Temperate *Streptococcus thermophilus* phages expressing superinfection exclusion proteins of the Ltp type

Yahya Ali1,2,3, Sabrina Koberg1, Stefanie Heßner1, Xingmin Sun1†, Björn Rabe1†, Angela Back1, Horst Neve1 and Knut J. Heller1*  

1 Department of Microbiology and Biotechnology, Max Rubner-Institut (Federal Research Institute of Nutrition and Food), Kiel, Germany  
2 Medical Biology Department, Faculty of Medicine, Jazan University, Jazan, Kingdom of Saudi Arabia  
3 Department of Biotechnology, Agricultural Research Center, Animal Health Research Institute, Cairo, Egypt  

**Edited by:** Jennifer Mahony, University College Cork, Ireland  
**Reviewed by:** Fabio Dal Bello, Sacco Srl, Italy  
**Evelien M. Adriaenssens, University of Pretoria, South Africa  
**Present address:** Xingmin Sun, Microbial Pathogenesis, Department of Infectious Disease and Global Health, Tufts University, North Grafton, USA  
**Correspondence:** Knut J. Heller, Department of Microbiology and Biotechnology, Max Rubner-Institut (Federal Research Institute of Nutrition and Food), Hermann-Weigmann-Strasse 1, D-24103 Kiel, Germany  
e-mail: knut.heller@mri.bund.de

Lipoprotein Ltp encoded by temperate *Streptococcus thermophilus* phage TP-J34 is the prototype of the wide-spread family of host cell surface-exposed lipoproteins involved in superinfection exclusion (sie). When screening for other *S. thermophilus* phages expressing this type of lipoprotein, three temperate phages—TP-EW, TP-DSM20617, and TP-778—were isolated. In this communication we present the total nucleotide sequences of TP-J34 and TP-778L. For TP-EW, a phage almost identical to TP-J34, besides the ltp gene only the two regions of deviation from TP-J34 DNA were analyzed: the gene encoding the tail protein causing an assembly defect in TP-J34 and the gene encoding the lysozyme, which in TP-EW contains an intron. For TP-DSM20617 only the sequence of the lysogeny module containing the ltp gene was determined. The region showed high homology to the same region of TP-778. For TP-778 we could show that absence of the attR region resulted in aberrant excision of phage DNA. The amino acid sequence of mature LtpTP-EW was shown to be identical to that of mature LtpTP-J34, whereas the amino acid sequence of mature LtpTP-778 was shown to differ from mature LtpTP-J34 in eight amino acid positions. LtpTP-DSM20617 was shown to differ from LtpTP-778 in just one amino acid position. In contrast to LtpTP-J34, LtpTP-778 did not affect infection of lactococcal phage P008 instead increased activity against phage P001 was noticed.

**Keywords:** *Streptococcus thermophilus*, prophage, superinfection exclusion, TP-J34, TP-778L, TP-EW, TP-DSM20617

**INTRODUCTION**

Superinfection exclusion (sie) is generally known as a mechanism by which a prophage residing in a host cell prevents infection of the lysogenic host cell by other phage through blocking DNA injection (Donnelly-Wu et al., 1993). This protects the host from being lysed by the infecting and multiplying incoming phage, and hence the prophage will not be destroyed in the process of phage multiplication (McGrath et al., 2002; Mahony et al., 2008).

Sie has been mostly described for prophages of Gram-negative bacteria: P22 residing in *Salmonella typhimurium* (Hofer et al., 1995), Lambda-like phages in *Escherichia coli* (Cumby et al., 2012), and kappa-phage K139 in *Vibrio cholerae* (Nesper et al., 1999). Interestingly, sie has also been described for lytic T-even phages of *E. coli* (Lu and Henning, 1994). In Gram-positive bacteria, sie has been identified in prophages of corynebacteria (Groman and Rabin, 1982), *Lactococcus lactis* (McGrath et al., 2002), and *Streptococcus thermophilus* (Sun et al., 2006). One common feature of many of these proteins appears to be their targeting to the external side of the cytoplasmic membrane by either an N-terminal membrane-spanning helix (Mahony et al., 2008; Cumby et al., 2012) or a lipid-anchor (Sun et al., 2006). One exception appears to be the Glo protein of *Vibrio cholerae*, which has been described to be a soluble periplasmic protein (Nesper et al., 1999).

