Virology 434 (2012) 222–232

Contents lists available at SciVerse ScienceDirect

Virology

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

Oral T4-like phage cocktail application to healthy adult volunteers from Bangladesh

Shafiqul Alam Sarker [a], Shawna McCallin [b], Caroline Barretto [b], Bernard Berger [b], Anne-Cécile Pittet [b], Shamima Sultana [a], Lutz Krause [b,1], Sayeda Huq [a], Rodrigo Bibiloni [b,2], Anne Bruttin [b], Gloria Reuteler [b], Harald Brüssow [b,*]

[a] International Centre for Diarrhoeal Diseases Research, Bangladesh (icddr,b), 68 Shaheed Tajuddin Ahmed Sharani, Mohakhali, Dhaka 1212, Bangladesh
[b] Nestlé Research Centre, Nestec Ltd., Vers-chez-les-Blanc, CH-1000 Lausanne 26, Switzerland

ARTICLE INFO

Available online 24 October 2012

Keywords:
Bacteriophages
Escherichia coli
Human phage therapy
Bioavailability
Safety
Fecal microbiota

ABSTRACT

The genomic diversity of 99 T4-like coliphages was investigated by sequencing an equimolar mixture with Illumina technology and screening them against different databases for horizontal gene transfer and undesired genes. A 9-phage cocktail was given to 15 healthy adults from Bangladesh at a dose of 3×10^9 and 3×10^7 plaque-forming units and placebo respectively. Phages were detected in 64% of the stool samples when subjects were treated with higher titer phage, compared to 30% and 28% with lower-titer phage and placebo, respectively. No *Escherichia coli* was present in initial stool samples, and no amplification of phage was observed. One percent of the administered oral phage was recovered from the feces. No adverse events were observed by self-report, clinical examination, or from laboratory tests for liver, kidney, and hematology function. No impact of oral phage was seen on the fecal microbiota composition with respect to bacterial 16S rRNA from stool.

Introduction

Eastern Europe has a long tradition of using bacteriophages against bacterial infectious diseases for both prophylactic and therapeutic application (Summers, 2001). Numerous phage cocktails are sold in Russian pharmacies as registered products. These phage cocktails target two main disease complexes: diarrheal diseases and pyogenic infections (Sulakvelidze et al., 2001). The rise of antibiotic-resistant bacteria has rekindled the interest of the Western medical community in the Eastern phage therapy tradition (Merril et al., 2003). However, the scarcity of scientific publications from Eastern Europe detailing the composition of phage cocktails, their biological safety in human applications, and their efficacy in controlled clinical trials makes it difficult to rationally evaluate phage therapy. Many Western scientists have therefore remained skeptical towards the potential of this therapy. Since alternatives to antibiotics are urgently needed for an increasing number of bacterial infections, a systematic and critical re-evaluation of phage therapy has become a public health priority.

The safety issue of bacteriophage therapy for topical application in humans has only recently been addressed in scientific reports. A mixture of bacteriophages directed against *Pseudomonas aeruginosa*, *Staphylococcus aureus*, and *E. coli* or a placebo was applied on venous leg ulcers in 39 patients, and no adverse effects were associated with the treatment (Rhoads et al., 2009). Likewise, a group of 24 patients with chronic otitis externa (''swimmer's ear'') were topically treated with a *P. aeruginosa* phage preparation (Wright et al., 2009) with no adverse effects. Furthermore, a mixture of *P. aeruginosa* and *S. aureus* phages was developed for topical application on the skin of burn patients (Merabishvili et al., 2009), but the outcome of the ongoing safety trial has not yet been reported.

In Eastern Europe, oral phage delivery is currently used against intestinal infections (Sulakvelidze et al., 2001; Sulakvelidze and Kutter, 2005). However, before the efficacy of phage preparations can be studied on large human cohorts, it is essential to first assess their safety in humans to ensure that they do not cause adverse effects when applied orally. Previous animal toxicology studies of laboratory mice receiving oral T4 phages have

* Corresponding author. Fax: +41 21 785 8549.
E-mail addresses: sasarker@icddrb.org (S.A. Sarker), bernard.berger@rdls.nestle.com (B. Berger), shamima@icddrb.org (S. Sultana), ltz.krause@gmail.com (L. Krause), sayeeda@mail.icddrb.org (S. Huq), Rodrigo.Bibiloni@agresearch.co.nz (R. Bibiloni), anne.bruttin@rdls.nestle.com (A. Bruttin), gloria.reuteler@rdls.nestle.com (G. Reuteler), harald.bruessow@rdls.nestle.com (H. Brüssow).
[1] Current address: Queensland Institute of Medical Research, Division of Genetics and Population Health, Herston, Brisbane, Queensland 4029, Australia.
[2] Current address: AgResearch Ltd., Ruakura Research Centre, East Street, Private Bag 3123, Hamilton 3240, New Zealand.

0042-6822/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.virol.2012.09.002

documented the presence of phage throughout the intestinal tract, evidence of *in vivo* phage replication, and the noted absence of adverse events (Chibani-Chennoufi et al., 2004a,b; Weiss et al., 2009). Finally, the reference T4 phage T4D (T4D refers to the individual classical T4 phage isolate NC_000866 herein to distinguish from T4 as a generic name for the whole phage group) was tested in healthy adult volunteers from Switzerland without adverse events (Bruttin and Brüssow, 2005). Since it is reported that temperate coliphages frequently carry virulence genes (Canchaya et al., 2003), we chose to restrict our approach to the best-characterized professional virulent (“lytic”) coliphage group, namely T4-like Myoviridae, for which genomic analyses have not revealed the presence of so-called virulence genes (Denou et al., 2009; Brüssow et al., 2004; Brüssow, 2010).

Before a T4-like phage cocktail can be evaluated for its therapeutic efficacy in children hospitalized at ICDDR,B with *E. coli* diarrhea, several safety and efficacy issues posed by clinicians had to be addressed. We present here further genetic and clinical safety data on oral delivery of phages against *E. coli* serotypes causing diarrhea. Phage cocktails are needed in order to achieve a reasonable coverage of pathogenic *E. coli* strains for a given clinical condition, and therefore the composition of the cocktail might differ for different geographical regions or need to be adjusted over time. Our first objective was therefore to extend the genetic safety analysis to our entire T4-like phage collection of 99 T4-like phage isolates instead of sequencing the individual phage genomes of those phages used in the current cocktail.

To address clinical safety and possible risks associated with dosage, safety studies were repeated in the target population from Bangladesh with a 100-fold higher dose of phage cocktail consisting of different T4-like subgroups than the dosage used in a previous safety study with Swiss healthy adult volunteers (Bruttin and Brüssow, 2005). This particular dosage and population were used in this study before clinical trials could be conducted with a 10-fold higher dose (than the Swiss study) in older children, with the standard dose in young healthy children, and to eventual administration of this dose to hospitalized children.

A major drawback of antibiotics is their collateral damage to bystander commensal microbiota, while phages are considered to be target species-specific leaving the commensals unaffected. To document this difference, we investigated the impact of oral coliphage on the fecal microbiota in phage-treated adult volunteers.

Results

The T4 phage collection

The 99 T4-like phages that comprise our collection were obtained from stool samples of pediatric diarrhea patients and healthy children, sewage samples, and environmental water samples from Bangladesh and Switzerland (Chibani-Chennoufi et al., 2004a,b). The phages were isolated and amplified on an *E. coli* strain K-12 derivative devoid of prophage lambda, and T4-like classification of phage was determined by a diagnostic PCR (Tétart et al., 2001) and/or electron microscopy (EM). The diagnostic *g23* PCR products were sequenced, and a tree was constructed from the aligned sequences (Fig. 1). The tree analysis allowed the distinction of three major groups in our phage collection: 45 phage isolates clustered around the T4D reference strain; 14 isolates were associated with RB69, the prototype of another subgroup of T4 phages; and 36 phage isolates belonged to still another subgroup of T4 phages represented by phage JS98. Only one phage aligned with the RB49-like phage subgroup, while

