Virology 388 (2009) 21–30

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

T4 phages against *Escherichia coli* diarrhea: Potential and problems

Emmanuel Denou, Anne Bruttin, Caroline Barretto, Catherine Ngom-Bru, Harald Brüssow *, Sophie Zuber

Nestlé Research Center, Nestec Ltd P.O. Box 44, CH-1000 Lausanne 26, Switzerland

ARTICLE INFO

Article history:
Received 23 June 2008
Returned to author for revision
15 December 2008
Accepted 1 March 2009
Available online 1 April 2009

Keywords:
T4 bacteriophage
Escherichia coli
Comparative genomics
Phage therapy
Safety evaluation

ABSTRACT

A combination of *in vitro* and *in vivo* experiments with comparative phage genomics was used for the rational design of a phage cocktail against *E. coli* diarrhea. Orally applied T4 coliphages representing three different subgroups (T4-, RB49- and JS98-like phages) had no negative impact on the murine gut microbiota. T4 phages were found with high titers in the cecum and colon and lower titers in the small intestine, but were not detected in the blood, liver or spleen. No adverse effects were observed after one-month exposure to phage nor were serum anti-T4 antibodies detected. T4 phages belonging to the same subgroup showed closely related genomes that differed by 12 (phage JS10 vs. JS98 reference) to 17 (phage JSE vs. RB49 reference) insertion/deletions mostly representing single small ORFs. Bioinformatic analysis did not reveal undesired genes in the T4 genomes. Sequence variability was seen over the tail fibre genes, but the variability did not correlate with phage host range. The investigated T4 phages were not only species- but also strain-specific, necessitating the use of phage cocktails consisting of 10 and 16 T4 phage isolates to cover half to two thirds of *E. coli* strains representing the five main pathotypes isolated from diarrhea patients.

Introduction

The increase of antibiotic resistance in a number of medically important bacterial pathogens limits the successful application of antimicrobial chemotherapy in clinical practice. The threat of an antibiotic crisis has revived the interest in diverse biological approaches against infectious diseases, including the use of phages against bacterial infections (Merril et al., 2003). The idea of phage therapy was introduced at the beginning of the last century by the discoverer of phages Felix d'Hérelle (Kutter and Sulakvelidze 2005). The first clinical trials gave variable results due to the unknown nature of phages and inadequate quality of the phage preparations. Nevertheless, these early efforts led to commercial phage products by US American pharmaceutical companies in the 1930s, sometimes, however, with doubtful claims. Subsequently, phage therapy against diarrheal diseases and wound infections was empirically, but systematically developed in the Soviet Union (Sulakvelidze et al. 2001). Numerous phage preparations are now sold as registered medicine in Russian pharmacies. However, neither the clinical trials conducted with these phage products nor the phages contained in these preparations were described in detailed scientific publications. Carefully documented veterinary work in the Western hemisphere demonstrated the potential of phage therapy against *E. coli* diarrhea in cattle (Smith et al. 1987a, 1987b) and *E. coli* respiratory infections in chicken (Huff et al., 2002; 2005). However, these phages were not described in detail either. Therefore it is not clear yet what phage properties are necessary for successful phage therapy.

Phage therapy approaches have to address a number of issues before human clinical trials can be considered (Brüssow 2005). To develop its full potential as a therapeutical or prophylactic antimicrobial agent, phage must infect the target pathogen in its epidemiological diversity, but spare commensal bacteria. For an easy oral application, phage must survive unprotected gastrointestinal passage. In addition, phage should stay sufficiently long and with high enough titer at the replication site of the pathogen to achieve its *in vivo* lytic activity. Finally, phages should not cause adverse side effects and their genomes should not carry virulence factors or undesired genes like antibiotic resistance genes, which they could transfer to the infected host. In the following we investigate whether these conditions are fulfilled for the molecularly well characterized (Karam 1994) T4-like coliphages.

Results and discussion

No impact of oral T4 phage on commensal microbiota in mice

Since phages are in general species-specific, phage application should not affect non-target bacteria thus avoiding the collateral damage induced by antibiotics on bystander bacteria. This aspect may not be a major concern when phages are used for the treatment of acute diarrhea since under these conditions the gut microbiota is already disturbed by the disease process. However, the specificity of phages becomes important during prophylactic oral application of phages.

* Corresponding author. Fax: +41 21 785 8544.
E-mail address: harald.bruessow@rdls.nestle.com (H. Brüssow).

0042-6822/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2009.03.009

We tested the impact of T4 phages on the gut microbiota when given orally to healthy conventional mice. On blood agar, the anaerobic cultivable microflora of control conventional C3H mice showed 10^{11} cfu (colony forming units) per gram of feces. On this medium mice yielded the following bacterial titers (in cfu/g of gut content) along the gastrointestinal tract: 10^7 (fore-stomach, mainly lactobacilli), 10^5 (duodenum), 10^7 (jejunum), 10^8 (ileum), 10^{10} (cecum), 10^9 (upper colon) and 10^{10} (lower colon). On Drigalski agar, a more specific medium for enterobacteria (Chibani-Chennoufi et al. 2004b), we detected 10^6 cfu/g gut content in the cecum and the large intestine, while only low numbers were counted along the small intestine ($\leq 10^2$ cfu/g in most samples) and intermediate values were found in the fore-stomach. Notably, mice treated with representatives of three subgroups of T4 phages (Comeau et al., 2007; Desplats et al., 2002; Nolan et al. 2006; Zuber et al., 2007) showed an identical intestinal distribution of bacterial counts on blood and Drigalski agar (data not shown) as untreated control mice thus excluding a gross effect of oral T4 phages on gut commensals in mice.

Simple plating methods give only a small picture of the microflora. One might ask whether plating tests are sensitive enough to reveal microbiota changes after interventions. Therefore we treated conventional mice with antibiotics (10 μg/ml erythromycin plus 25 μg/ml chloramphenicol) in their drinking water. These mice showed a marked reduction of the fecal microbial flora. For example, the count of lactobacilli dropped from 10^9 cfu/g feces in control mice to 10^4 cfu/g feces in antibiotics-treated mice (Denou et al., 2008). However, due to substantial antibiotic resistance of diarrhea-causing *E. coli* strains (Djie-Maletz et al., 2008), antibiotics are now infrequently used clinically for the treatment of *E. coli* diarrhea, which is mainly symptomatic.

This lack of an impact of oral T4 coliphage on the murine gut microbiota was not surprising since an API sugar fermentation gallery identified the major colony types recovered from Drigalski agar plates as *Klebsiella sp*. No characteristic *E. coli* colonies were detected when fecal material was plated on the more selective EMB medium. Furthermore, when we picked 500 colonies, none showed *in vitro* susceptibility towards these three T4 phages (data not shown).

In vitro host range of T4 phages on pathogenic E. coli strains

While the host species-specificity of phages is a clear asset for phage therapy, many phages are not only species- but also strain-specific. The narrow host range of phages becomes a potential liability when a reasonable coverage of the targeted pathogen has to be achieved with a small phage cocktail. Table 1 illustrates this point with T4 phages representing four of the five known subgroups of T4 coliphages whose genomes were all sequenced (Comeau et al., 2007; Desplats et al., 2002; Nolan et al. 2006; Zuber et al., 2007; this manuscript). These phages were tested against a panel of entero-pathogenic (EPEC) and enterotoxigenic (ETEC) *E. coli* strains of known O, H and K serotype collected at the Division of Enteric Pathogens of the Central Public Health Laboratory, Enteric division, Colindale, London/UK during the 1970s (kindly provided by B. Rowe) associated with infant diarrhea (Levine and Edelman, 1984; Qadri et al., 2005). Only a few T4 phages showed a relatively broad host range on the pathogenic *E. coli* test strains like phage RB49, which lysed 8 of the 25 test strains. The other investigated T4 phages lysed only four or less of the test strains. In this example a cocktail of six T4 phages lysed 13 out of the 25 (52%) investigated *E. coli* pathogens.