In temperate *S. thermophilus* phage TP-J34, a sie system is encoded by the ltp gene, residing within the lysogeny module. ltp is transcribed in the prophage state and encodes a lipoprotein, which is tethered to the outside of the cytoplasmic membrane, where it prevents injection of the DNA of the infecting phage into the cytoplasm of the host cell (Sun et al., 2006). Besides its rather weak activity against *S. thermophilus* phages, Ltp shows high activity against lactococcal phage P008 (Sun et al., 2006).

Ltp has been shown to consist of three different functional units: a lipid moiety for membrane anchoring, a serine-rich spacer region, and a repeat domain responsible for sie (Sun et al., 2006; Bebeacua et al., 2013). When expressed without its lipid-anchor, its host-range is extended to phages P335 and P001 belonging to different lactococcal phage species (Bebeacua et al., 2013). Thus, the active domain of Ltp may represent a broad-spectrum phage-resistance protein.

**ORIGINAL RESEARCH ARTICLE**

**doi:** 10.3389/fmicb.2014.00098

**published:** 13 March 2014

www.frontiersin.org

March 2014 | Volume 5 | Article 98 | 1
bacteria and phages. No such gene has been described for *L. lactis* strains and phages, respectively (Sun et al., 2006), although lactococci and streptococci and their phages are very closely related (Proux et al., 2002). Within the 11 publicly available sequenced genomes of *S. thermophilus* phages 2972, 5093, 7201, 858, ALQ13.2, Abc2, DT1, O1205, Sf11, Sf19, Sf21 <http://www.ncbi.nlm.nih.gov/genomes/GenomesGroup.cgi?opt =phage&taxid=102399&host=bacteria>, ltp determinants have not been identified. Phages O1205 (Stanley et al., 1997) and Sf21 (Brüssow and Bruttin, 1995) are the only temperate among the 11 phages. However, they are closely related to the virulent *S. thermophilus* phages (Brüssow and Bruttin, 1995; Lucchini et al., 1999; Desiere et al., 2002). They all together may form just one species (Quiberoni et al., 2010). A differentiation of the 11 phages according to their DNA-packaging mechanism resulted in two sub-species (Quiberoni et al., 2010), represented by Sf21 (cos-type) and Sf11 (pac-type) (Proux et al., 2002). O1205 belongs to the pac-type (Stanley et al., 1997), indicating that the type of infection is of minor importance for the relatedness of phages.

To investigate the distribution and diversity of members of the Ltp protein family among strains of *S. thermophilus* and to analyze the relatedness of phages carrying an ltp gene, we screened among *S. thermophilus* strains for prophages carrying genes similar to ltp. For two temperate phages - TP-J34L and TP-778L, we analyzed the whole genome sequences. Of the two other phages, TP-EW and TP-DSM20617, we determined the sequences of some selected DNA regions: ltp gene for both phages, lysogeny module for TP-DSM20617, and putative host specificity gene and lysin gene for TP-EW. The two Ltp proteins of phages TP-J34 and TP778 were functionally compared and found to differentially inhibit lactococcal phages. The differences in inhibition are discussed with respect to the differences found in the amino acid sequences of the two Ltp proteins.

**MATERIALS AND METHODS**

**BACTERIA AND PHAGES**

*S. thermophilus* strains used in this study were: J34 (lysogenic wild type), J34-6 (prophage-cured J34), SK778 (lysogenic wild type), DSM20617 (lysogenic wild type, German Collection of Microorganisms and Cell Cultures - DSMZ), and EW (lysogenic wild type).

The following phages were used: TP-J34 (wild type lysate, obtained by induction of the prophage) (Neve et al., 2003), TP-J34L (deletion derivative of TP-J34) (Neve et al., 2003), TP-778 (wild type lysate, obtained by induction of the prophage; this study), TP-778L (single plaque isolate from wild type lysate, this study), TP-DSM20617 (wild type lysate, obtained by induction of the prophage; this study), TP-EW (wild type lysate, obtained by induction of the prophage; this study).