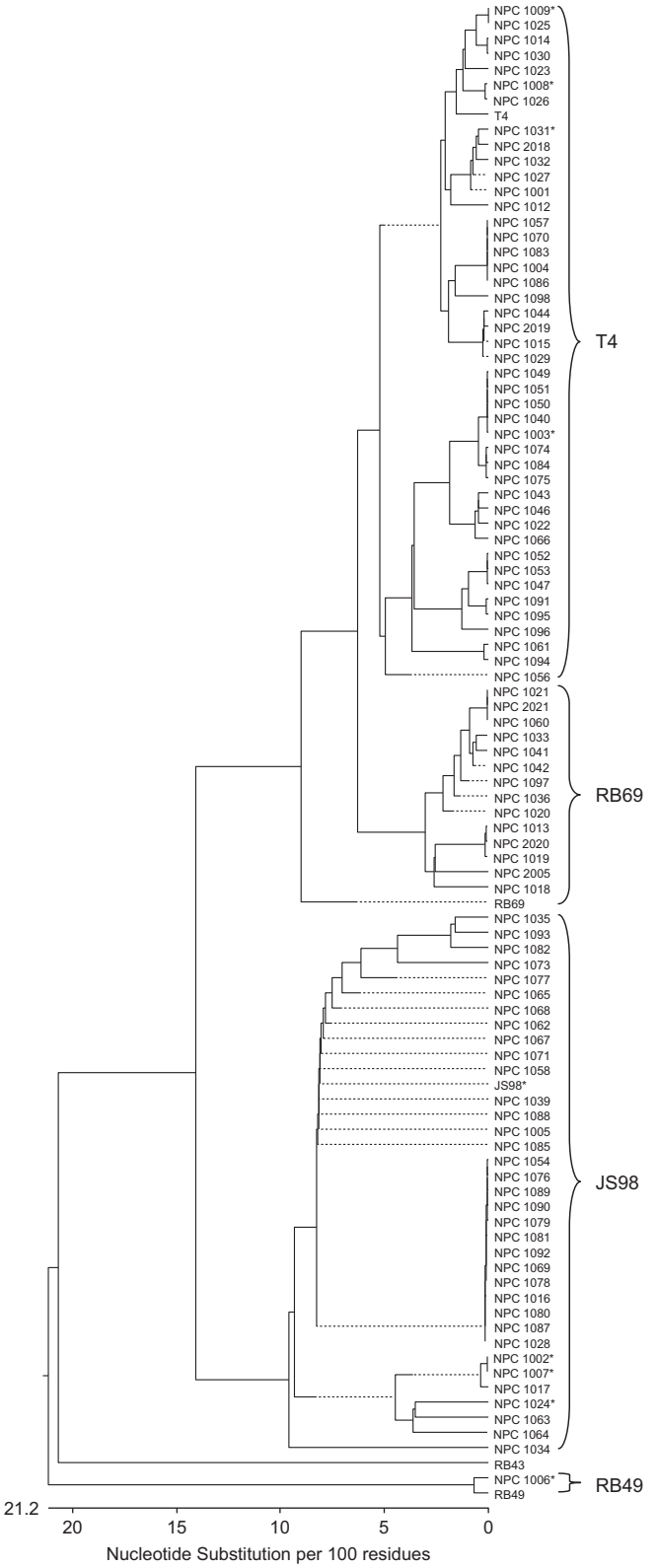

Fig. 1. Phylogenetic tree of *g23* sequences of 98 phages in the NPC. All phages in the NPC are represented here, with the exception of NPC 2022 due to poor *g23* sequence quality. The *g23* of RB43 is included in the tree, although it was not sequenced with the collection, and JS98 corresponds to NPC 1000. * indicates phages that were individual components of the phage cocktail administered to volunteers within this study. Dotted lines in the phenogram refer to negative branch lengths.

no phages related to the reference phage RB43 were part of our collection. Based on the *g23* sequence, the T4D-like phage branch showed a degree of sequence diversification similar to that of

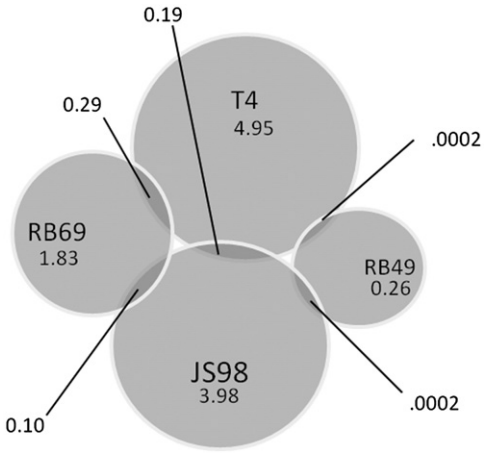

Fig. 2. Overlap of reads mapped to reference genomes of phage classes, in millions. No reads were common among all four reference genomes or between the RB69 and RB49 mappings.

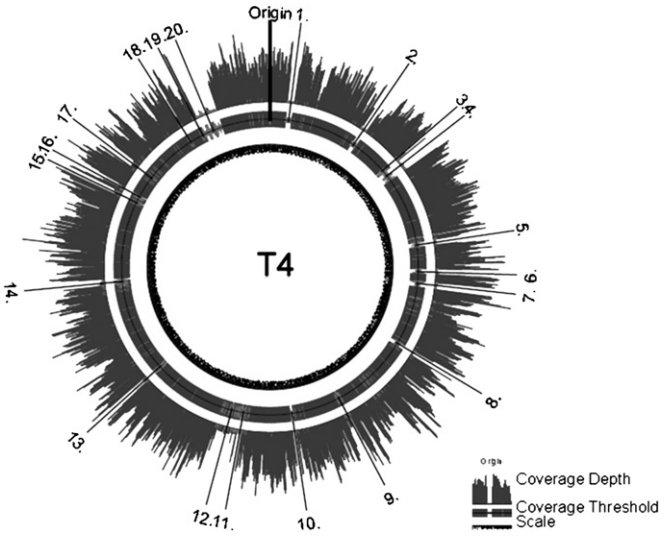

Fig. 3. T4 sequence mapping showing areas of low coverage. Low coverage areas are numerated and the corresponding genes found in that region are shown in bold, followed by a short description. **1: gp60** topoisomerase II; **gp60.1** hypothetical protein; **mobA** DNA endonuclease; **gp39** DNA topoisomerase II. **2: segF** homing endonuclease; **gp56** dCTP pyrophosphatase. **3: usvX** RecA-like recombination protein. **4: segA** homing endonuclease; **b-gt** beta glucosyl transferase; **5: mobB** DNA endonuclease. **6: nrdG** anaerobic NTP reductase; **mobC** homing endonuclease. **7: I-TevII** homing endonuclease. **8: mobD** homing endonuclease. **9: segB** homing endonuclease; **tRNA. 10: segC** homing endonuclease; **gp5.3** conserved hypothetical protein. **11: gp10** baseplate wedge subunit and tail pin. **12: gp11** baseplate wedge subunit and tail pin ; **gp12** short tail fibers; **wac** fibritin (whisker fibers). **13: segD** DNA endonuclease. **14: alt** ADP-ribosyltransferase. **15: I-TevIII** homing endonuclease (defective). **16: nrdB** aerobic NDP reductase; **nrdB.1** hypothetical protein; **mobE** homing endonuclease. **17: I-TevI** homing endonuclease. **18: gp34** proximal long tail fiber. **19: gp36** hinge connector of long tail fiber. **20: gp37** distal long tail fiber; **gp38** adhesin.

RB69-like phages while JS98-like phages were more closely related to each other.

Sequencing of the T4 phage collection

Health authorities such as the FDA request the sequencing of phage genomes in order to assess the genetic risk associated with the application of phages. Foreseeing that phage cocktail composition may need to be adjusted in the future, we sequenced our entire T4 phage collection instead of only sequencing the individual phages components of the cocktail tested within this study. The 99 phages of our collection were propagated individually, their DNA was extracted, and equal amounts of phage DNA were mixed for sequencing with the Illumina technique. Thirty-one million 50 bp reads with a high sequencing quality above Q30 were obtained.

Sequence mapping was initially done with the ensemble of reads in order to gain insight into the relatedness of NPC phages to several reference genomes representing the main subgroups of T4-like phages, namely T4, JS98, RB69, and RB49 (Supplementary Table 1). The ratio of the reads aligned with the T4D compared to the RB49 genome was 19:1 and thus very similar to the ratio of T4D-like over RB49-like phages of 22:1 as determined by *g23* sequencing. The same was true for the ratio of T4D over RB69 (2.7:1 and 3:1) and the ratio of T4D over JS98 (1.2:1 and 1.2:1) obtained by aligned Illumina reads and *g23* sequence-based phage typing, respectively. This close association suggests first, that the DNA sequences were read proportional to the representation of the phages in the collection and second, that *g23* sequences are a good predictor for attribution of the entire genome sequences to distinct subgroups. Shared reads between the subgroups were minimal: only a small proportion of sequences could be attributed to two different subgroups (Fig. 2).

Sequence alignments with reference genomes

When projected on the T4D reference genome, an overall coverage of 1460 reads per 50 bp region was obtained with our Illumina sequence set. However, the coverage was not homogeneous: 20 regions were covered poorly (Fig. 3). Three long T4D genome segments with low coverage were identified, including a 4-kb region encoding a baseplate wedge protein, tail pin, short tail fibers, and whisker fibers. The high coverage over major parts of the T4D reference genome demonstrates that many phages in our phage collection belonged to this group and shared close bp identity with T4D.