Clinicians from the International Center for Diarrheal Diseases Research at Dhaka/Bangladesh (ICDDR,B) provided us with 46 epidemiologically independent pathogenic *E. coli* isolates, which were recovered from patients hospitalized in 2008 with acute diarrhea at their institution (F. Qadri, S. Sarker and A. Cravioto, ICDDR,B). These isolates represented five different pathotypes of *E. coli*, namely 10 enteroaggregative strains (EAggEC), 3 enteroinva-

Table 1
Susceptibility of the indicated O, K, and H serotypes of pathogenic *E. coli* towards the indicated sequenced T4-like coliphage.

Serotype	O18: K77	O26: K60	O44: K74	O55: K59	O86: K61	O111: K58	O112: K66	O114: H49	O124: K72	O125: K70	O127: K63	O119: K69	O128: K67	O6: H16	O8:H9	O15: H11	O25: H42	O78: H12	O115: H51	O20: H11	O27: H7	O63: H–	O128: H18	O148: H28	O153: H12	K-12
T4	+																			+				+		+
RB69	+																		+	+						+
RB49	+								+	+				+		+				+		+		+		+
JSE	+										+			+												+
JS98	+	+						+			+															+
JS10	+	+				+										+										+

+: lysis in test tube on indicated strain; O18 to O128 (left side) are all EPEC strains; O6 to O153 (right side) are all ETEC strains; K-12 at the far right side is a non-pathogenic laboratory strain.

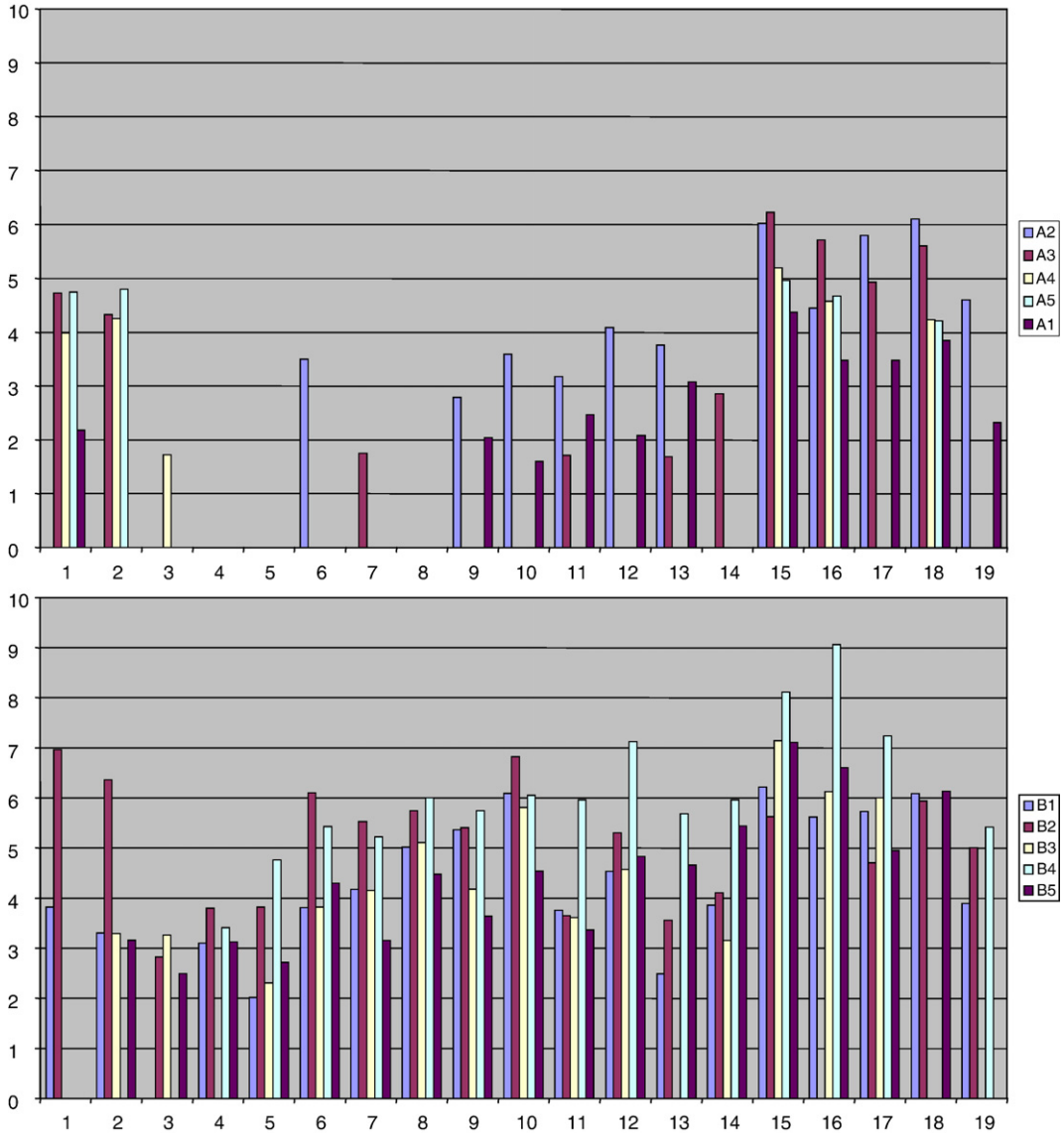

Fig. 1. T4 phage distribution along the different gut segments of mice fed with 10^9 pfu T4 phage cocktail in drinking water (top panel) and in mice fed with the same amount of phage, which received in addition 10^{10} cfu of *E. coli* K-12 by intra-gastric feeding (bottom panel). The ordinate indicates the $\log_{10}$ pfu/g for five individual mice (A1 to A5 and B1 to B5 represent individual mice); the numbers at the abscissa indicate the gut segments (1: extract and 2: scraping of stomach; 3: wash and 4: scraping of duodenum; 5, 6 and 7, 8 wash and scraping of proximal and distal jejunum, respectively; 9, 10 and 11, 12 and 13, 14 wash and scraping of proximal, medial and distal ileum, respectively; 15: extract of cecum; 16, 17 and 18, 19 extract and scraping of proximal and distal colon, respectively). Where no bar is indicated in the figure, no phage was detected.

sive strains (EIEC), 15 EPEC strains, 3 verotoxin-producing *E. coli* strains (VTEC) and 15 ETEC strains. Against these pathogenic *E. coli* strains, the cocktail showed only 18% coverage suggesting temporal changes in the *E. coli* epidemiology (Harris et al., 2008; Stoll et al., 1983). We screened our T4 phage collection, which comprises about 140 ecologically independent T4 phage isolates (Chibani-Chennoufi et al., 2004a; Zuber et al. 2007) against the pathogenic *E. coli* collection from ICDDR,B. With a cocktail consisting of 10 phages we achieved 52% coverage of this pathogen collection. Only two phage isolates lysed each 6 distinct pathogenic *E. coli* strains. One, three and four phage isolates lysed three, two and one pathogenic *E. coli* strains, respectively. A coverage of two thirds of the pathogenic *E. coli* strains from the ICDDR,B collection necessitated a cocktail consisting of sixteen T4 phage isolates demonstrating that pathogen coverage is an issue in phage therapy approaches against *E. coli* diarrhea. This same 16-phage cocktail showed a 53% coverage against a collection of 40 distinct pathogenic *E. coli* strains isolated from 321 children living in Dhaka/Bangladesh followed prospectively in the community for diarrhea episodes not leading to hospitalization over their first two years of life (Qadri et al., 2007).