The following lactococcal phages from our collection were used to test for infection-blocking activities of Ltp-derivatives: P197, P220, P624, P653, P684 (c2-species); P955, P957, P983, P993, P996 (936-species); P615 (P335-species). They had been assigned to species by electron microscopic inspection of their morphologies.

**GROWTH MEDIA, GROWTH CONDITIONS, PHAGE PROPAGATION, PROPHAGE INDUCTION, PHAGE-CURING, AND RELYSOGENIZATION**

*S. thermophilus* strains were routinely grown at 40°C in modified M17 medium containing lactose (th-LM17) (Krusch et al., 1987). For phage propagation, glycine-lysis medium was used: thM17 supplemented with 8 mM CaCl₂ and 1% glycine (Sun et al., 2006). Prophage induction was carried out with UV-light or mitomycin C. For UV-light induction, cells from a growing culture in log-phase were harvested by centrifugation, re-suspended in ½ volume of 0.1 M MgSO₄ and pumped through a quartz tube (internal diameter, 1.3 mm; length, 75 cm) placed under a laboratory 254 nm UV lamp (Schütz, Göttingen, Germany) at short distances (maximum 5 cm). Thereafter, the cell suspensions were mixed with another ½ volume of double-concentrated th-LM17 medium and incubated in the dark at 40°C. Induction was considered successful, when complete lysis was seen after ca. 3–4 h. For mitomycin C induction, different concentrations of mitomycin C (between 0.1 and 1 μg/ml) were added to growing cultures at early log-phase. Induction was considered successful, when turbidity increased for ca. 90 min after mitomycin C addition and then dropped to low turbidity levels.

Phage lysates were routinely centrifuged (Beckmann J2-21 centrifuge, 6000 rpm, 20 min, 4°C) and subsequently sterile filtered (nitrocellulose filters, pore size 0.45 μm).

Efficiency of plating was determined as described by Sun et al. (2006). Spot assays for determining the effects of Ltp-derivatives on phage infection were carried out by spotting 10 μl each of serial dilutions of phage lysates on agar plates overlaid with 0.75% top agar seeded with appropriate host bacteria.

All other relevant and specific information can be found in Neve et al. (2003).

**DNA TECHNIQUES**

Isolation of chromosomal DNA followed the method of Leenhouts et al. (1990) with some modifications. Ten ml th-LM17 medium (supplemented with 40 mM DL-threonine) was inoculated with *S. thermophilus*. Incubation proceeded at 40°C until an optical density at 620 nm (OD₆₂₀) of ca. 0.8 was reached. From 2 ml of the culture, cells were sedimented by centrifugation (Eppendorf microcentrifuge) and washed once with 0.5 ml buffer pH 8.0, containing 20% sucrose, 10 mM Tris-EDTA buffer of pH 8.0. Mixing by inversion, incubation proceeded for 1 h at 60°C. Finally, DNA was taken up in 200 μl Tris-EDTA buffer of pH 8.0.

Phage DNA was isolated from CsCl-purified phage with subsequent phenol extraction following the procedure described by Sambrook and Russel (2000).

Restriction analyses were done according to Sambrook and Russel (2000). Enzymes and recommended buffers were purchased from New England Biolabs (Frankfurt, Germany).

Agarose gel electrophoresis and Southern blot analysis were carried out as described by Sambrook and Russel (2000).
For digoxigenin-labeling of DNA, the “DIG DNA Labeling Kit” of Roche Diagnostics (Mannheim, Germany) was applied, following the manual of the supplier.