According to the *g23* sequence analysis, only two RB49 phages were part of our collection. The observed average coverage of 80 for RB49 phages suggested that a single phage accounted for 40-fold coverage. Based on this approximation, regions of average coverage suggest that at least 37 (1460 reads: 40 fold coverage) T4D-related phages are in our collection which is close to the number of 45 estimated from the *g23* sequences (Fig. 4). The regions of low coverage of the T4D genome indicated regions which are specific to T4D or are found in only one further phage from our cocktail. The sequence alignment with the T4D sequence revealed 18,000 small nucleotide polymorphisms (SNPs). Localized clusters of sequence variation emerged from a homogeneous background of bp diversity (Fig. 4). When analyzed for coding sequences, there was a clear bias for bp changes to occur at the third codon position, which did not lead to amino acid changes in the encoded proteins (Fig. 5).

The average coverage of the JS98 genome with reads was 1150, which, assuming homogeneous sequencing efficacy, suggests that at least 29 (1150 reads: 40 fold coverage) JS98-like phages were in our collection and that is again close to the 36 JS98-like phages in our collection diagnosed by *g23* sequencing. However, the coverage of the reference JS98 genome was, with the exception of two small regions, even and fewer SNPs were observed than in the T4D alignment (Fig. 4). These observations suggest less diversification in the JS98 subgroup than in the T4D subgroup, which coincides with the analysis of the *g23* tree structure (Fig. 1).

The RB69 genome showed only an average 550-fold coverage, suggesting 14 (550 reads: 40 fold coverage) RB69-like phages were in our collection, which comes very close to the 16 RB69-like phages diagnosed by *g23* sequencing. Only one RB69 genome

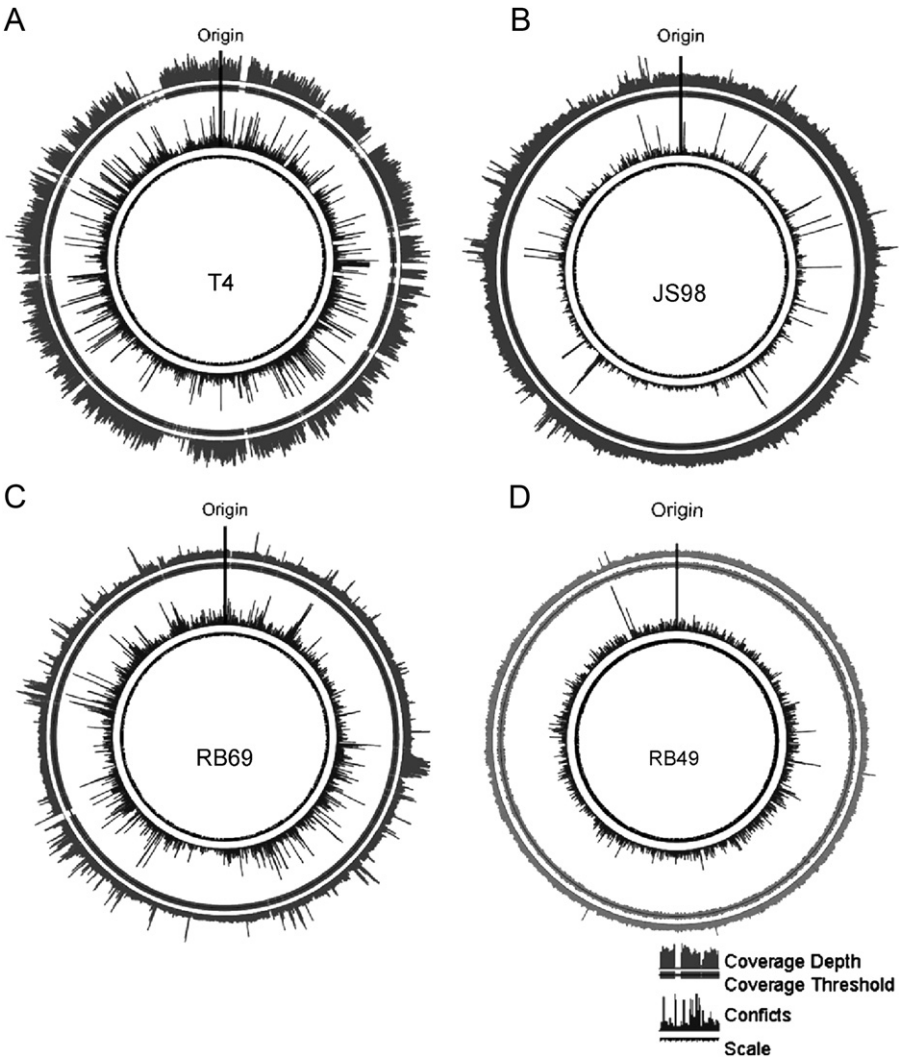

Fig. 4. Reference mapping to phage genomes (A) T4, (B) JS98, (C) RB69, and (D) RB49. Coverage of reference genome when mapped with the full set of Illumina reads and conflict report of SNPs for those mappings. From outer to inner most ring: Coverage depth; coverage threshold; conflicts; scale.

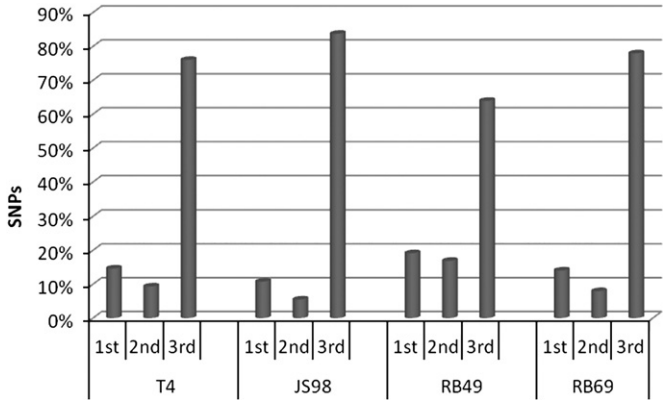

Fig. 5. SNP distribution of reads mapped to the four subgroup reference genomes T4, JS98, RB49, and RB69 by codon position, shown in %.

region had a low coverage (115,000 bp position) indicative of a region specific to the reference strain, and the RB69 SNP density showed an intermediate level between T4D- and JS98-like phages (Fig. 4). Only seven small regions of the RB43 phage genome were covered with reads (data not shown) demonstrating that this T4-like coliphage was not sampled in our survey.

Sequence analysis of the unmapped phage DNA

After removing repeated attributions, 10.5 million reads could be aligned with the reference genomes of the four T4 phage subgroups, accounting for 33% of total reads. It was considered that the unattributed reads could have been from contaminating bacterial DNA from the *E. coli* strain K-12 derivative on which the T4-like phages from our collection were propagated. This was, however, not the case, as only 0.02 million reads aligned with the K-12 genome. Notably, these reads were not distributed evenly throughout the *E. coli* genome, but were mapped to two small regions of the bacterial genome (Supplementary Fig. 1). One group of reads overlapped a tRNA-Lys gene, while the other group mapped to *yfiD*, a gene that encodes a pyruvate formate lyase and catalyzes a step in anaerobic glycolysis. In the T4D reference phage genome the corresponding gene is annotated as *vs.6* (Miller et al., 2003). It is located within a cluster of T4 genes (*vs.3* to *vs.8*), which otherwise lack database matches. Sequences related to *yfiD* were found in the sequenced genomes of several T4- and RB69-like phages, but not in JS98-like phages. We tested whether the possession of this gene provided a growth advantage

under anaerobic conditions to the corresponding T4 phages by comparing plaque counts under aerobic and anaerobic growth. No relative titer increase under anaerobic culture was seen in T4 phages possessing *vs.6* (data not shown).

The plasmid pDM30 was added at a low concentration to the T4-like phage DNA pool. This plasmid contained three antibiotic resistance genes that served as positive controls for the screening of undesired genes. The 4200 plasmid-specific reads achieved a 19-fold coverage of the entire plasmid genome, and such a small amount of reads excluded plasmid DNA as a source for the unattributed DNA reads.

In order to investigate the identity of unattributed reads, redundant reads were removed from the total set of sequences and reference mapping was repeated. The total number of 31 million reads was reduced to 3.7 million non-redundant reads, of which 2.3 million reads could be mapped on the four T4 subgroup reference phage genomes. Not accounting for shared reads, 24%, 15%, 17%, and 6% could be aligned with the reference phages T4D, RB69, JS98 and RB49 of the four subgroups, respectively (Supplementary Table 1). Reads that mapped to addition phage genomes (JS-10 and JSE, data not shown), to *E. coli* K-12, or to the plasmid were removed, and 1.37 million reads were left unattributed.

We projected these non-attributed reads on coliphage genomes deposited in the NCBI database. Approximately 0.2 million reads were attributed to T4-like phages wV7, RB51, and Shfl2. These reads were not evenly distributed along the genome (which would have suggested new subgroups), but clustered over distinct genome regions (Supplementary Fig. 2). The conclusion that these reads were not associated with new subgroups was further confirmed with *g23* analysis, which placed these phages within the T4D-like branch. An additional 0.1 million reads were aligned with T4-like phages SP18 and vB-EcoM-VR7. They showed a relatively even, but low coverage distribution on these genomes (Supplementary Fig. 2). Their *g23* sequence identified them as distant members of the JS98 branch. Finally, 0.06 million reads were aligned with T4-like phages IME08 and Bp7. Together, these matches accounted for 0.29 million T4-related DNA reads.