Gastrointestinal passage of oral phage

Next we asked for the gut survival of phage added to the drinking water. When mice received a phage preparation consisting of three different subgroups of T4 phages at a dose of 10^9 plaque forming units (pfu)/ml drinking water, phage titers up to 10^6 pfu/g gut content were detected in the cecum and the large intestine. The titers were thus 1000-fold lower than in the drinking water. Even when accounting for dilution effects by secretions into the gastrointestinal tract, a sizable loss of phage must have occurred. A possible barrier could be the mouse stomach where the acidity can drop to a pH of 3. However, a much lower dose of 10^5 pfu/ml of phage in drinking water led to phage detection in all feces samples of adult human volunteers (Bruttin and Brüssow 2005), who display stronger gastric acidity drops than mice. It is thus not clear whether phage

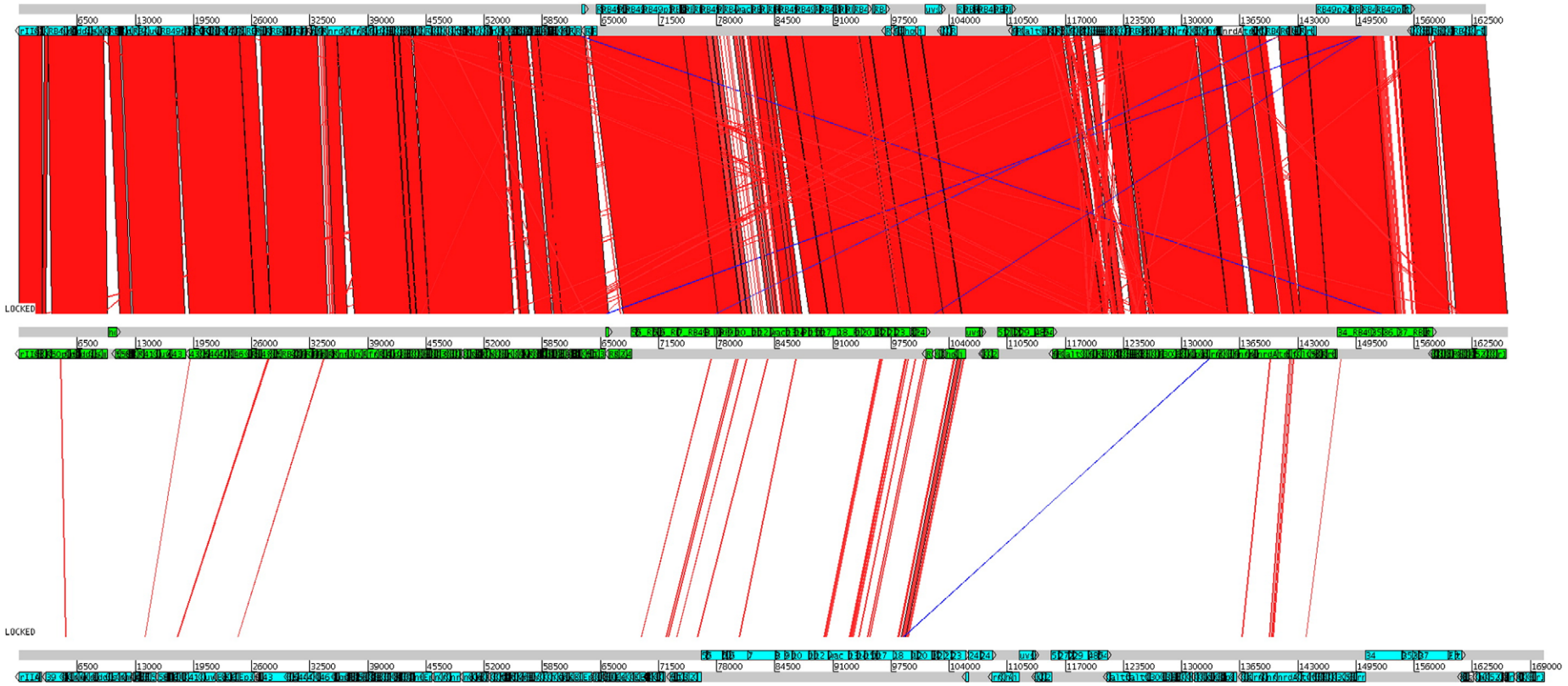

Fig. 2. Genome alignment of the RB49-like coliphage JSE isolated from sewage in Switzerland (center) with coliphage RB49 (top), the reference phage for a subgroup of T4 coliphages, which was isolated from sewage in the USA, and coliphage JS98 (bottom), the reference phage from another subgroup of T4 coliphage, which was isolated from a pediatric diarrhea patient in Dhaka/Bangladesh. The DNA sequence comparison is visualized with the Artemis Comparison Tool (ACT) developed by the Sanger Center (http://www.sanger.ac.uk/Software/ACT/) (Rutherford et al., 2000) using BlastN *E*-values $<10^{-3}$. The ORFs are indicated by arrows starting with the *rII* gene at the left side of the map.

must be given together with a bicarbonate buffer to achieve high *in situ* concentrations.

In the phage-fed mice, phage titers in excess of 10^4 pfu/g content were found in the stomach. In the small intestine titers were equal or less than 10^4 pfu/g gut content and more than half of the samples did not show phages (detection limit was 50 pfu/g material) (Fig. 1A). Due to the rapid transit time only low amounts of luminal material was recovered from the small intestine. In patients with acute diarrhea, oral phages will thus only have a very short *in situ* contact time with the *E. coli* pathogen.

In vivo replication of oral phage

Subsequently we explored the replication of oral phage on target *E. coli* cells in the gut. As our conventional mice lacked *E. coli* target cells in the gut, we introduced 10^{10} cfu of the phage-susceptible indicator *E. coli* strain K-12 by intra-gastric force-feeding. Four hours later the mice were dissected. Mice which received K-12, but no oral phage showed *E. coli* titers between 10^5 and 10^8 cfu/g in the cecum and colon and 10^5 cfu/g or less in the small intestine. Test mice (K-12 fed plus T4 phage in the drinking water) showed higher intestinal phage titers than mice, which received identical phage concentrations, but no K-12 feeding (compare Fig. 1B with Fig. 1A). The phage titer increase was 1000-fold in the jejunum and proximal ileum, but only 10 to 100-fold in the distal ileum and the large intestine. Oral phage had no impact on *E. coli* titers in the stomach, cecum and colon, while no K-12 was observed in the small intestine of phage-treated mice (data not shown). There is currently no mouse model to test the *in situ* replication of phage under conditions resembling an *E. coli* diarrhea. A mouse model exists for *Citrobacter rodentium* (Borenshtein et al., 2008) but not for *E. coli* in mice (but see Tanji et al. 2005). For the enterohemorrhagic *E. coli* strain O157:H7 a model was developed in rabbits (Ritchie et al., 2003; Ritchie and Waldor, 2005). Our phage cocktails displayed no lytic activity against these pathogens precluding the use of these models.