PCR was carried out on an Eppendorf Mastercycler 5333 or on a Perkin Elmer GeneAmp PCR System 9600. Primers (Table 1) were purchased from MWG Biotech (Ebersberg, Germany). The following pipetting scheme was used: 5 μl 10 × (NH₄)₂SO₄ buffer, 5 μl dNTPs (2 mM), 2 μl M-Twee 20 (2.5%), 1 μl of each of both primers (100 μM), DNA polymerase [10 parts Taq-polymerase (Qiagen, Hilden, Germany) plus 1 part Pfu-polymerase (Strategene, Amsterdam, The Netherlands), diluted 1:5 with distilled water], 1 μl template-DNA, bi-distilled water 34 μl. PCR was carried out as “hot start” PCR (D’Aquila et al., 1991), starting with 5 min at 95°C for denaturation, holding at 80°C for addition of polymerase, followed by 30 cycles involving denaturation (95°C for 1 min), annealing (at mean Tm of primer pair for 1 min) and elongation (72°C for variable duration: ca. 1 min for 1 kb expected length). Finally, PCR concluded with an elongation at 72°C for 5 min.

An internal 384 bp fragment of *ltp* was amplified by PCR as follows. The reaction solution in the thermal cycler contained 10 μl of 10× PCR kit buffer (Appligene Oncor, USA), 10 μl of dNTP-mix (Appligene Oncor, USA), 4 μl of Tween-20, 1 μl of both primers B and D (100 pmol/ml), 5 μl (0.1 μg) of DNA, 66.5 μl of H₂O and 2.5 μl of Taq DNA polymerase (1 unit/μl, Roche). Negative controls were set up similarly except that template DNA was omitted. Prior to cycling, the reaction mixture was heated to 95°C for 5 min, followed by 35 cycles of 30 s at 95°C, 30 s at 50°C, 30 s at 72°C and a final extension at 72°C for 7 min.

For “long-range” PCR (expected PCR products of up to ca. 4 kb), amplification was done following the “touchdown” protocol of Don et al. (1991). Primer pair D8+ and D12+ was applied. Annealing temperature in the first cycle was 10°C higher than the mean Tm of the primer pair. In the following 29 cycles, annealing temperature was reduced by 0.5°C per cycle. Finally, 10 cycles were added with an annealing temperature 1°C lower than the mean Tm of the primer pair. Elongation in that case was always 4 min.

Sequencing of the TP-J34 genome was done on a LI-COR 4200 system (MWG Biotech) according to the instructions of the supplier. Sequencing-PCR was done using the “Thermo Sequenase fluorescent labeled primer cycle sequencing kit with 7-deaza-dGTP (RPN 2438)” (Amersham Pharmacia Biotech, Freiburg, Germany), following the instructions of the supplier. Sequencing primers were labeled with fluorescence dye IRD800 (MWG Biotech). The sequence was completely determined for both DNA strands. It is available under EMBL accession number HE861935.1.

Sequencing of genomic DNA of TP-778L was done by AGOWA (Berlin, Germany) using 454 sequencing with an average coverage of approximately 20 fold. The sequence is available under EMBL accession number HG380752.1.

For sequencing of terminal ends of the integrated prophage and host DNA regions flanking the insertion sites, the following primers were applied: primer pair primer4 (targeting the gene encoding 50S ribosomal protein L19) (Bruttin et al., 1997) and int.cro.2 (targeting the cro gene of temperate *Streptococcus* phages) for amplification of the left primer pair lys.1 (targeting the lysin gene of temperate *Streptococcus* phages) and primer 3 (targeting an untranslated DNA region) (Bruttin et al., 1997) for amplification of the right flanking region. Both sequences are available under EMBL accession numbers HG917969 (left) and HG917970 (right).

The sequence of the DSM20617 prophage lysogeny module defined by primers 4 and Mz12.R binding sites was completely determined on both strands by primer walking. The sequence is available under EMBL accession number HG917971.