Heterologous coliphage genomes from the NCBI database did not contribute sequence matches with DNA from our collection, except for Myovirus P1 and Siphovirus T1. The T4 phages from our collection shared sequence relatedness with two genes from P1. However, these were not phage genes, but rather mobile genetic elements (insertion sequences IS1 and IS5). In contrast, the T1 genome was covered with 0.03 million reads (Supplementary Fig. 3). The T1 genome was not covered over its entire length, the coverage depth of the aligned regions was low, and showed many SNPs.

We next explored the 1.37 million reads that could not be attributed to the established four subgroups of T4-like phages by yet another approach: 0.74 million could be *de novo* assembled into 1523 contigs, with the longest being 21-kb long. Most contigs were small ($n=862$ were only 0.2–0.5 kb long), and only $n=234$ contigs were larger than 0.5 kb. For 600 contigs the closest homolog was shown to be a T4 phage from the four reference subgroups (identity at the protein level). Almost exclusively, the remaining contigs shared their best hits with T4-like phages from enterobacteria. Most numerous were matches with *E. coli* phages IME08, wV7, Bp7, VB_EcoM-vR7, RB32, and RB51, followed by matches with T4-like Shigella phages SP18 and Shfl2 (data not shown). The exceptions were matches with Siphovirus T1 DNA encoding terminase, portal protein, major head, tail, tape measure, tail fiber genes and DNA primase, helicase, and Dam methylase. Only one contig shared significant sequence identity with bacterial DNA (52% aa identity with the *Enterobacter cloacae* enzyme deoxyheptonate aldolase, which is involved in aromatic amino acid metabolism).

Finally, BLASTN searches were used to investigate the 0.63 million non-redundant reads that could not be assembled into contigs. When a random subset of 30,000 reads were analyzed by BLASTN homology searches, 65% met stringency criteria to be considered a positive hit. Of the positive hits, 95% of them gave a hit with phages, practically all of which were represented by T4-like phages.

Search for undesired genes

We screened all sequences from the phage collection sequencing against three different databases; an internally maintained database of undesired genes (DUG) and two public databases compiling antibiotic resistance genes (ARDB) (Liu and Pop, 2009) and bacterial virulence factor genes (VFDB) (Yang et al., 2008). When the Illumina sequences were screened against the DUG, the largest of the three databases, we obtained hits with 1555 entries into the database. Practically all hits were represented by a single 50-bp match of phage sequence with the undesired genes. When we asked for two or more 50-bp alignments from the phage collection sequence pool with the database entries to qualify for a hit, only 66 genes fulfilled this inclusion criterion, and these matches were all investigated manually. Fifty-five of these hits were with genes annotated as holin, lysin, tail fiber adhesion, tail fiber assembly, integrases, and transposase genes, and thus represented conventional phage genes. Of the 11 hits that remained, these genes included several metabolite transporter genes, a transcriptional regulator, and an elongation factor. The search also identified an acridine resistance gene. However, this is not an antibiotic resistance gene, but a T4 gene involved in phage DNA replication (Rappaport et al., 1974; Wang and Ripley, 1998). Weak and partial matches were observed with capsular polysaccharide synthesis, ferric citrate transport system of *E. coli* (Fec system), and ferric enterobactin transport proteins (Fep genes).

No further hits were obtained when screened against the ARDB and VFDB databases.

Clinical observation of adult volunteers for adverse effects

To assess the biological safety of our T4-like phages, a cocktail of nine phage isolates from our collection was selected for application in healthy adult volunteers from Bangladesh. The phages were propagated individually on the non-pathogenic *E. coli* strain K-12 and isolated by a combination of differential centrifugation and sterile filtration steps (Bruttin and Brüssow, 2005). The phages were chosen based on their lytic activities on pathogenic *E. coli* serotypes associated with pediatric diarrhea (data not shown). Based on a previous safety test in Swiss adult volunteers who were treated with 3×10^7 pfu of total phage and showed a good fecal bioavailability of the oral phage (Bruttin and Brüssow, 2005), we decided to use a 100-fold higher dose to test for possible adverse effects before considering the extension the safety tests to healthy and then hospitalized children.

The safety study group consisted of 15 healthy adult volunteers recruited by ICDDR,B. The volunteers received the phage at a high dose (3×10^9 pfu of total phage), at low dose (3×10^7 pfu of total phage), and placebo. Each subject received all three treatments in a random order and thus served as his/her own control, allowing for the detection of adverse effects with a small group of exposed volunteers. However, this type of design, known as repeated measures, uses internal controls and has no formal control group, *i.e.* subjects were studied before the intervention and also when given placebo to represent internal subject-specific controls. This type of design presents some problems of analysis since certain types of delayed effects cannot always be excluded.

Table 1
Characteristics of the study subjects and observation of abnormal observations in physical examination and clinical chemistry or hematology analyses according to treatment mode for each individual volunteer.

Subject	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Normal range
Characteristics of subjects																
Age (yr)	27	32	40	40	22	25	25	27	24	22	25	26	26	23	39	
Sex	m	m	m	m	m	m	f	f	f	f	f	m	m	m	m	
Weight (kg)	63	59	60	57	57	67	49	51	53	54	54	65	68	59	59	
Abnormal physical measures																
Weight changes > 0.7 kg																< 0.7 kg (1.5%)
Pulse rate																60–100/min
Respiration rate		1,2		1	0,1									0		12–25/min
Fever														1,2		≤ 37.5 °C
Systolic blood pressure	0,1,2		2		2	0		0,1				0	2			110–140 mmHg
Diastolic blood pressure						2										60–85 mmHg
Clinical chemistry																
Na^+							1									136–145 mEq/L
K^+			*0*	1	1								0			3.5–5.1 mEq/L
Cl^-	*0*	1											*0*			98–107 mEq/L
Total CO_2	0,2	0,2					2	*0*			1,2	*2*	0,1,2	1	*0,1,2*	23–29 mEq/L
Ca^{2+}																2.1–2.6 mmol/L
Creatinine													1			80–115 μmol/L[a]
Blood urea nitrogen	0							*0*				*0*				6–20 mg/dL
Total serum protein	0,1		2		*0,1,2*		*0,1*	*0,2*	*0*	*0,1*	*0,2*	*0,2*				6.0–8.3 g/dL
Albumin																3.4–5.4 g/dL
Total cholesterol															*0,1*	< 5.18 mmol/L
Triglycerides			0,1,2								*0,1,2*					< 1.7 mmol/L
Total bilirubin								0,1,2								5.1–17 mmol/L
Direct bilirubin									0,1							1.7–5.1 mmol/L
Alanine aminotransferase					1							0,1,2			2	< 56 U/L
Aspartate aminotransferase			2		1							0,2				< 40 U/L[b]
γ – Glutamyl Aminotransferase																< 40 U/L
Fasting blood glucose				1	2							0			0,2	3.9–5.6 mmol/L
Hematology																
Platelet count										0,1						170–500,000/μL
Red blood cell count																4.3–6.2 × 10^6/μL[c]
Mean corpuscular volume				0						0,1						82–102 fL[d]
Mean corp. hemoglobin conc.			0	0	0	0										31–35 g/dL
Hemoglobin									2							13.2–16.2 g/dL[e]
Hematocrit									0,2							40–52%[f]
White blood cell count	0		0,1,2	1						*0,2*	*0*					3.9–11 × 10^9/L
Neutrophils	0					1,2					2					45–74% WBC
Lymphocytes	0			1		2					2			0		25–45% WBC
Monocytes	0					2								0	1,2	3–7% WBC
Eosinophils			0,1,2	*2*		*0,1,2*	*1*		*0,2*				*0,1,2*			1–7% WBC
Adverse events	0[g]							1[h]				1[i]				

The first three rows give the physical characteristics (age in years, weight in kg, sex: m, male) of the 15 study subjects (columns).
The next six rows note individually for each subject if abnormal measurements were made during the physical examination. No indication means that the volunteers showed only values that were within the normal range as defined for the respective measurement in the last column. Abnormal measurements were noted with 0, 1 or 2 if they were observed during the treatment with placebo (0), low phage (1) or high phage (2) preparation. If the abnormal values are italicized, an abnormal value for the given value was already observed in the corresponding volunteer during the screening before any application of the test products. Many normal subjects who were judged healthy for the local standard had "abnormal" values already at baseline, *e.g.* high eosinophil percentage, which might reflect local parasite and worm exposure.
The next rows give in a similar way observation of clinical chemistry values and hematology values which are outside of the normal range.
Notes: For a number of values the normal range is gender-specific. The last column gives the normal range for men. The respective values for women are as indicated here: a, 53–97 μmol/L; b, < 34 IU/L; c, 3.8–5.5 × 10^6/μL; d, 78–101 fL; e, 12.0–15.2 g/dL; f, 37–46%.
Adverse events: g, abscess; h, mild dyspepsia; i, mild abdominal discomfort.
Reading examples: Volunteer 1 shows 0,1,2 for systolic blood pressure which means that he had blood pressure outside of the norm value when on placebo, low phage and high phage. Volunteer 2 has no entry in systolic blood pressure meaning normal measurements during all treatment periods.