Safety of oral T4 phage in mice

Representatives of three different subgroups of T4 coliphages were given to conventional 6-week-old female C3H mice both individually and as a cocktail. The phage was orally applied at the high dose of 10^9 pfu/ml drinking water over one month. No adverse events were observed, and mice showed normal weight gain and normal behaviour (data not shown). At the end of the experiment, no phages were detected in blood, liver and spleen samples when using plaque assays (detection limit 50 pfu/g). Serum antibodies against T4 phage were searched in phage-exposed mice using two tests (murine Ig-specific ELISA on T4 phage-coated plates and neutralization test using the plaque reduction test with T4). No significant T4-specific antibody increase was seen for paired serum samples taken before and four weeks after oral phage exposure (data not shown). Furthermore, by macroscopic observation the gastrointestinal tract of the phage-treated mice did not reveal evidence of inflammation. The innocuous nature of phage T4 was previously also demonstrated in human volunteers receiving this phage in their drinking water (Bruttin and Brüssow, 2005). An integral part of a safety analysis of phages should also be their in silico genome analysis as detailed in the following paragraphs.

Genome sequence of the RB49-like phage JSE

T4-like phages were the target of a genome sequencing project (http://phage.bioc.tulane.edu). At least one representative of each of the five subgroups was sequenced and comparative genomics analyses were published (Comeau et al., 2007; Chibani-Chennoufi et al., 2004c; Desplats et al., 2002; Nolan et al. 2006; Zuber et al., 2007). A detailed "genomic safety" analysis of one representative T4 phage did not revealed undesired genes in its genome (Zuber et al., 2007). Now we asked what genetic diversity is encountered in T4-like coliphages belonging to the same subgroup. To answer this

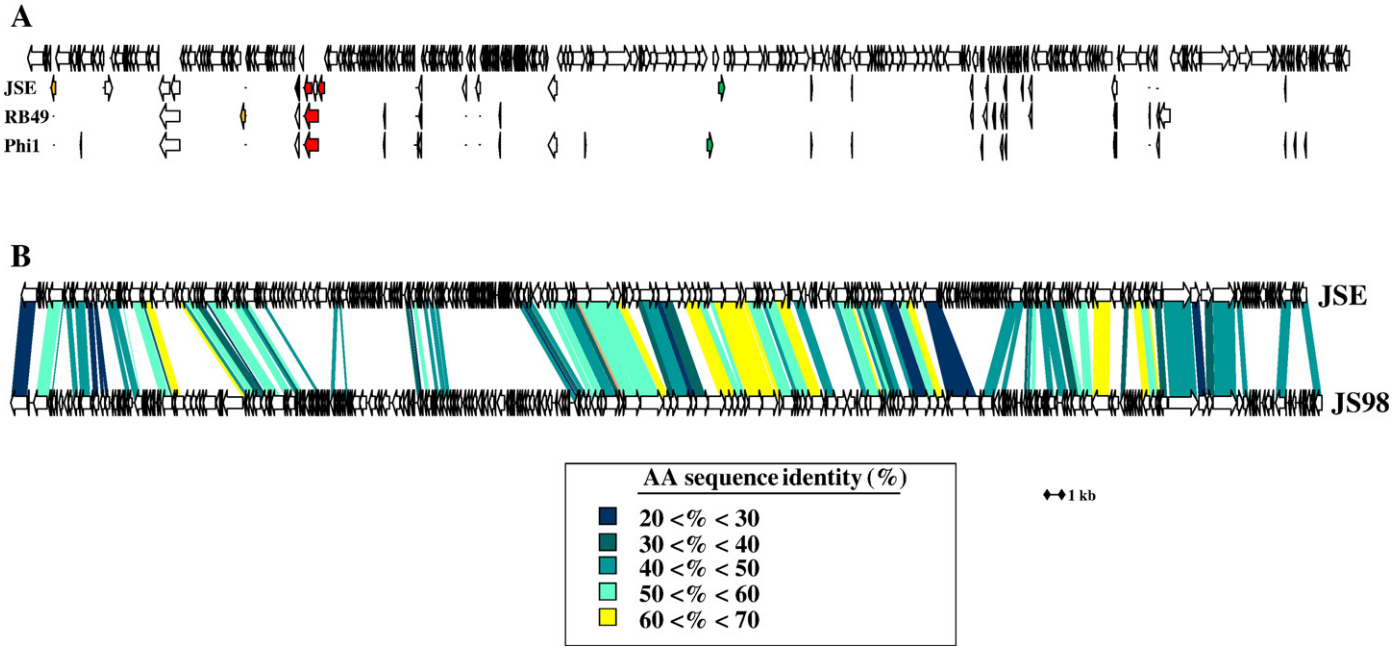

Fig. 3. Genome map alignment of RB49-like coliphages. (A) Alignment of the genome maps of the RB49-like coliphages JSE, RB49 and phi1 (http://phage.bioc.tulane.edu). The phi1 sequence is from Nolan, J.M., Department of Biochemistry, Tulane University Health Sciences Center, New Orleans, LA 70112, USA listed in T4-like genome database in NAR Molecular Biology Database Collection entry number 750 (http://www3.oup.co.uk/nar/database/summary/750). The top row shows those parts of the genome map that are shared between all three phages. The lower three rows highlight the regions where differences in the genome map were observed between the three phages depicted individually. Where related ORFs are found at different positions along the map, they are highlighted with the same color. (B) Alignment of genome maps of coliphages JSE and JS98, which represent two different subgroups of T4 coliphages. The ORFs are indicated by arrows starting with the *rII* gene at the left side. ORFs sharing aa identity are linked by color shading; the degree of sequence sharing is explained by the color scale at the bottom of the figure.