**Table 1** | Primers used for amplification of genomic DNA.

| Primer | Sequence | References |
|-------|----------|-----------|
| D8    | GGGTGGACCATAGGAGG | This study |
| D12   | ACCACTGAAATGCTACC | This study |
| D8+   | GGGTGGACCATAGGAGGGA | This study |
| D12+  | TCTACACCAACATGAAATGCTACC | This study |
| LysSup | GAACGAGCATTGAACACTAC | This study |
| Lysdown | CAGTTCAGCATCAGGTC | This study |
| terS-F  | GCTCATTTTGGGCGTCT | This study |
| terS-R  | CAAGGTCTATCTCGTCTC | This study |
| Itp-F  | TAGCAACACGGCTAGTCAGC | This study |
| Pri.C1-R | AAGCAAAAGAGTACGACATC | This study |
| lys1     | CCAAGCGCTTAAAGAGGCA | This study |
| 3       | CACATCTCCATCAAGGC | Bruttin et al., 1997 |
| 4       | GCAAGTGAACGCTGCAC | Bruttin et al., 1997 |
| Int.cro.2 | TTTTCTCCCATGCACTAACC | This study |
| Mz12.R  | ATAGCATATTATCGAATCGTCACG | This study |
| 8F     | AGAGTGTACCGTGCTAGCAG | Beumer and Robinson, 2005 |
| 1525R  | AAGCGAGGCTGACAGGC | Beumer and Robinson, 2005 |
| B      | GGCAAGCGTTGCTCGTGCTTGCTC | This study |
| D      | GGCGAATTCGCAACACGGCTAGTCAGC | This study |

**CLONING OF ltp**

Using primers *ltp*-XbaI and *ltp*-HindIII binding upstream and downstream, respectively, the *ltp*-TP–778 open reading frame was amplified by PCR. After restriction with the corresponding restriction enzymes the *ltp* orf was ligated into XbaI/HindIII-cleaved pMG36e. After transformation into *L. lactis* Bu2-60, transformed cells were selected and plasmids extracted. By DNA sequencing plasmid pYAL1-3 was confirmed to be the correct construct.

**SEQUENCE ANALYSIS**

For identification of open reading frames “orf finder” <http://www.ncbi.nlm.nih.gov/gorf/gorf.html> and “Artemis” (Rutherford et al., 2000) were applied. To obtain an overview over the major directions of transcription, only *orfs* with coding capacities larger than 100 amino acids were considered in a first draft. Gaps between *orfs* were inspected for potential *orfs* as small as ca. 50 amino acids by searching for appropriate start codons in connection with potential ribosome binding sites. For annotation
“blast” analyses were performed directly on the genes predicted by “orf finder” or “Artemis.”

tRNA genes were searched for by applying the “tRNAscan-SE” program of Lowe and Eddy (1997), and the “Tandem Repeat Finder” (Benson, 1999) was applied for searching for tandem repeats.

Functional assignment of gene products to protein families and identification of motifs of functional significance was done online <http://smart.embl-heidelberg.de/smart/set_mode.cgi?NORMAL=1> using SMART (Simple Modular Architecture research Tool) (Schultz et al., 1998; Letunic et al., 2009).

Dot plots were performed online <http://www.vivo.colostate.edu/molkit/dnadot/index.html>, (Maizel and Lenk, 1981) with the window size set to 13 and the mismatch limit set to 0.

For multiple sequence alignment, ClustalW at the EMBL-EBI website <http://www.ebi.ac.uk/Tools/msa/clustalw2/> (Larkin et al., 2007) or BLAST <http://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch&BLAST_SPEC=blast2seq&LINK_LOC=align2seq> (Altschul et al., 1990) was applied.

CRISPR spacer sequences were searched for at the “CRISPRs web server” by blasting phage genomic DNA sequences against the CRISPR database <http://crispr.u-psud.fr/crispr/BLAST/CRISPRsBlast.php> (Grissa et al., 2007).

RESULTS

S. thermophilus temperate phage TP-J34 carrying an ltp gene has been described in some detail (Neve et al., 1998, 2003; Sun et al., 2006). Isolation of TP-778 has also been described (Neve et al., 2004). It has been identified as related to but considerably different from TP-J34 by subjecting DNAs extracted from 142 S. thermophilus strains and digested by HindIII to Southern blots using digoxigenin-labeled TP-J34 DNA as a probe. In a further screening, more than 100 strains were tested by Southern hybridization with a probe generated from the ltpTP-J34 gene using primers B and D. Positive signals were obtained from three strains. Upon induction with mitomycin C two strains gave rise to phages with DNA restriction patterns identical to TP-J34 (data not shown). The third strain, S. thermophilus DSM20617, a strain from DSMZ collection which had been included in the screen, had originally been considered non-inducible (Sun, 2002).

