resistance.

TABLE 4. Actual MOI and calculated and actual bacterial survivors of *B. cenocepacia* strain PC184 following exposure to phage BcepIL02[a]

Replicate no.	Actual MOI	Predicted % surviving cells	Measured % surviving cells	Fold excess of survivors vs. prediction
1	5.7	0.32	11.7	36
2	6.3	0.19	17.6	94
3	6.4	0.16	21.3	133

[a] Predicted survivors were calculated from the Poisson distribution for the measured actual MOIs. Results of three replicate experiments are shown.

To interpret these results and reconcile them with the genomic observations above, we propose that phage BcepIL02 and, by extension, Bcep22, may be temperate phages that are unable to successfully integrate into the host chromosome. Thus, repression can be established, thereby allowing a significant fraction of infected cells to survive, but stable prophages cannot be formed. In this scheme, the unintegrated repressed prophage is lost upon cell division. This situation may be analogous to the behavior of phage lambda *int* mutants, which are able to form abortive lysogens that give rise to sensitive cells upon subculture (19, 21). In a study of a known temperate *B. cenocepacia* phage, KS9, only 18% of the phage-insensitive strains isolated following phage challenge were true lysogens, the remainder becoming insensitive presumably by some other mechanism (49). Phages Bcep22 and BcepIL02 may be a more extreme version of this behavior, in which lysogen formation, if it occurs, is below the rate of bacterial mutation to phage insensitivity. The predicted *attP O* sites of both Bcep22 and BcepIL02 are not found in the genomes of the hosts used to isolate them (AU1054 and PC184, respectively). In Bcep22, regions that exactly match the predicted *attP* site were located in *Burkholderia vietnamensis* G4 chromosome 3 (CP000616; positions 203779 to 203799) and also *Edwardsiella tarda* (CP002154; positions 1524102 to 1524127), suggesting that the natural lysogenic hosts of these phages may not be *B. cenocepacia* at all but instead other soil microorganisms.

Conclusions. Due to frequent horizontal gene transfer and the mosaic nature of phage genomes, bacteriophages pose a major challenge to meaningful taxonomy. Phage morphology, a long-standing taxonomic criterion, is of limited usefulness in this regard, prominently exemplified by the relatedness of lambda, a siphophage, and P22, a podophage. A functional taxonomic approach, focusing on protein similarity, has produced some biologically meaningful classification of podophages (43) and myophages (44). This method, however, has thus far remained confined to comparisons of phages within a given

morphotype and has not thoroughly examined intermorphotype relationships. A useful grouping, "phage type," has been advanced by Casjens (11), based on sequence similarity of individual proteins in addition to similarities in genome size, gene order, and transcriptional pattern, irrespective of phage morphology. Using the direct criteria of protein similarity, phages Bcep22 and BcepIL02 form a novel phage group with no evident relationships to other phages. Using the phage type system, these phages can be seen to share some characteristics with the lambdoid phages, such as the organization of the DNA replication genes and the similarity of the TerL, portal, and major capsid proteins to those of the lambdoid phage 933W. Even under these expanded criteria, however, these phages lack significant relationships to any other characterized phages and are characterized instead by physically uncoupled *terS*/*terL* gene pairs, unusual clusters of tail fiber orthologs, and the presence of a unique DarB-like protein with an N-terminal SLT domain. Due to their inability to form lysogens in the hosts used here, it is not clear if these phages are temperate, but their genome organization and behavior when infected under high-MOI conditions strongly suggest that they are temperate phages. The fact that BcepIL02 was able to produce significant treatment effects in the mouse lung model (10) may be due to its behavior as a *de facto* virulent phage in at least certain strains, unable to give rise to homoimmune lysogens in the host population. Resurgent interest in the use of phages for the treatment of bacterial infections, and the general consensus that temperate phages are not optimal candidates for use in phage therapy, has brought into focus the issue of determining the lifestyle of novel phages as either temperate or virulent. On the basis of simple experiments, such as plaque morphology or the isolation of phage-insensitive strains, these phages would appear to be virulent. Upon further experimentation and genomic analysis, however, these phages exhibit all the hallmarks of temperate phages, save for their ability to produce stable lysogens. The behavior of BcepIL02 is a topic deserving of further study as a means to define the working criteria for virulence and temperance in the context of phage therapy.

ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health (grant R01-AI64512), by the National Science Foundation (grants EF 0523951 and EF 0949351), and by funding provided to the Center for Phage Technology, an Initial University Multidisciplinary Research Initiative of Texas A&M University. J.J.L. is supported by the Cystic Fibrosis Foundation.