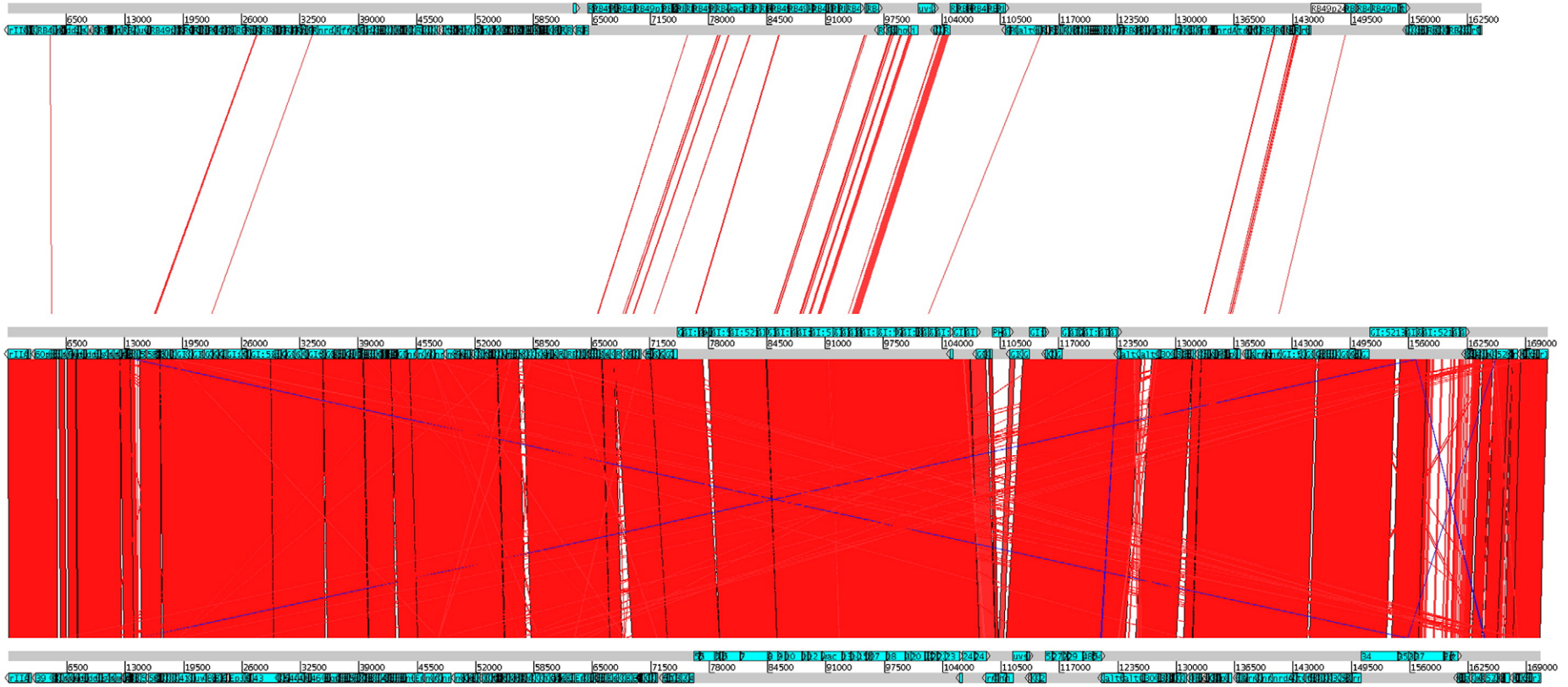

Fig. 4. Genome alignment of the JS98-like coliphage JS10 isolated from a pediatric diarrhea patient in Dhaka/Bangladesh (center) with coliphage RB49 (top) and coliphage JS98 (bottom), which was also isolated from a pediatric diarrhea patient in Dhaka/Bangladesh. The DNA sequence comparison is visualized with the Artemis Comparison Tool (ACT) using BlastN *E*-values $<10^{-3}$. The ORFs are indicated by arrows starting with the *rII* gene at the left side of the map.

question we sequenced the genome from the T4-like coliphages JSE and JS10.

Phage JSE is a member of the RB49-subgroup of T4 coliphages. JSE was isolated in 2000 from a sewage station in Switzerland, while RB49 was isolated more than 40 years ago from a sewage plant in the USA (Russell 1974). JSE yielded a 166,418 bp-long genome sequence in a single contig. In an ACT nucleotide comparison the JSE genome sequence could be aligned with the DNA sequence of phage RB49 (Comeau et al. 2007; Desplats et al. 2002) essentially over the entire genome length (Fig. 2). In sharp contrast, phages JSE and JS98, which belong to two different subgroups of T4-like coliphages, showed only limited DNA sequence sharing in the ACT comparison (Fig. 2). However, at the protein level, substantial sequence identity was still observed between JSE and JS98, particularly over the structural gene cluster (Fig. 3B) which represents the core region of the T4 phage family (Comeau et al., 2007; Filée et al., 2006). The conserved genome region shared between phage JSE and JS98 was, however, smaller than that shared across other T4 subgroups (Zuber et al., 2007).

Phages JSE and RB49 differed by 17 insertions/deletions (Fig. 2). Most of them covered a single ORF (Fig. 3A) and represented small hypothetical ORFs and two encoded homing endonucleases. Overall the difference between independent RB49-like phages was small even when including a third phage phi1 from the Tulane database into this comparison (Fig. 3A). One larger variable region encoded the homologues of the T4 baseplate wedge protein gp11, the short tail fibre gp12 and the neck whisker gpwac. Another region with substantial sequence variability covered the tail fibre genes *g34* to *g37*. Two split genes (Petrov et al., 2006) were observed when JSE was compared with RB49. The first was JSE ORF27 and 28, encoding, respectively, the N-terminal and the C-terminal halves of gp43 T4 DNA polymerase, separated by 480 bp without database match. The second split gene in JSE was ORF57 and 59 encoding two parts of an anaerobic ribonucleotide reductase, separated by an endonuclease gene (ORF58).

Genome sequences of JS98-like phages JS10 and JSD1-4

Like JS98 phage JS10 was isolated from the feces of a diarrhea patient in Bangladesh while JSD1-4 is an environmental water isolate from Bangladesh. JS10 yielded a 171,451 bp-long genome sequence in a single contig. In an ACT nucleotide comparison the JS10 genome sequence could be aligned with the DNA sequence of phage JS98 over the entire genome length, while only small regions of DNA sequence identity were shared with phage RB49 (Fig. 4). The JS10/JS98 alignment showed 12 insertions/deletions (Fig. 5). Most of them represented single and small ORFs including the large outer capsid protein gene *hoc*, the *alt* gene-homologue encoding an adenosylribosyltransferase, and *motB*, whose gene product interacts with MotA (Miller et al. 2003). One large region of variability covered the adjacent tail fibre genes *g37* and *g38*, gene *t*, encoding the phage holin, and *asiA*, encoding a protein that facilitates interaction of the MotA activator with T4 phage middle promoters (Miller et al., 2003).

Table 2
Percentage of amino acid identity between the tail fibre proteins gp34 (above the diagonal) and gp37 (below the diagonal) of the indicated T4-like coliphages.

Gp37/gp34	T4	RB32	RB69	JS98	JS10	RB49	JSE
T4	–	**91**	70	51	51	37	37
RB32	(76)	–	70	51	51	36	43
RB69	34	(71)	–	51	52	36	36
JS98	34	(67)	***75***	–	**96**	42	42
JS10	32	(35)	***64***	***61***	–	38	38
RB49	35	49	36	32	21	–	**98**
JSE	((44))	(49)	(35)	(38)	(35)	***71***	–

Numbers in parenthesis: identity is only over part of the proteins; double parenthesis: only short region of alignment. Bold numbers are the highest identity percentages within a subgroup of T4 coliphages, bold numbers in italics are the highest identities between subgroups of T4 coliphages.

Phage JSD1-4 yielded 148,778 bp of partial sequence information coming in 82 contigs. Due to its close similarity with the JS98 genome sequence, the numerous JSD1-4 contigs could be easily ordered along the JS98 genome (Supplementary Fig. 1).

The tail fibres of T4 contain the anti-receptor reacting with the cellular receptor, which is frequently the lipopolysaccharide (LPS) of *E. coli* (Tétart et al., 1996). The large proximal tail fibre subunit protein gp34 showed high protein sequence identity between phages from the same subgroup (91–98%) and much lower identity across the subgroups displaying a clear gradient (Table 2). Surprisingly, the highest level of aa identity of the large distal tail fibre subunit protein gp37, which is annotated as the phage adhesin, was observed between JS98 and RB69, which represent two different T4 phage subgroups and which show clearly distinct host specificities (Table 2). Apparently, the interaction of the tail fibre with the bacterial receptor is not the only host range determinant, although it definitively plays a primary role (Tétart et al., 1998).