We determined whole genome sequences for TP-J34 and TP-778L. In addition, left and right genome regions flanking prophage TP-778 were sequenced. For TP-EW, the two genome regions differing from those of TP-J34 (orf48 and the lysis gene) were sequenced in addition to the ltp gene. For TP-DSM20617, only the genomic region corresponding to the lysogeny module of TP-J34, bearing the ltp sequence, was amplified from genomic DNA by PCR and sequenced.

In this section, we will address features TP-J34 and TP-778L genomes that have in common, before we present in more detail those data, which are specific for the four phages and distinguish them from other S. thermophilus phages. TP-J34 and TP-778L DNAs share the same typical organization of functional modules characteristic for temperate S. thermophilus phages. Starting with the gene encoding the integrase, the order is: lysogeny module followed by modules for replication, DNA packaging, head morphogenesis, tail morphogenesis, lysis and finally lyogenic conversion (Figure 3A). While the lysogeny modules are transcribed from right to left, transcription of all other genes is from left to right. In none of the two genomes tRNA genes were detected. Sequences identical or highly similar to CRISPR spacer sequences in S. thermophilus strains were found in both genomes (Table 2). Their positions are indicated in Figure 3A. Orientations of the sequences are such that they correspond with the directions of transcription. Both phage genomes share with some other S. thermophilus phage genomes a site of a potential -1 translational frame-shift (Xu et al., 2004), which fuses orf41 with orf42 (TP-J34: bp 22942–23087) and orf38 with orf39 (TP-778L: bp 22560–22705), the two orfs in front of the gene encoding the tape measure protein (TMP). This frame-shift is known to result in formation of the tail assembly chaperone (Xu et al., 2013). TP-J34 has been shown to be a pac-type phage (Neve et al., 2003).

By the same experimental approach, namely showing that minor DNA restriction bands were not affected by heat treatment of digested DNA, TP-778L was shown to be a pac-type phage as well. This corresponds with the rather high similarity seen between both large terminase units (Figure 3A).

We compared the nucleotide sequence of TP-J34 with those of other S. thermophilus phages, for which complete genomes were available: O1205 (Stanley et al., 1997), Sf12 and Sf19 (Desiere et al., 1998), Sf11 (Lucchini et al., 1999), 7201 (Stanley et al., 2000), DT1 (Tremblay and Moineau, 1999), 2972 (Levesque et al., 2005), 858 (Deveau et al., 2008), ALQ13.2, Abc2 (Guglielmotti et al., 2009), and 5093 (Mills et al., 2011). The alignments by DotPlot analysis are shown in Figure 3B. It appears that virulent phage Sf11 and temperate phage TP-778 and O1205 are the most closely related to TP-J34. This is further reflected by the large number of putative gene products of these phages sharing highest homologies with those of TP-J34 (see Table 3).

TP-J34 DNA

The nucleotide sequence was determined for DNA isolated from purified phage particles obtained by mitomycin C treatment of lysogenic S. thermophilus J34, as described before (Neve et al., 1998, 2003). TP-J34 DNA consists of 45,606 bp, and thus it is the largest of the S. thermophilus phage DNAs sequenced so far (http://www.ncbi.nlm.nih.gov/genomes/GenomesGroup.cgi?opt= virus&taxid=10699). It has a G+C content of 38.8%, which is similar to the 39% of its host (Bolotin et al., 2004). The sequence is accessible under NC_020197. Numbering of the TP-J34
sequence starts with the last nucleotide of the stop codon of the \textit{int} gene.