We thank Marilia Cordova for her assistance in sequencing the Bcep22 genome.

REFERENCES

1. **Adams, M. H.** 1959. Bacteriophages. Interscience Publishers, New York, NY.
2. **Bartual, S. G., et al.** 2010. Structure of the bacteriophage T4 long tail fiber receptor-binding tip. Proc. Natl. Acad. Sci. U. S. A. **107:**20287–20292.
3. **Bell, C. E., P. Frescura, A. Hochschild, and M. Lewis.** 2000. Crystal structure of the lambda repressor C-terminal domain provides a model for cooperative operator binding. Cell **101:**801–811.
4. **Belogurov, A. A., et al.** 2000. Antirestriction protein Ard (type C) encoded by IncW plasmid pSa has a high similarity to the "protein transport" domain of TraC1 primase of promiscuous plasmid RP4. J. Mol. Biol. **296:**969–977.
5. **Belogurov, A. A., E. P. Delver, and O. V. Rodzevich.** 1993. Plasmid pKM101 encodes two nonhomologous antirestriction proteins (ArdA and ArdB) whose expression is controlled by homologous regulatory sequences. J. Bacteriol. **175:**4843–4850.
6. **Bendtsen, J. D., H. Nielsen, G. von Heijne, and S. Brunak.** 2004. Improved prediction of signal peptides: SignalP 3.0. J. Mol. Biol. **340:**783–795.
7. **Berry, J., E. J. Summer, D. K. Struck, and R. Young.** 2008. The final step in the phage infection cycle: the Rz and Rz1 lysis proteins link the inner and outer membranes. Mol. Microbiol. **70:**341–351.
8. **Boyd, D. A., et al.** 2006. VanG-type vancomycin-resistant *Enterococcus faecalis* strains isolated in Canada. Antimicrob. Agents Chemother. **50:**2217–2221.
9. **Camacho, C., et al.** 2009. BLAST+: architecture and applications. BMC Bioinformatics **10:**421.
10. **Carmody, L. A., et al.** 2010. Efficacy of bacteriophage therapy in a model of *Burkholderia cenocepacia* pulmonary infection. J. Infect. Dis. **201:**264–271.
11. **Casjens, S. R.** 2008. Diversity among the tailed-bacteriophages that infect the Enterobacteriaceae. Res. Microbiol. **159:**340–348.
12. **Casjens, S. R., et al.** 2005. The generalized transducing *Salmonella* bacteriophage ES18: complete genome sequence and DNA packaging strategy. J. Bacteriol. **187:**1091–1104.
13. **Chilley, P. M., and B. M. Wilkins.** 1995. Distribution of the *ardA* family of antirestriction genes on conjugative plasmids. Microbiology **141:**2157–2164.
14. **Coenye, T., and P. Vandamme.** 2003. Diversity and significance of *Burkholderia* species occupying diverse ecological niches. Environ. Microbiol. **5:**719–729.
15. **Croucher, N. J., et al.** 2009. Role of conjugative elements in the evolution of the multidrug-resistant pandemic clone *Streptococcus pneumoniae* Spain 23F ST81. J. Bacteriol. **191:**1480–1489.
16. **Davies, M. R., J. Shera, G. H. Van Domselaar, K. S. Sriprakash, and D. J. McMillan.** 2009. A novel integrative conjugative element mediates genetic transfer from group G *Streptococcus* to other β-hemolytic streptococci. J. Bacteriol. **191:**2257–2265.
17. **Ding, F., et al.** 2009. Genome evolution driven by host adaptations results in a more virulent and antimicrobial-resistant *Streptococcus pneumoniae* serotype 14. BMC Genomics **10:**158.