Clinicians measured the vital parameters, assessed the kidney and liver functions as well as the hematology status of the study subjects, and questioned the subjects for gastrointestinal complaints. The results of these analyses are summarized in Table 1.

Overall we did not observe a consistent pattern suggestive of an adverse effect of phage therapy. Particularly, subjects showing shifts in white blood cell proportions (volunteers 4, 6, 11, 15) were not the same as those who showed elevated aminotransferase levels (volunteers 3, 5, 15); nor did they fit to the subjects who reported adverse events (volunteers 1, 8, 12). For details see Supplementary text 1.

Phage detection

Twenty-seven out of 42 investigated stool samples (64%) from subjects treated with the high-titered oral phage dose yielded phage on the *E. coli* indicator cell K-12. To be scored as phage-positive, the stool sample had to contain > 100 pfu/g stool. In comparison, 13 out of 44 stools (30%) from subjects treated with the low-titered phage dose and 12 out of 43 stool samples (28%) from subjects treated with placebo were phage positive. The same prevalence of phage-positive stool samples was found using the *E. coli* indicator strain WG-5. The data are heterogeneous,

supporting a significant difference between the treatments ($p=0.0006$, Chi-square test). Moreover, pair wise comparisons indicate statistical differences between high dose phage and the other two groups ($p<0.001$), but no difference was detected between low dose phage and placebo ($p>0.05$).

Two volunteers showed relatively high phage titers in the feces when on placebo. This is not carry-over of high phage application from the previous treatment since in one case the high phage titer was detected in the first week of treatment (Supplementary Table 2, subject 15).

Only a single stool sample contained a higher phage titer than that which was orally applied. In all other fecal samples, titers were well below the level of orally applied phage. The median titer of all phage-positive stools samples was 2×10^2 pfu/g feces. We estimated that less than 1% of the orally applied phage was recovered from the stool (Supplementary Text 2).

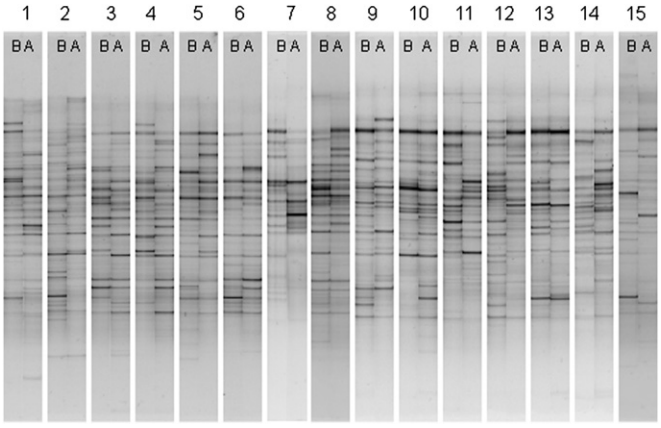

Fig. 6. DGGE profiles obtained with universal primers HDA1-GC and HDA2 targeting the V3 region of the 16S rRNA gene in bacteria from the stool of 15 adults from Bangladesh at baseline (B) and after (A) receiving all three treatments over a 3-wk intervention period.

In vitro experiments simulating stomach and duodenal gut conditions suggested that the low pH value expected in the adult stomach are detrimental to phage (Supplementary Table 3). With this low survival of oral phage in Bangladeshi adults, and the low amount of protein that even the high phage dose represents (less than a microgram), it was not surprising that neither phage, nor LPS (as contamination of the phage preparation), nor phage-specific or LPS-specific antibodies were detected in the serum of the study subjects.

DGGE

A frequently quoted advantage of oral phage over antibiotic application is the lack of impact of phage on the gut microbiota beyond the target species (Sulakvelidze et al., 2001). This is particularly the case for *E. coli* phages, where the target pathogenic bacterium in healthy subjects represents only a small minority of the gut bacteria. The enrolled volunteers did not show *E. coli* colonies in their stool samples ($<10^2$ pfu/g feces on EMB or McConkey agar). Lack of effect of oral phage on the gut microbiota has, however, not been demonstrated in humans. To investigate this effect, we used two different experimental approaches to assess changes in the gut microbiota composition. In the first approach, bacterial DNA was extracted from the fecal samples of the volunteers at admission and at the end of the 3-wk safety study. PCR was performed targeting the V3 region of the 16S rRNA gene using universal primers. The PCR products were then analyzed by denaturing gradient gel electrophoresis (DGGE) for band pattern variation for each individual volunteer before and after application of all three test products (Fig. 6). Profiles were unique for each individual, as expected. Pair wise calculations using the Dice's similarity scores showed that profiles from the same subject were on the average Dsc=85.4% similar. The difference between the two stool samples from the same subject collected before and after the intervention was rather small for some volunteers (Dice's similarity scores Dsc=92.1, volunteer 2), while it was more substantial for other individuals (Dsc=78.0, volunteer 11).

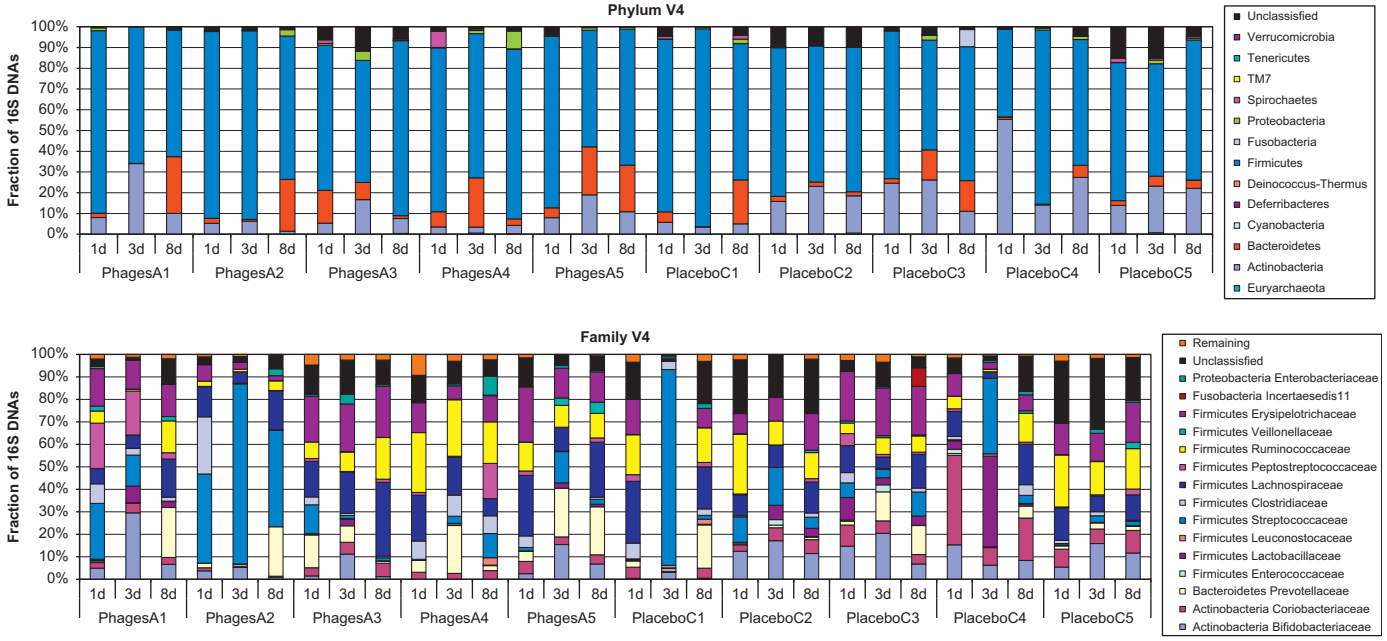

Fig. 7. Bacterial community structure profiles for fecal samples from five human volunteers receiving the high dose T4 phage cocktail (subjects A1–A5) and five volunteers receiving placebo (subjects C1–C5). The profiles are given for day 1 (before first intervention) and for days 3 and 8 (*i.e.* 1 and 6 days after the intervention). The top panel gives the fraction of 16S rRNA (in percent) amplified with universal primers targeting the V4 region at the phylum level and the bottom panel at the family level. The sequences attributed to a specific phylum or family are depicted in different colors, the color code is given at the right side.

Microbiota profiling by pyrosequencing

Next, we investigated stool samples from five volunteers treated with placebo and five volunteers treated with the high dose of oral phage. Stool samples were taken from the volunteers just before phage administration, and then one and five days after the 2-d application of the test products to the volunteers. We only used stool samples collected during the first week of intervention to avoid interference by previous treatments. When DNA was extracted and sequenced, approximately 1200 sequence reads were obtained per stool sample and were assigned to bacterial taxonomic categories. Sequencing data from the V1+V2 and the V4 region gave comparable results for the same stool samples. Hence, only results from the V4 region are presented in detail.