Sixty \textit{orfs} were predicted by the Artemis programme (Rutherford et al., 2000), all of which were considered as protein-encoding genes (Table 3) with protein sizes varying between 46 (\textit{orf}9) and 1647 amino acids (\textit{orf}48). The predominant start codon appears to be AUG (57 out of 60); one UUG (\textit{orf}23), one AUU (\textit{orf}28), and one CUG (\textit{orf}55) were additionally predicted as start codons. AUU is a very unusual start codon (Blattner et al., 1997) normally coding for isoleucine. By repeated sequencing of PCR products generated with primers terS-F and terS-R using TP-J34 and TP-EW DNA, respectively, as templates, we excluded sequencing errors in this genomic region.

We have previously shown that upon induction of prophage TP-J34, mostly defective particles were released from the lysed host cells, and we have attributed the defect to a repeat region within \textit{orf}48 encoding the receptor binding protein (Neve et al., 2003). TP-J34L, an isolate forming clear plaques has been shown to have suffered a deletion of ca. 2.7 kb within the 4.4 kb HindIII fragment, thus reducing its size to 1.7 kb (Neve et al., 2003). In a Southern blot with HindIII-cleaved DNAs using a 1.0 kb PCR product (internal to the 1.7 kb HindIII fragment, obtained with primer pair D8/D12) of TP-J34L DNA as a probe, TP-J34 DNA extracted from lysates obtained by prophage induction yielded a major hybridization signal with the 4.4 kb fragment (Figures 1A,B). Two smaller signals at 3.5 and 2.6 kb were seen, indicating that the DNA was heterogeneous with respect to the 4.4 kb fragment, with 0.9 kb either one or two times deleted. As expected, TP-J34L DNA yielded a major signal at 1.7 kb. To confirm these results, the respective DNA regions of a TP-J34 lysate obtained by induction of the prophage and a TP-J34L lysate obtained by lytic propagation, were amplified by PCR, using primers D8+ and D12+ targeting sequences within the 4.4 kb HindIII fragment of TP-J34 but located outside of the repeat sequences. As expected, TP-J34L DNA gave rise to only one PCR product of ca. 1 kb. In case of the TP-J34 lysate, however, the DNA extracted yielded four products of ca. 1.0, 1.9, 2.8, and 3.7 kb (Figure 1C). This confirmed that TP-J34 DNA obtained by induction of the prophage was apparently heterogeneous with respect to the 4.4/1.7 kb HindIII fragment.
FIGURE 3 | (A) Alignment of gene maps and functional gene regions of TP-J34 and TP-778L. On the genetic maps, genes, and direction of transcription are indicated by arrows (very small genes are shown as boxes, the directions of transcription correspond to adjacent genes). Numbers or gene abbreviations refer to ORFs or genes as listed in Tables 3, 4. A scale indicating nucleotide positions is shown above the TP-J34 map. Approximate positions of functional regions (modules) are indicated by horizontal bars below the TP-778L map. Positions of CRISPR spacer sequences are indicated by dots above and below the maps of TP-J34 and TP-778L, respectively. (B) Dot plots of the TP-J34 nucleotide sequence compared to those of other S. thermophilus phages, including TP-778L. The horizontal line of each dot plot represents the 45,605 bp of TP-J34 DNA, whereas the vertical lines represent the numbers of bp for each phage, as indicated within each dot plot. Temperate (t) and virulent (v) phages are indicated.

Inspection of the TP-J34 genome sequence in this region revealed a 912 bp repeat structure within orf48 (Figure 4), located between genome positions 34,630 and 37,677. The triplicated sequence (3 × 912 bp) was found to be entirely in frame with the coding sequence of orf48 encoding the putative host specificity protein. Theoretically, a gene product should be produced, which—according to the defective morphology of TP-J34—should be either inactive in the tail assembly process or physically unstable. We like to point out that when the TP-J34 prophage was induced and the resulting lysate was inspected by transmission electron microscopy after fractionation in a CsCl gradient, no tail structures were detected anywhere in the gradient (Neve et al., 2003).