18. **Dopazo, J., et al.** 2001. Annotated draft genomic sequence from a *Streptococcus pneumoniae* type 19F clinical isolate. Microb. Drug Resist. **7:**99–125.
19. **Gingery, R., and H. Echols.** 1967. Mutants of bacteriophage lambda unable to integrate into the host chromosome. Proc. Natl. Acad. Sci. U. S. A. **58:**1507–1514.
20. **Goodner, B., et al.** 2001. Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58. Science **294:**2323–2328.
21. **Gottesman, M. E., and M. B. Yarmolinsky.** 1968. Integration-negative mutants of bacteriophage lambda. J. Mol. Biol. **31:**487–505.
22. **Goudie, A. D., et al.** 2008. Genomic sequence and activity of KS10, a transposable phage of the *Burkholderia cepacia* complex. BMC Genomics **9:**615.
23. **Groth, A. C., and M. P. Calos.** 2004. Phage integrases: biology and applications. J. Mol. Biol. **335:**667–678.
24. **Guo, L., et al.** 1998. Lipid A acylation and bacterial resistance against vertebrate antimicrobial peptides. Cell **95:**189–198.
25. **Hall, S. D., and R. D. Kolodner.** 1994. Homologous pairing and strand exchange promoted by the *Escherichia coli* RecT protein. Proc. Natl. Acad. Sci. U. S. A. **91:**3205–3209.
26. **Hendrix, R. W.** 2003. Bacteriophage genomics. Curr. Opin. Microbiol. **6:**506–511.
27. **Holden, M. T., et al.** 2009. Genomic evidence for the evolution of *Streptococcus equi*: host restriction, increased virulence, and genetic exchange with human pathogens. PLoS Pathog. **5:**e1000346.
28. **Holden, M. T., et al.** 2009. The genome of *Burkholderia cenocepacia* J2315, an epidemic pathogen of cystic fibrosis patients. J. Bacteriol. **191:**261–277.
29. **Hunter, S., et al.** 2009. InterPro: the integrative protein signature database. Nucleic Acids Res. **37:**D211–D215.
30. **Iida, S., R. Hiestand-Nauer, H. Sandmeier, H. Lehnherr, and W. Arber.** 1998. Accessory genes in the *darA* operon of bacteriophage P1 affect antirestriction function, generalized transduction, head morphogenesis, and host cell lysis. Virology **251:**49–58.
31. **Iida, S., M. B. Streiff, T. A. Bickle, and W. Arber.** 1987. Two DNA antirestriction systems of bacteriophage P1, *darA*, and *darB*: characterization of *darA* phages. Virology **157:**156–166.
32. **Juhala, R. J., et al.** 2000. Genomic sequences of bacteriophages HK97 and HK022: pervasive genetic mosaicism in the lambdoid bacteriophages. J. Mol. Biol. **299:**27–51.
33. **Juncker, A. S., et al.** 2003. Prediction of lipoprotein signal peptides in Gram-negative bacteria. Protein Sci. **12:**1652–1662.
34. **Kaguni, J. M.** 2006. DnaA: controlling the initiation of bacterial DNA replication and more. Annu. Rev. Microbiol. **60:**351–375.
35. **Kasman, L. M., et al.** 2002. Overcoming the phage replication threshold: a mathematical model with implications for phage therapy. J. Virol. **76:**5557–5564.
36. **Kersulyte, D., et al.** 2009. *Helicobacter pylori*'s plasticity zones are novel transposable elements. PLoS One **4:**e6859.
37. **Kingsford, C. L., K. Ayanbule, and S. L. Salzberg.** 2007. Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. Genome Biol. **8:**R22.
38. **Kourilsky, P.** 1973. Lysogenization by bacteriophage lambda. I. Multiple infection and the lysogenic response. Mol. Gen. Genet. **122:**183–195.