At the phylum level, the majority of 16S rRNA sequences were assigned to Firmicutes, Actinobacteria, and Bacteroidetes (in decreasing order) with small contributions coming from Proteobacteria and Spirochetes (Fig. 7, Top). Firmicutes were present in all 30 investigated stool samples and represented, with one exception (subject C4, day 1), the numerically dominant bacterial phylum. At family rank, Firmicutes were mainly represented by Streptococcaceae, Lachnospiraceae, Ruminococcaceae, and Erysipelotrichaceae (Fig. 7, Bottom). Subjects with high fecal read numbers for Streptococcaceae showed at the operational taxonomic unit (OTU) level annotations of *Streptococcus lutetiensis* (10% of all reads). Actinobacteria and Bacteroidetes were the next most frequent fecal annotations. Bacteroidetes were at the genus level mostly Prevotella. Actinobacteria were represented mainly with Bifidobacteriaceae and Coriobacteriacieae. At the OTU level the dominant Bifidobacterium species was *Bifidobacterium ruminantium* (5% of all reads).

Many subjects showed changes in the relative abundance of Actinobacteria and Bacteroidetes over the 8-d observation period. However, there were no consistent trends and no differences were seen between placebo and phage recipients. Enterobacteriaceae were detected in all but one subject, but they represented, with one exception, only 0.5–2.0% of the reads. Phage application was not associated with an abundance decrease of Enterobacteriaceae.

The ratio of bacteria before and after the intervention was analyzed for the major fecal bacterial families (data not shown). Data points from phage recipients overlapped those from placebo recipients. When the fecal microbiota data were compared for overall bacterial diversity and richness at strain level (99% identity), phage-treated subjects displayed the same range as placebo-treated patients (Supplementary Fig. 4).

Discussion

Our field survey of coliphages in Bangladesh and Switzerland led to a phage collection containing 99 T4-like isolates belonging to four previously described subgroups (T4D-, RB69-, JS98- and RB49-like phages). Our survey covers only a short time period and only two, although very diverse geographical regions (Chibani-Chennoufi et al., 2004a,b), and thus might not explore the full diversity of T4-like coliphages. As we used a *g23*-based PCR test as inclusion criterion for phages into our collection, we might have selected against more distantly related T4-like coliphages which gave only weak *g23* PCR signals. A further limitation might be the use of a single *E. coli* indicator strain.

When considering the genome structure of lambdoid coliphages (Casjens and Thuman-Commike, 2011), one might wonder whether the *g23* sequence is a reliable marker for the genome diversity of T4-like phages. By projecting the sequences from our collection on the genomes of established T4 subgroups, we could get an independent estimation for the contribution of established subgroups to our collection. This approximation came very close to the attribution of the phages based on their *g23* sequences, therefore confirming *g23* as a reliable marker for the entire genome classification. This result might appear trivial, but it underlines that the modular mode of evolution displayed by lambdoid phages does not apply to T4-like phages. In the latter, the vast majority of the genome is conserved within a subgroup ("conserved core") (Comeau et al., 2007, 2010), which diversifies mostly by the accumulation of point mutations, the shuffling of single genes via mini-circles (Arbiol et al., 2010), and only in some genome segments by modular exchanges (Tétart et al., 1998). From the genetic safety point of view, this characteristic of T4-like phages is a clear advantage over the pervasive modular mode of evolution in lambdoid phages. The knowledge of the *g23* sequence alone thus predicts a vast majority of the genome sequence.

This monotonous mode of evolution does not exclude horizontal gene transfers (HGT) between *E. coli* and T4 phages and between different T4 phages (Arbiol et al., 2010). In fact, T4 phages shared a tRNA gene and a gene encoding a glycolytic enzyme. The latter gene is probably a HGT from *E. coli* to some (T4, RB69), but not other (JS98) T4-like phages. The 98% bp identity between *yfiD* of *E. coli* and *v.6* of T4 suggests a recent HGT event. Whether this gene provides a selective benefit to T4 phages possessing *v.6* could not be determined, as increased growth under anaerobic conditions was not associated with this gene. A phage possessing a glycolytic enzyme does not represent a genetic risk, and T4D containing the *v.6* gene was tested in Swiss volunteers without producing adverse effects (Bruttin and Brüssow, 2005).

When considering the non-redundant DNA reads, 2.3 out of the 3.7 million reads (62%) were mapped on the four established T4 subgroups, and subsequently raised the question about the source for the remaining 38% of reads. Part of them projected on T4-like phages like vB_EcoM-VR7 (Kaliniene et al., 2011), which is according to the *g23* tree analysis a peripheral member of the JS98 subgroup. Another part projected on T4-like phages like RB51, an inner member of the T4D *g23* branch, which showed unique DNA sequences over particular genome regions. Half of the remaining 1.4 million reads could be assembled into 1523 contigs, the majority of which matched T4-like phage genomes. In addition, the obtained DNA sequences were also screened against the genomes of non-T4 coliphages deposited in the database. We found sequences that matched two genes of phage P2, but otherwise genes from other taxonomical groups of coliphages were conspicuously absent, suggesting that T4 phages are relatively resistant against HGT from other coliphage classes. We detected sequences that covered more than half of the T1 reference genome through sequence mapping, and in parallel analysis of *de novo* contig assembly. T1 genes within assembled contigs were found in the vicinity of other T1 genes, frequently in the same order as in the T1 genome map, and never in the vicinity of T4 phage genes. This display pattern of T1 genes therefore did not provide evidence of gene exchange between T1 and T4 phages, but most likely of a T1 phage contaminating one of our T4 phage isolates. T1 was never introduced into our laboratory, but a phage was once isolated from stool which resembled T1 in morphology (Chibani-Chennoufi et al., 2004a,b). T1 is difficult to remove from laboratories working with *E. coli* due to its high resistance to desiccation (Rotman et al., 2010). However, T1 does not represent a genetic risk since it does not carry undesired genes (Roberts et al., 2004). In addition, with primers based on the obtained T1-like sequences, it would be possible to screen the individual phage isolates for the T1 contaminant and eliminate the identified phage isolates from our collection.

The establishment of the safety of a given phage cocktail cannot be achieved by genome analysis, but ultimately from *in vivo* tests in animals and in humans. In the present study we gave a cocktail of nine T4-like phages from our collection representing three of the four subgroups to adult volunteers, and we did not observe adverse effects. We used a large number of clinical and laboratory tests to detect harmful effects on liver, kidney, gut, and hematology functions. The phage dose was 100-fold higher than in a previous safety test with Swiss volunteers. One may still contend that the absolute amount of phage given to the volunteers was still low, corresponding to only a few micrograms of phage protein, and indeed, this dose was lower than that determined to be necessary to achieve treatment effects in previous animal studies (Jamalludeen et al., 2009). However, phage doses as high as 10^{11}–10^{12} pfu would present problems with product availability and costs.

Corresponding *in vitro* experiments suggested that stomach acidity was a major block to a higher phage transit and may account for the fact that less than of 1% of orally applied phage was recovered in the feces. This current study figures into a series of planned clinical trials aimed ultimately towards pediatric hospitalized diarrhea patients, for which a much lower dose than that of the current study would be used. Since only the high and not the low dose of phage led to fecal phage detection and was significantly higher than in placebo recipients, such a phage dose might seem problematic. This problem can be remediated by applying the phage together with bicarbonate, which leads to a transient neutralization of stomach acidity and much better fecal phage recovery (Jamalludeen et al., 2009). This approach is also common practice for human application of phages in Russia and Georgia (Sulakvelidze and Kutter, 2005). However, since stomach acidity is also a major hurdle against enteric over-infection in patients hospitalized on a diarrhea ward, we aimed to avoid bicarbonate application. Children have a lesser acidic stomach content than adults and phage transit might be correspondingly better when administered to children, even without the use of bicarbonate. Therefore a lower dose may prove to be sufficient.

The fact that we achieved gut transit with the high phage dose, even in the absence of bicarbonate, led us to believe that one can achieve therapeutically efficient doses in *E. coli* diarrhea patients. In theory, phage is a self-amplifying antimicrobial agent (Payne and Jansen, 2003) and even the low doses used in the present test should reach therapeutic levels after replicating on the target pathogen in the gut. If the same amount of phage is given orally to patients suffering from acute *E. coli* diarrhea, we cannot, however, automatically anticipate safety of the treatment. Phage may amplify to high titers in the intestine and cause side effects not observed under the conditions of the present test, such as a massive release of lipopolysaccharide (“endotoxin”) from phage-lysed *E. coli* cells. Therefore, further safety trials will have to be conducted with T4-like phages in subjects acutely infected with *E. coli* diarrhea.