"Genomic safety" of T4 phages

For all but ten of the 277 predicted JSE proteins the closest database match was with phage RB49 (Supplementary Table 1); five of these ten proteins were annotated as homing endonucleases, suggesting an introduction of selfish DNA elements into the JSE genome as was previously observed in T4 phage (Quirk et al., 1989). Three proteins showed links with T4-like coliphage phi1, two proteins lacked database matches.

No DNA sequence identity was detected between JSE and *E. coli*, thus excluding recent gene transfer between phage and its bacterial host. Nineteen JSE proteins shared protein sequence identity <60% with *E. coli* (Supplementary Table 2); 17 of these 19 proteins also had a match with T4 phage genes. Fifteen JS10 proteins shared <54% aa sequence identity with *E. coli* proteins; 13 of them showed a better alignment with T4-like phages (Supplementary Table 3). Finally, we screened the JSE and JS10 genome sequences against a database of harmful or undesired genes (DUG), as well as against a list of virulence genes identified in prophages from bacterial pathogens (Boyd and Brüssow, 2002; Brüssow et al., 2004; Canchaya et al. 2003) and against an allergen database (see Materials and methods). No suspicious gene matches were thus identified.

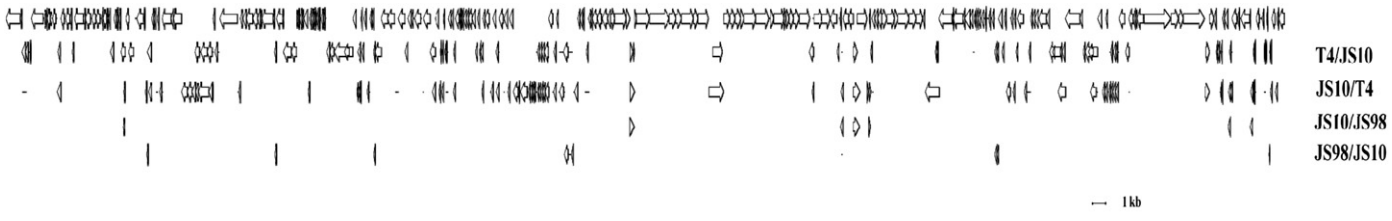

Fig. 5. Inter- and intra-subgroup alignments between T4-like coliphages. The top row shows those parts of the genome map that are shared between T4, JS10 and JS98 phages. The second and third rows give the T4-specific and JS10-specific ORFs, respectively, from the inter-subgroup comparison between phages T4 and JS10. The fourth and fifth rows give the JS10-specific and JS98-specific ORFs, respectively, from the intra-subgroup comparison between phages JS10 and JS98. The ORFs are indicated by arrows starting with the *rII* gene at the left side of the map.

In the construction of the DUG database more than 300 keywords were used. Most of them covered antibiotic resistance genes, while toxin and virulence genes were only represented by generic terms. Since virulence is not well defined, one might question the efficacy of such searches compared to direct searches of the individual phage genes against the large online public databases. We used prophages from the *E. coli* strain O157:H7, which contain biologically proven virulence genes (Tobe et al., 2006), as a test case with the DUG database. The DUG screening demonstrated Shiga toxin, cytotoxin, verocytotoxin, and a complement resistance factor in a short hit list. Blast searches against the NCBI database identified Shiga toxin, cytotoxin and an LEE effector. Both methods are of comparable sensitivity, but a DUG search is done quicker. However, this becomes only a clear advantage when numerous phages or larger bacterial genomes (for which DUG was developed) are screened for undesired factors.

Conclusion

T4 phages are professional virulent phages that propagate by serial lytic infection cycles and do not establish a lysogenic state. Since T4 phages destroy the bacterial DNA during the infection process, the transfer of undesired genes via lysogenic conversion or transduction is thus unlikely. Within subgroups of T4 coliphages the phage genomes are well conserved and no undesired genes were identified by bioinformatic analysis. In addition, oral T4 phages did not lead to adverse effects in animals or humans making T4 suitable candidates for phage therapy. More questions have still to be answered with respect to the efficacy of T4 use. Since no T4 phages with a broad host range on pathogenic *E. coli* were identified, phage approaches against *E. coli* diarrhea necessitate the use of phage cocktails. Phages are a potentially self-amplifying antimicrobial agent when they meet their target cells *in vivo*. Therefore it is not clear whether the low phage titers established in the small intestine, i.e. the site where the disease process of *E. coli* diarrhea occurs in humans, are sufficient to assure an *in situ* replication on the pathogen. The only modest amplification of T4 phages in the gut, which we observed, reflects more the limitation of our mouse model with respect to an *E. coli* infection in the gut than an intrinsic low replication potential of T4 phages *in vivo*.

Materials and methods

Sequencing

Phage DNA was amplified by an emulsion-based method, sequenced by synthesis using a pyro-sequencing protocol (Margulies et al., 2005) and phages JS10 and JSE were *de novo* assembled into a single contig by 454 LifeSciences Corp (Branford, Connecticut, USA). About 60,000 reads per phage were assembled into a single contig. Since the sequence did not contain repeats, the assembly was straightforward. Most reads showed lengths of between 100 and 140 bp (mean at 110 bp) and the coverage of the sequencing varied from 10- to 70-fold (mean 30-fold). In regions of <30-fold coverage, a 30-bp overlap of at least 10 readings was achieved. Previously we compared the sequence of a T4-like phage obtained with the pyro-sequencing protocol with that obtained for sheared phage DNA cloned into the plasmid vector pUC18, propagated in *E. coli* and sequenced by the Sanger method. We observed 99.97% sequence identity; discrepancies were mostly resolved in favor of the pyro-sequencing method (Zuber et al., 2007). The sequence data were deposited at GenBank under accession number EU863408 for JSE and EU86409 for JS10.

Bioinformatics analysis and annotation

Genome sequence comparisons were carried out using the MUMmer package (version 3.18) (http://www.tigr.org/software/mummer) (Kurtz et al., 2004) and EMBOSS (The European Molecular Biology Open Software Suite, version 4). We used the Artemis software (version 9) as sequence viewer and annotation tool (Rutherford et al., 2000). ORFs were predicted using the Glimmer software (version 3.02) (Delcher et al., 1999) and based on nucleotide and amino acid sequence alignment searches (BlastN, TBlastN, BlastX and BlastP), using the T4-like genome database available through the Tulane website (http://phage.bioc.tulane.edu) and the non-redundant database (nr) from NCBI. The basic prerequisites for an ORF were the presence of one of the three potential start codons ATG, TTG or GTG, and a length of at least 25 encoded amino acids. A search for tRNA genes was done with the tRNAscan-SE program (version 1.23) (Lowe and Eddy, 1997). Homology assignments between genes from other T4-like phages and predicted ORFs of phage JSE and JS10 were based on amino acid sequence alignment searches (BlastP) and were only accepted if the statistical significance of the sequence similarities (*E*-value) was $\leq$0.001, the bit score $\geq$50 and the percent identity between the aligned sequences (% id) $\geq$30%. Functional annotations were based on the homology assignments with the T4 genes and functional classifications were performed using the COG (Clusters of Orthologous Groups of proteins) database (Tatusov et al. 1997). We used the ACT (Artemis Comparison Tool, version 6) to compare closely related genomes. Comparisons between two genomes were based on nucleotide sequence alignments searches (BlastN) (*E*-value $\leq$0.001 and bit score $\geq$100). The minimal length to determine supplementary and variable DNA regions was 100 nucleotides.