39. **Kuwahara, T., et al.** 2004. Genomic analysis of *Bacteroides fragilis* reveals extensive DNA inversions regulating cell surface adaptation. Proc. Natl. Acad. Sci. U. S. A. **101:**14919–14924.
40. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:**680–685.
41. **Langley, R., D. T. Kenna, P. Vandamme, R. Ure, and J. R. Govan.** 2003. Lysogeny and bacteriophage host range within the *Burkholderia cepacia* complex. J. Med. Microbiol. **52:**483–490.
42. **Lavigne, R., P. J. Ceyssens, and J. Robben.** 2009. Phage proteomics: applications of mass spectrometry. Methods Mol. Biol. **502:**239–251.
43. **Lavigne, R., et al.** 2009. Classification of Myoviridae bacteriophages using protein sequence similarity. BMC Microbiol. **9:**224.
44. **Lavigne, R., D. Seto, P. Mahadevan, H. W. Ackermann, and A. M. Kropinski.** 2008. Unifying classical and molecular taxonomic classification: analysis of the Podoviridae using BLASTp-based tools. Res. Microbiol. **159:**406–414.
45. **Leiman, P. G., et al.** 2010. Morphogenesis of the T4 tail and tail fibers. Virol. J. **7:**355.
46. **Letek, M., et al.** 2008. Evolution of the *Rhodococcus equi vap* pathogenicity island seen through comparison of host-associated *vapA* and *vapB* virulence plasmids. J. Bacteriol. **190:**5797–5805.
47. **Lobocka, M. B., et al.** 2004. Genome of bacteriophage P1. J. Bacteriol. **186:**7032–7068.
48. **Lukashin, A. V., and M. Borodovsky.** 1998. GeneMark.hmm: new solutions for gene finding. Nucleic Acids Res. **26:**1107–1115.
49. **Lynch, K. H., K. D. Seed, P. Stothard, and J. J. Dennis.** 2010. Inactivation of *Burkholderia cepacia* complex phage KS9 gp41 identifies the phage repressor and generates lytic virions. J. Virol. **84:**1276–1288.
50. **Mahenthiralingam, E., T. A. Urban, and J. B. Goldberg.** 2005. The multifarious, multireplicon *Burkholderia cepacia* complex. Nat. Rev. Microbiol. **3:**144–156.
51. **Mahillon, J., and M. Chandler.** 1998. Insertion sequences. Microbiol. Mol. Biol. Rev. **62:**725–774.
52. **Marino-Ramirez, L., and J. C. Hu.** 2002. Isolation and mapping of self-assembling protein domains encoded by the *Saccharomyces cerevisiae* genome using lambda repressor fusions. Yeast **19:**641–650.
53. **Moak, M., and I. J. Molineux.** 2004. Peptidoglycan hydrolytic activities associated with bacteriophage virions. Mol. Microbiol. **51:**1169–1183.
54. **Moak, M., and I. J. Molineux.** 2000. Role of the Gp16 lytic transglycosylase motif in bacteriophage T7 virions at the initiation of infection. Mol. Microbiol. **37:**345–355.
55. **Morgan, G. J., G. F. Hatfull, S. Casjens, and R. W. Hendrix.** 2002. Bacteriophage Mu genome sequence: analysis and comparison with Mu-like prophages in *Haemophilus*, *Neisseria* and *Deinococcus*. J. Mol. Biol. **317:**337–359.
56. **Muyrers, J. P., Y. Zhang, F. Buchholz, and A. F. Stewart.** 2000. RecE/RecT and Redα/Redβ initiate double-stranded break repair by specifically interacting with their respective partners. Genes Dev. **14:**1971–1982.
57. **Nzula, S., P. Vandamme, and J. R. Govan.** 2000. Sensitivity of the *Burkholderia cepacia* complex and *Pseudomonas aeruginosa* to transducing bacteriophages. FEMS Immunol. Med. Microbiol. **28:**307–312.
58. **Ogura, Y., et al.** 2009. Comparative genomics reveal the mechanism of the parallel evolution of O157 and non-O157 enterohemorrhagic *Escherichia coli*. Proc. Natl. Acad. Sci. U. S. A. **106:**17939–17944.
59. **Pang, T., C. G. Savva, K. G. Fleming, D. K. Struck, and R. Young.** 2009. Structure of the lethal phage pinhole. Proc. Natl. Acad. Sci. U. S. A. **106:**18966–18971.
60. **Park, T., D. K. Struck, C. A. Dankenbring, and R. Young.** 2007. The pinholin of lambdoid phage 21: control of lysis by membrane depolarization. J. Bacteriol. **189:**9135–9139.
61. **Paulsen, I. T., et al.** 2003. Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. Science **299:**2071–2074.
62. **Plunkett, G., III, D. J. Rose, T. J. Durfee, and F. R. Blattner.** 1999. Sequence of Shiga toxin 2 phage 933W from *Escherichia coli* O157:H7: Shiga toxin as a phage late-gene product. J. Bacteriol. **181:**1767–1778.
63. **Projan, S. J.** 2003. Why is big pharma getting out of antibacterial drug discovery? Curr. Opin. Microbiol. **6:**427–430.
64. **Riede, I.** 1987. Receptor specificity of the short tail fibres (gp12) of T-even type *Escherichia coli* phages. Mol. Gen. Genet. **206:**110–115.
65. **Rosas-Acosta, G., W. K. Russell, A. Deyrieux, D. H. Russell, and V. G. Wilson.** 2005. A universal strategy for proteomic studies of SUMO and other ubiquitin-like modifiers. Mol. Cell Proteomics **4:**56–72.