The current volunteers did not contain *E. coli* in the feces, and thus the lack of phage amplification was not surprising. Considering that phages are target species-specific, the lack of a significant impact of oral phage application on the fecal microbiota composition was expected. In DGGE analysis, the Dice band-based coefficient values ranged from 78% to 92%. Similar values were reported for stool samples from four Belgian subjects collected over three months (Vanhoutte et al., 2004). This conservation of the fecal microbiota strongly contrasts with the deleterious impact of antibiotics on the gut commensals that fulfill not only physiological, but also protective functions.

Finally, the small number of volunteers and the use of a repeated measure approach on the same volunteers presented certain limitations to our safety evaluation. However, based on the presented data we obtained the approval of the ethical committee to pursue the safety evaluation first in healthy children and, if uneventful, also in diseased children. In addition, one should not forget that unintended large scale safety test with orally applied phages are conducted daily in millions of consumers of fermented food like yoghurt, cheese, and sauerkraut, which contain inevitably bacteriophages directed against the starter bacteria used in food production. No adverse effects linked to dairy phage exposure were ever reported. Furthermore, T4-like phages are, as also demonstrated by our ecological survey, natural members of the human gut ecosystem and we only isolated and amplified what was already naturally present in the gut.

Subjects and methods

Recruitment of volunteers

Thirty healthy adult volunteers were screened in a periurban community of Nandipara, located 12 km from the International Center for Diarrheal Diseases Research (ICDDR,B) in Dhaka, Bangladesh. Subjects were recruited if they had not suffered from diarrhea or other illness in the preceding two weeks, if their blood count, liver and kidney functional tests were normal, and if they were negative for antibodies to hepatitis A, B, C and E virus. Further exclusion criteria were systemic infection, body-mass index (BMI) below 21, use of prescription drugs during the last two weeks, antibiotic use during the last four weeks, drug or alcohol abuse, a clinically significant medical history, or abnormal vital signs. The purpose of the study was explained to the volunteers in the local language and those willing to comply with the study procedure and who signed a written informed consent were enrolled.

Safety study

The study was a single-centered, randomized, double-blind, placebo-controlled trial consisting of a three period cross-over comparison for two different doses of an oral T4-like phage cocktail. The dose of 3×10^7 pfu was chosen because a previous trial with Swiss adults showed that this dose resulted in phage-positive stools in all volunteers with peak stool titers of 3×10^4 pfu/g. This dosage was estimated sufficient to control bacterial targets within the gut. The higher dose of 3×10^9 pfu was chosen to test for dose-dependent adverse effects. The phages were purified as described previously (Bruttin and Brüssow, 2005): the stock was diluted into 150 mL mineral water and given as three divided doses of 50 mL during daytime, 6 h apart. This treatment was given over two days followed by a wash-out period of five days. The placebo consisted of the corresponding volume of mineral water (Vittel, Nestlé) without added phage. Vittel water has a pH of 7.3 and contains 258 mg bicarbonate per liter as a weak buffer against gastric acidity. The time line for the product application and medical examination of the subjects is given in Supplementary Fig. 5. The product attribution to the subjects was made according to a randomization table elaborated by a biostatistician.

Adverse events were defined as signs, symptoms or laboratory analyses in the volunteer that could not be related to their current health status whether or not they have a causal relationship with the application of the study product. During the days of observation, the volunteers were on the usual diet of the ICDDR,B hospital; the loss of wage was compensated, but no financial remuneration was paid to discourage participation for economic needs. The clinical protocol (ICDDR,B research protocol #2008-062) was first evaluated by external scientific reviewers and subsequently approved by two independent Institutional Review

Boards: first the Research Review Committee (RRC) and then the Ethical Review Committee (ERC) (RRC approval on May 11, 2006 and ERC approval on July 26, 2006). The study started March 7, 2007 and finished on April 30, 2007.

All clinical and laboratory observations were compiled in an electronic file, which was sent to the Nestlé Research Centre (NRC). When all data had been compiled, the code was broken to prepare the report requested by the ERC.

Phage cocktail

The cocktail consisted of nine independent isolates of T4-like *E. coli* phages, including four phages closely related to the T4D reference phage: phages NPC (Nestlé Phage Collection) 1008, NPC 1009, NPC 1031 and NPC 1003; four JS98-like coliphages: NPC 1000, NPC 1024, NPC 1002 and NPC 1007 and one RB49-like coliphage: NPC 1006. All phages were tested for T4-like phage morphology and purity by negative staining electron microscopy; *g23* gene sequencing (Tétart et al., 2001) confirmed that the chosen phages belonged to defined T4-like phage subgroups (Zuber et al., 2007) with at least two completely sequenced phage representatives showing no undesired genes with respect to human or environmental safety (Brüssow, 2010).

Phage tests

Phage titration in stool samples was done with plaque assays on the two *E. coli* indicator strains K-12 and WG-5, as described previously (Bruttin and Brüssow, 2005). Total daily stool volumes were not measured. To convert the measured phage density in the stool (expressed as pfu/g stool) into a stool phage output, we used data on stool output of healthy adults from a comparable population. Compared to Western communities, the population from developing countries showed higher stool weight, higher frequency of defecation and shorter transit times linked probably to their diet which is higher in fiber and lower in fat. In this control population the mean 24 h stool weight was 408 ± 133 g and 295 ± 102 g in men and women, respectively. Daily stool weight increased by 42 g for every 10 kg increase in body weight and 27 g for every 10 years increase in age (〈www.ams.ac.ir/aim/index.html〉).

Phage T4-neutralizing antibodies were determined by the plaque reduction test as described previously (Denou et al., 2009).

Stability test

The phage cocktail or individual T4-like phages were incubated in the following solutions (0.05 M Tris–HCl pH 7 and 8; 0.05 M glycine–HCl pH 3; 0.5 M KCl–HCl pH 2) at 37°C for the time indicated in the text. As specified, trypsin (Type XI from bovine pancreas, T-1005 Sigma, St. Louis), pepsin (P-7000, Sigma) pancreatin (from porcine pancreas, P-1750, Sigma) or porcine bile extract (B-8631, Sigma) were added singly or in combination. The outcome measure was the survival of phage infectivity as determined in plaque assays.

LPS

Lipopolysaccharide (LPS, endotoxin) was detected in the serum samples with the Limulus amoebocyte lysate test (Charles River, Wilmington, Mass.). *E. coli* LPS-specific antibodies were determined by ELISA with isotype-specific anti-IgA-, anti-IgM- and anti-IgG-antibody conjugates as described previously (Brüssow et al., 1990).

Denaturing gradient gel electrophoresis (DGGE)

Fecal bacteria profiles of predominant groups were assessed by DGGE at baseline and after three weeks. Bacterial DNA extraction, PCR amplification of the 16S rRNA gene and DGGE analysis was done as described previously (Jaquet et al., 2009; Tannock et al., 2004).

Microbiota profiling by pyrosequencing

DNA was extracted from fecal samples using the GNOME kit (BIO 101, La Jolla, CA, USA) as previously described (Firmesse et al., 2008). The primers for PCR were designed as recently described (Frank et al., 2008; Hamady et al., 2008). The details for PCR, pyrosequencing and the analysis of the sequence data for microbiota analysis were done as described previously (Claus et al., 2011).

Sample preparation, sequencing, and sequence mapping

Phages were amplified and their DNA extracted as described previously (Chibani-Chennoufi et al., 2004a,b). A sample pool was created of equimolar concentrations of phage DNA from the 99 phages in the NPC and the pDM30 plasmid. Samples were processed *in vitro* to generate a DNA template library of short inserts following a genomic shotgun protocol (Fasteris SA, Switzerland). High-throughput DNA sequencing was performed using Hi-Seq 2000 technology (Illumina) to produce 50 bp reads. Sequencing quality of the reads was verified using FASTQC (version 0.10).

All reads were mapped to phage reference genomes (NC_000866, EF469154, NC_004928, NC_005066, NC_007023, EU863409, and EU863408), to the bacterial genome of the host strain used for phage amplification (NC_000913), and to the pDM30 plasmid, which contained resistance genes to chloramphenicol, ampicillin, and tetracyclin. Additional sequence mapping were done of the following phage reference genomes: wV7 (HM997020), RB51 (NC_012635), Shfl2 (NC_015457), IME08 (NC_014260), Bp7 (HQ829472), SP18 (NC_014595), vB_EcoM_VR7 (NC_014792), Shfl1 (NC_015456), T1 (NC_005833), P1 (NC_005856), T5 (NC_005859), TLS (NC_009540), lambda (NC_001416), Stx1 (NC_004913), HK97 (NC_002167), Mu (NC_000929), P4 (NC_001609), phiEco32 (NC_010324), P22 (NC_002371), T7 (NC_001604), N15 (NC_001901), T3 (NC_003298), EPS7 (NC_010583), and WV8 (NC_012749). The mappings were performed using Seqman NGen software (version 3, DNAstar) and were further analyzed for coverage, redundancy, and conflicts.