Genetic safety screening

The Database of Undesirable Genes (DUG) was built in the following way. A list of terms related to antibiotics, virulence, and toxins, comprising a total of 314 items, was defined. Then, the EMBL public sequence database (only the prokaryote and phage subsets) is screened based on these terms. For all EMBL coding sequence (CDS) files, the terms are searched in the following features: “description”, “product”, and “note”. If one of the terms is found, the protein and its related nucleotide sequences are added to the DUG_protein and the DUG_nucleotide databases respectively. Next a set of experimentally proven sequences known to be involved in antibiotic resistance is added to the sequences collected from EMBL. All sequences in the DUG databases are in FASTA format. Their title line contains the description of the gene or protein, the specific term which was found and the organism name. Then, the two DUG databases are made non-redundant (http://blast.wustl.edu/pub/nrdb/). The database is updated at bimonthly intervals. BlastP and BlastN searches were performed against the DUG and against 15 *E. coli* genomic sequences. Only hits with an *E*-value $\leq$0.01 (BlastP and BlastN) and a bit score $\geq$50 (BlastP) were considered to constitute significant matches.

The predicted phage proteins were screened for similarities against known protein food allergens in the Food Allergy Research and Resource Program at http://www.allergenonline.com.

Host range

The lysis test in tubes was done against a representative collection of enteropathogenic and enterotoxigenic *E. coli* serotypes as described previously (Chibani-Chennoufi et al. 2004b).

Animal experiments

Adult female conventional C3H mice were separated into insulators. All mice received the specified phage in sterile Vittel mineral water. Phages were propagated on *E. coli* strain K-12 inoculated into Hershey medium. The lysate was cleared by centrifugation and passed through a 0.22 μm Millipore filter. The T4 phages were then pelleted by ultracentrifugation (35,000 ×g for 25 min) and resuspended in Vittel water.

After the oral feeding of phage with the drinking water and if indicated the intra-gastric feeding of the indicator K-12 *E. coli* strain, fecal samples were collected and the mice were at the indicated time points anaesthetized with isoflurane and killed by bleeding. A blood sample was collected into a heparin tube. Liver and spleen were dissected as well as the stomach, duodenum, jejunum (2 segments), ileum (3 segments), cecum, and colon (2 segments). The stomach, cecum and colon contents were extracted and the small intestine content was recovered by washing with a syringe. After extraction and washing the mucosa layer was collected from all gut segments (except the cecum) by scraping. The presence of phage was determined by the double-layer agar plaque assay on Hershey agar. The bacterial counts were determined on blood agar, Drigalski or EMB medium according to the indications of the supplier (Difco). The identity of selected colonies was further determined on an API sugar fermentation gallery. Food intake, mobility, change in behaviour and physical appearance of the mice was noted by the workers in the animal house. The gross anatomy of the gut and other vital internal organs was noted during the dissection of the mice.

ELISA was essentially done as described (Brüssow et al. 1988) except that purified T4 phage antigen was coated on the plates and a polyvalent anti-murine Ig conjugate was used. Neutralizing anti-T4 serum antibodies were determined with a plaque reduction test. A T4 phage preparation titrated to yield 100 pfu on a plaque assay was incubated before plating with a series of de-complemented serum dilutions at 37 °C for 30 min. The serum dilution yielding a greater than 50% plaque reduction was considered as neutralizing.

Acknowledgments

We thank Henry Krisch (CNRS Toulouse) for critically reading the comparative genomics part of the manuscript. We also thank B2. Rowe (PHL) and F. Qadri, S. Sarker and A. Cravioto (ICDDR,B) for providing bacterial pathogens associated with diarrhea.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.virol.2009.03.009.

References

Borenshtein, D., McBee, M.E., Schauer, D.B., 2008. Utility of the *Citrobacter rodentium* infection model in laboratory mice. Curr. Opin. Gastroenterol. 24, 32–37.

Boyd, E.F., Brüssow, H., 2002. Common themes among bacteriophage-encoded virulence factors and diversity among the bacteriophages involved. Trends Microbiol. 10, 521–529.

Bruttin, A., Brüssow, H., 2005. Human volunteers receiving *Escherichia coli* phage T4 orally: a safety test of phage therapy. Antimicrob. Agents Chemother. 49, 2874–2878.

Brüssow, H., 2005. Phage therapy: the *Escherichia coli* experience. Microbiology 151, 2133–2140.

Brüssow, H., Werchau, H., Liedtke, W., Lerner, L., Mietens, C., Sidoti, J., Sotek, J., 1988. Prevalence of antibodies to rotavirus in different age groups of infants in Bochum, Germany. J. Infect. Dis. 157, 1014–1022.

Brüssow, H., Canchaya, C., Hardt, W.-D., 2004. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. Microbiol. Mol. Biol. Rev. 68, 560–602.

Canchaya, C., Proux, C., Fournous, G., Bruttin, A., Brüssow, H., 2003. Prophage genomics. Microbiol. Mol. Biol. Rev. 67, 238–276.

Chibani-Chennoufi, S., Sidoti, J., Bruttin, A., Dillmann, M.L., Kutter, E., Qadri, F., Sarker, S.A., Brüssow, H., 2004a. Isolation of *Escherichia coli* bacteriophages from the stool of pediatric diarrhea patients in Bangladesh. J. Bacteriol. 186, 8287–8294.

Chibani-Chennoufi, S., Sidoti, J., Bruttin, A., Kutter, E., Sarker, S., Brüssow, H., 2004b. In vitro and in vivo bacteriolytic activities of *Escherichia coli* phages: implications for phage therapy. Antimicrob. Agents Chemother. 48, 2558–2569.

Chibani-Chennoufi, S., Canchaya, C., Bruttin, A., Brüssow, H., 2004c. Comparative genomics of the T4-like *Escherichia coli* phage JS98: Implications for the evolution of T4 phages. J. Bacteriol. 186, 8276–8286.

Comeau, A.M., Bertrand, C., Letarov, A., Tétart, F., Krisch, H.M., 2007. Modular architecture of the T4 phage superfamily: a conserved core genome and a plastic periphery. Virology 362, 384–396.

Delcher, A.L., Harmon, D., Kasif, S., White, O., Salzberg, S.L., 1999. Improved microbial gene identification with GLIMMER. Nucleic Acids Res. 27, 4636–4641.

Desplats, C., Dez, C., Tétart, F., Eleaume, H., Krisch, H.M., 2002. Snapshot of the genome of the pseudo-T-even bacteriophage RB49. J. Bacteriol. 184, 2789–2804.

Denou, E., Pridmore, R.D., Berger, B., Panoff, J.-M., Arigoni, F., Brüssow, H., 2008. Identification of genes associated with the long-gut-persistence phenotype of the probiotic *Lactobacillus johnsonii* strain NCC533 using a combination of genomics and transcriptomics analysis. J. Bacteriol. 190, 3161–3168.