66. **Rutherford, K., et al.** 2000. Artemis: sequence visualization and annotation. Bioinformatics **16:**944–945.
67. **Sambrook, J., E. F. Fritsch, and T. Maniatis (ed.).** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
68. **Schattner, P., A. N. Brooks, and T. M. Lowe.** 2005. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. Nucleic Acids Res. **33:**W686–W689.
69. **Sebaihia, M., et al.** 2006. The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. Nat. Genet. **38:**779–786.
70. **Seed, K. D., and J. J. Dennis.** 2009. Experimental bacteriophage therapy increases survival of *Galleria mellonella* larvae infected with clinically relevant strains of the *Burkholderia cepacia* complex. Antimicrob. Agents Chemother. **53:**2205–2208.
71. **Seed, K. D., and J. J. Dennis.** 2005. Isolation and characterization of bacteriophages of the *Burkholderia cepacia* complex. FEMS Microbiol. Lett. **251:**273–280.
72. **Sulakvelidze, A., and E. Kutter.** 2005. Bacteriophage therapy in humans, p. 510. *In* E. Kutter and A. Sulakvelidze (ed.), Bacteriophages: biology and applications. CRC Press, Boca Raton, FL.
73. **Summer, E. J.** 2009. Preparation of a phage DNA fragment library for whole genome shotgun sequencing. Methods Mol. Biol. **502:**27–46.
74. **Summer, E. J., et al.** 2007. Rz/Rz1 lysis gene equivalents in phages of Gram-negative hosts. J. Mol. Biol. **373:**1098–1112.
75. **Summer, E. J., J. J. Gill, C. Upton, C. F. Gonzalez, and R. Young.** 2007. Role of phages in the pathogenesis of *Burkholderia*, or "where are the toxin genes in *Burkholderia* phages?" Curr. Opin. Microbiol. **10:**410–417.
76. **Summer, E. J., et al.** 2006. Divergence and mosaicism among virulent soil phages of the *Burkholderia cepacia* complex. J. Bacteriol. **188:**255–268.
77. **Summer, E. J., et al.** 2004. *Burkholderia cenocepacia* phage BcepMu and a family of Mu-like phages encoding potential pathogenesis factors. J. Mol. Biol. **340:**49–65.
78. **Summers, W. C.** 2001. Bacteriophage therapy. Annu. Rev. Microbiol. **55:**437–451.
79. **Sun, Q., et al.** 2009. Regulation of a muralytic enzyme by dynamic membrane topology. Nat. Struct. Mol. Biol. **16:**1192–1194.
80. **Takeda, Y., et al.** 1986. Different interactions used by Cro repressor in specific and nonspecific DNA binding. J. Biol. Chem. **261:**8608–8616.
81. **Tanamoto, K., and S. Azumi.** 2000. *Salmonella*-type heptaacylated lipid A is inactive and acts as an antagonist of lipopolysaccharide action on human line cells. J. Immunol. **164:**3149–3156.
82. **Tettelin, H., et al.** 2002. Complete genome sequence and comparative genomic analysis of an emerging human pathogen, serotype V *Streptococcus agalactiae*. Proc. Natl. Acad. Sci. U. S. A. **99:**12391–12396.
83. **Thomassen, E., et al.** 2003. The structure of the receptor-binding domain of the bacteriophage T4 short tail fibre reveals a knitted trimeric metal-binding fold. J. Mol. Biol. **331:**361–373.
84. **Valentine, R. C., B. M. Shapiro, and E. R. Stadtman.** 1968. Regulation of glutamine synthetase. XII. Electron microscopy of the enzyme from *Escherichia coli.* Biochemistry **7:**2143–2152.
85. **Vandamme, P., and P. Dawyndt.** 2011. Classification and identification of the *Burkholderia cepacia* complex: past, present and future. Syst. Appl. Microbiol. **34:**87–95.
86. **Vuilleumier, S., et al.** 2009. *Methylobacterium* genome sequences: a reference blueprint to investigate microbial metabolism of C1 compounds from natural and industrial sources. PLoS One **4:**e5584.
87. **Wang, I. N., D. L. Smith, and R. Young.** 2000. Holins: the protein clocks of bacteriophage infections. Annu. Rev. Microbiol. **54:**799–825.
88. **Weigel, C., and H. Seitz.** 2006. Bacteriophage replication modules. FEMS Microbiol. Rev. **30:**321–381.
89. **Xu, M., et al.** 2005. Disulfide isomerization after membrane release of its SAR domain activates P1 lysozyme. Science **307:**113–117.
90. **Xu, M., D. K. Struck, J. Deaton, I. N. Wang, and R. Young.** 2004. A signal-arrest-release sequence mediates export and control of the phage P1 endolysin. Proc. Natl. Acad. Sci. U. S. A. **101:**6415–6420.
91. **Ye, C., et al.** 2009. Clinical, experimental, and genomic differences between intermediately pathogenic, highly pathogenic, and epidemic *Streptococcus suis.* J. Infect. Dis. **199:**97–107.
92. **Zeng, L., et al.** 2010. Decision making at a subcellular level determines the outcome of bacteriophage infection. Cell **141:**682–691.