In order to reduce the complexity, reads with a 100% sequence identity were removed from the total reads to generate a set of non-redundant (nr) reads, and this set of reads was used again for sequence mapping against various reference genomes. Nr reads were then mapped against all genes present in an in-house Database for Undesirable Genes (DUG), described previously (Denou et al., 2009) in a safety screening, in addition to being screened against two public databases compiling antibiotic resistance genes (ARDB) and bacterial virulence factor genes (VFDB). Only hits having two or more reads aligned and covering more than 30% of the gene were considered as positive hits and were further investigated.

A *de novo* assembly was performed with the non-redundant, non-mapping reads using default parameters of SeqMan NGen. The contigs obtained were manually inspected for their closest homolog from blastx matches in the nr protein NCBI database. Reads not assembled into contigs were analyzed for their closest homolog from blastn matches within the EMBL database (e-value $\leq$ 1.00E-4, coverage $\geq$ 80%, percent identity $\geq$ 95).

Phage subgroup typing

The *g23* of NPC phages were amplified using MziaI (5′ TGTTATsGGTATGGTwCGyCGTGCTAT 3′) and CAP8 (5′ TGAAGTTACCTTCACCACGACCGG 3′) primers as follows: 30 cycles of 30 s denaturation at 96 °C, a 2 min annealing at 62 °C, and 3 min extension at 72 °C. PCR products were sequenced by Fasteris, SA. Sequences were trimmed to an averaged quality score threshold of 12 and then aligned by the Clustal W method (Thompson et al., 1994) in MegAlign (DNAstar).

Funding

The study was paid by the Nestlé Research Centre.

Acknowledgments

We thank Bertrand Bourqui and Isabelle Rochat for help in the DGGE analysis, Shahram Emady-Azar for help in statistical analysis and Ms. Rekha Chanda, Rownock Ara and Abdul K. Azad for their excellent help in recruitment of volunteers from the Nandipara Field Clinic of ICDDR,B.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.virol.2012.09.002.

References

Arbiol, C., Comeau, A.M., Kutateladze, M., Adamia, R., Krisch, H.M., 2010. Mobile regulatory cassettes mediate modular shuffling in T4-type phage genomes. Genome Biol. Evol. 2, 140–152.

Brüssow, H., 2010. Control of bacterial diarrhea with phages: coverage and safety issues of bacteriophage therapy. In: Sarbour, P.M., Griffiths, M.W. (Eds.), Bacteriophages in the Control of Food- and Waterborne Pathogens. ASM Press, Washington DC, pp. 273–295.

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.

Brüssow, H., Sidoti, J., Link, H., Hoang, Y.K., Barclay, D., Dirren, H., Freire, W.B., 1990. Age-specific prevalence of antibody to enterotoxigenic *Escherichia coli* in Ecuadorian and German children. J. Infect. Dis. 162, 974–977.

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.

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

Casjens, S.R., Thuman-Commike, P.A., 2011. Evolution of mosaically related tailed bacteriophage genomes seen through the lens of phage P22 virion assembly. Virology 411, 393–415.

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

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

Claus, S.P., Ellero, S.L., Berger, B., Krause, L., Bruttin, A., Molina, J., Paris, A., Want, E.J., Nicholson, J.K., 2011. Colonization-induced host-gut microbial metabolic interaction. mBio 2, e00271-10.

Comeau, A.M., Arbiol, C., Krisch, H.M., 2010. Gene network visualization and quantitative synteny analysis of more than 300 marine T4-like phage scaffolds from the GOS metagenome. Mol. Biol. Evol. 27, 1935–1944.

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.

Denou, E., Bruttin, A., Barretto, C., Ngom-Bru, C., Brüssow, H., Zuber, S., 2009. T4 phages against *Escherichia coli* diarrhea: potentials and problems. Virology 388, 21–23.

Firmesse, O., Mogenet, A., Bresson, J.L., Corthier, G., Furet, J.P., 2008. *Lactobacillus rhamnosus* R11 consumed in a food supplement survived human digestive transit without modifying microbiota equilibrium as assessed by real-time polymerase chain reaction. J. Mol. Microbiol. Biotechnol. 14, 90–99.

Frank, J.A., Reich, C.I., Sharma, S., Weisbaum, J.S., Wilson, B.A., Olsen, G.J., 2008. Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes. Appl. Environ. Microbiol. 74, 2461–2470.

Hamady, M., Walker, J.J., Harris, J.K., Gold, N.J., Knight, R., 2008. Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. Nat. Methods 5, 235–237.

Jamalludeen, N., Johnson, R.P., Shewen, P.E., Gyles, C.L., 2009. Evaluation of bacteriophages for prevention and treatment of diarrhea due to experimental enterotoxigenic *Escherichia coli* O149 infection of pigs. Vet. Microbiol. 136, 135–141.

Jaquet, M., Rochat, I., Moulin, J., Cavin, C., Bibiloni, R., 2009. Impact of coffee consumption on the gut microbiota: a human volunteer study. Int. J. Food Microbiol. 130, 117–121.

Kaliniene, L., Klausa, V., Zajančkauskaite, A., Nivinskas, R., Truncaite, L., 2011. Genome of low-temperature T4-related bacteriophage vB_EcoM-VR7. Arch. Virol. 156, 1913–1916.

Liu, B., Pop, M., 2009. ARDB—Antibiotic Resistance Genes Database. Nucleic Acids Res. 37, D443–D447.

Merabishvili, M., Pirnay, J.P., Verbeken, G., Chanishvili, N., Tediashvili, M., et al., 2009. Quality-controlled small-scale production of a well-defined bacteriophage cocktail for use in human clinical trials. PLoS One 4, e4944.

Merril, C.R., Scholl, D., Adhya, S., 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., Rüger, W., 2003. Bacteriophage T4 genome. Microbiol. Mol. Biol. Rev. 67, 86–156.

Payne, R.J., Jansen, V.A., 2003. Pharmacokinetic principles of bacteriophage therapy. Clin. Pharmacokinet. 42, 315–325.

Rappaport, H., Russel, M., Susman, M., 1974. Some acridine-resistant mutations of bacteriophage T4D. Genetics 78, 579–592.

Rhoads, D.D., Wolcott, R.D., Kuskowski, M.A., Wolcott, B.M., Ward, L.S., Sulakvelidze, A., 2009. Bacteriophage therapy of venous leg ulcers in humans: results of a phase I safety trial. J. Wound Care 18, 237–243.

Roberts, M.D., Martin, N.L., Kropinski, A.M., 2004. The genome and proteome of coliphage T1. Virology 318, 245–266.

Rotman, E., Amado, L., Kuzminov, A., 2010. Unauthorized horizontal spread in the laboratory environment: the tactics of Lula, a temperate lambdoid bacteriophage of *Escherichia coli*. PLoS One 5, e11106.

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

Sulakvelidze, A., Kutter, E., 2005. Bacteriophage therapy in humans. In: Kutter, E., Sulakvelidze, A. (Eds.), Bacteriophages—Biology and Applications. CRC Press, Boca Raton, pp. 381–436.

Summers, W.C., 2001. Bacteriophage therapy. Annu. Rev. Microbiol. 55, 437–451.

Tannock, G.W., Munro, K., Bibiloni, R., Simon, M.A., Hargreaves, P., et al., 2004. Impact of consumption of oligosaccharide-containing biscuits on the fecal microbiota of humans. Appl. Environ. Microbiol. 70, 2129–2136.

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.

Tétart, F., Desplats, C., Kutateladze, M., Monod, C., Ackermann, H.W., Krisch, H.M., 2001. Phylogeny of the major head and tail genes of wide-ranging T4-type bacteriophages. J. Bacteriol. 183, 358–366.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Vanhoutte, T., Huys, G.E., de Brandt, G., Swings, J., 2004. Temporal stability analysis of the microbiota in human feces by denaturing gradient gel electrophoresis using universal and group-specific 16S rRNA gene primers. FEMS Microbiol. Ecol. 48, 437–446.

Wang, F.J., Ripley, L.S., 1998. The spectrum of acridine resistant mutants of bacteriophage T4 reveals cryptic effects of the tsL141 DNA polymerase allele on spontaneous mutagenesis. Genetics 148, 1655–1665.

Weiss, M., Denou, E., Bruttin, A., Serra-Moreno, R., Dillmann, M.L., Brüssow, H., 2009. *In vivo* replication of T4 and T7 bacteriophages in germ-free mice colonized with *Escherichia coli*. Virology 393, 16–23.

Wright, A., Hawkins, C.H., Anggard, E.E., Harper, D.R., 2009. A controlled clinical trial of a therapeutical bacteriophage preparation in chronic otitis due to antibiotic-resistant *Pseudomonas aeruginosa*; a preliminary report of efficacy. Clin. Otolaryngol. 34, 349–357.

Yang, J., Chen, L., Sun, L., Yu, J., Jin, Q., 2008. VFDB 2008 release: an enhanced web-based resource for comparative pathogenomics. Nucleic Acids Res. 36, D539–D542.

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.