Djie-Maletz, A., Reither, K., Danour, S., Anyidoho, L., Saad, E., Danikuu, F., Ziniel, P., Weitzel, T., Wagner, J., Bienzle, U., Stark, K., Seidu-Kokor, A., Mockenhaupt, F.P., Ignatius, R., 2008. High rate of resistance to locally used antibiotics among enteric bacteria from children in Northern Ghana. J. Antimicrob. Chemother. 61, 1315–1318.

Filée, J., Bapteste, E., Susko, E., Krisch, H.M., 2006. A selective barrier to horizontal gene transfer in the T4-type bacteriophages that has preserved a core genome with the viral replication and structural genes. Mol. Biol. Evol. 23, 1688–1696.

Harris, A.M., Chowdhury, F., Begum, Y.A., Khan, A.I., Faruque, A.S.G., Svennerholm, A.-M., Harris, J.B., Ryan, E.T., Cravioto, A., Calderwood, S.B., Qadri, F., 2008. Shifting prevalence of major diarrheal pathogens in patients seeking hospital care during floods in 1998, 2004, and 2007 in Dhaka, Bangladesh. Am. J. Trop. Med. Hyg. 79, 708–714.

Huff, W.E., Huff, G.R., Rath, N.C., Balog, J.M., Donoghue, A.M., 2002. Prevention of *Escherichia coli* infection in broiler chickens with a bacteriophage aerosol spray. Poult. Sci. 81, 1486–1491.

Huff, W.E., Huff, G.R., Rath, N.C., Balog, J.M., Donoghue, A.M., 2005. Alternatives to antibiotics: utilization of bacteriophages to treat colibacillosis and prevent foodborne pathogens. Poult. Sci. 84, 655–659.

Karam, J.D. (Ed.), 1994. Molecular Biology of Bacteriophage T4. American Society for Microbiology Press, Washington, DC.

Kurtz, S., Phillippy, A., Delcher, A.L., Smoot, M., Shumway, M., Antonescu, C., Salzberg, S.L., 2004. Versatile and open software for comparing large genomes. Genome Biol. 5, R12.

Kutter, E., Sulakvelidze, A. (Eds.), 2005. Bacteriophages: Biology and Applications. CRC Press, Boca Raton.

Levine, M.M., Edelman, R., 1984. Enteropathogenic *Escherichia coli* of classical serotypes associated with infant diarrhea: epidemiology and pathogenesis. Epidemiol. Rev. 6, 31–51.

Lowe, T.M., Eddy, S.R., 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25, 955–964.

Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., et al., 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature 437, 376–380.

Merril, C.R., Scholl, D., Adhya, S.L., 2003. The prospect for bacteriophage therapy in Western medicine. Nat. Rev. Drug Discov. 2, 489–497.

Miller, E.S., Kutter, E., Mosig, G., Arisaka, F., Kunisawa, T., Ruger, W., 2003. Bacteriophage T4 genome. Microbiol. Mol. Biol. Rev. 67, 86–156.

Nolan, J.M., Petrov, V., Bertrand, C., Krisch, H.M., Karam, J.D., 2006. Genetic diversity among five T4-like bacteriophages. Virol. J. 3, e30.

Petrov, V.M., Nolan, J.M., Bertrand, C., Levy, D., Desplats, C., Krisch, H.M., Karam, J.D., 2006. Plasticity of the gene functions for DNA replication in the T4-like phages. J. Mol. Biol. 361, 46–68.

Qadri, F., Svennerholm, A.-M., Faruque, A.S.G., Sack, R.B., 2005. Enterotoxigenic *Escherichia coli* in developing countries: epidemiology, microbiology, clinical features, treatment, and prevention. Clin. Microbiol. Rev. 18, 465–483.

Qadri, F., Saha, A., Ahmed, T., Tarique, A.A., Begur, Y.A., Svennerholm, A.-M., 2007. Disease burden due to enterotoxigenic *Escherichia coli* in the first 2 years of life in an urban community in Bangladesh. Infect. Immun. 75, 3961–3968.

Quirk, S.M., Bell-Pedersen, D., Belford, M., 1989. Intron mobility in the T-even phages: high frequency inheritance of group I introns promoted by intron open reading frames. Cell 56, 455–465.

Ritchie, J.M., Waldor, M.K., 2005. The locus of enterocyte effacement-encoded effector proteins all promote enterohemorrhagic *Escherichia coli* pathogenicity in infant rabbits. Infect. Immun. 73, 1466–1474.

Ritchie, J.M., Thorpe, C.M., Rogers, A.B., Waldor, M.K., 2003. Critical roles for *stx2*, *eae*, and *tir* in enterohemorrhagic *Escherichia coli*-induced diarrhea and intestinal inflammation in infant rabbits. Infect. Immun. 71, 7129–7139.

Russell, R.L., 1974. Comparative genetics of the T-even bacteriophages. Genetics 78, 967–988.

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M.A., Barrell, B., 2000. Artemis: sequence visualization and annotation. Bioinformatics 16, 944–945.

Smith, H.W., Huggins, M.B., Shaw, K.M., 1987a. The control of experimental *Escherichia coli* diarrhoea in calves by means of bacteriophages. J. Gen. Microbiol. 133, 1111–1126.

Smith, H.W., Huggins, M.B., Shaw, K.M., 1987b. Factors influencing the survival and multiplication of bacteriophages in calves and in their environment. J. Gen. Microbiol. 133, 1127–1135.

Stoll, B.J., Rowe, B., Glass, R.I., Huq, I., 1983. Changes in serotype of enterotoxigenic *Escherichia coli* in Dhaka over time: usefulness of polyvalent antisera. J. Clin. Microbiol. 18, 935–937.

Sulakvelidze, A., Alavidze, Z., Morris, J.G., 2001. Bacteriophage therapy. Antimicrob. Agents Chemother. 45, 649–659.

Tanji, Y., Shimada, T., Fukudomi, H., Nakai, Y., Unno, H., 2005. Therapeutic use of phage cocktail for controlling *Escherichia coli* O157:H7 in gastrointestinal tract of mice. J. Biosci. Bioengin. 100, 280–287.

Tatusov, R.L., Koonin, E.V., Lipman, D.J., 1997. A genomic perspective on protein families. Science 278, 631–637.

Tétart, F., Repoila, F., Monod, C., Krisch, H.M., 1996. Bacteriophage T4 host range is expanded by duplications of a small domain of the tail fibre adhesion. J. Mol. Biol. 258, 726–731.

Tétart, F., Desplats, C., Krisch, H.M., 1998. Genome plasticity in the distal tail fiber locus of the T-even bacteriophage: recombination between conserved motifs swaps adhesin specificity. J. Mol. Biol. 282, 543–556.

Tobe, T., Beatson, S.A., Taniguchi, H., Abe, H., Bailey, C.M., Fivian, A., Younis, R., Matthews, S., Marches, O., Frankel, G., Hayashi, T., Pallen, M.J., 2006. An extensive repertoire of type III secretion effectors in *Escherichia coli* O157 and the role of lambdoid phages in their dissemination. Proc. Natl. Acad. Sci. U.S.A. 103, 14941–14946.

Zuber, S., Ngom-Bru, C., Barretto, C., Bruttin, A., Brüssow, H., Denou, E., 2007. Genome analysis of phage JS98 defines a fourth major subgroup of T4-like phages in *Escherichia coli*. J. Bacteriol. 189, 8206–8214.