To genetically prove that the defect in orf48 was responsible for the tail assembly defect, we used the lysate obtained by induction of the TP-J34 prophage, which contained mostly defective particles, for re-lysogenization of prophage-cured S. thermophilus J34-6. From 11 lysogens obtained, chromosomal DNA was isolated, restricted with HindIII and subjected to Southern blotting using the 1.0 kb PCR product of TP-J34L DNA as probe. Of the 11 strains, seven showed a hybridization signal at 1.7 kb, three a signal at 2.6 kb and one a strong signal at 1.7 and a weaker signal at 2.6 kb. Genomic DNA isolated from lysogenic S. thermophilus J34 yielded three signals at 2.6, 3.5, and 4.4 kb (Figure 5). Of two of the re-lysogenized strains, J34-6-RL2 (signal at 2.6 kb) and J34-6-RL4 (signal at 1.7 kb), prophage were induced with mitomycin C. The lysates obtained were subjected to electron microscopy and compared with lysates obtained by prophage induction of S. thermophilus J34 and by lytic propagation of TP-J34L. The vast majority of phage particles of TP-J34 and TP-J34-6-RL2 were defective, whereas about half of the TP-J34L and TP-J34-6-RL4 looked morphologically intact, when analyzed in the electron microscope.
Table 2 | CRISPR spacer sequences present in genomes of TP-J34 and TP-778L.

| Sequence ID/ S. thermophilus strain phage | Spacer sequenceb | Identity | E-value |
|------------------------------------------|------------------|----------|---------|
| **TP-J34**                               |                  |          |         |
| NC_008532_5_4 /LMD-9                      | agagtacaatattgctctatggagacac 5882 5911 | 1        | 7e-07   |
| NC_008532_4_3 /LMD-9                      | catcatagggccagcttgaggtgtacac 44252 44281 | 1        | 7e-07   |
| NC_006449_1_31 /NC_006449_1_5 /CNRZ1066  | gtgggcaacctgaaacttgaaccgcggg 40182 40211 | 1        | 7e-07   |
| NC_017563_1_29 /ND03                     | gaaagattgcgtctatggagtttcaaa 5245 5274 | 0.97     | 1e-04   |
| NC_006449_1_6 /CNRZ1066                  | aaggttgaagcttatgcaacagaaaatc 33041 33070 | 0.97     | 1e-04   |
| NC_006449_1_41 /CNRZ1066                 | attggctaaatgcagccacaggttgaaa 38388 38416 | 0.97     | 4e-04   |
| **TP-778L**                              |                  |          |         |
| NC_017563_3_3 /ND03                      | cggcacagctataacacttatacagaga 12541 12571 | 1        | 2e-07   |
| NC_017927_3_4 /MN-ZLW-002                | attgaacctgattggtggctgattgaa 38358 38387 | 1        | 7e-07   |
| NC_008532_2_3 /LMD-9                     | agataagttgctgcttttacagatac 36793 36822 | 0.97     | 1e-04   |
| NC_006448_1_17 /LMG 18311                | cattaagctggtgaagcagcatgtgaag 4072 4101 | 0.97     | 1e-04   |
| NC_008532_2_16 /LMD-9                    | aaacgttcatctaattacagattcc 35046 35430 | 1        | 4e-04   |

a. Only sequences with E-values < 0.001 are shown.

b. The spacer sequences are shown with positions of first and last nucleotide.

microscope. When measuring plaque formation, phage lysates of TP-J34L and TP-J34-6-RL4 each yielded ca. 10^5 pfu/ml, while TP-J34 and TP-J34-6-RL2 each yielded ca. 10^5 pfu/ml. It thus appears that even an insertion of one 912 bp repeat is sufficient for inactivation of the tail assembly function of orf3 gene product.

**TP-778**

The nucleotide sequence was determined for DNA isolated from CsCl-purified TP-778L, lytically propagated on S. thermophilus B106, as described in Materials and Methods. TP-778L DNA consists of 41,757 bp. It has a G+C content of 39%, which is identical to the 39% of its host (Bolotin et al., 2004). The sequence is accessible under NC_022776. Numbering of the TP-J34 sequence starts with the last nucleotide of the stop codon of the int gene. Of the 52 orfs predicted by the Artemis programme (Rutherford et al., 2000), all were considered as protein-encoding genes (Table 4) with protein sizes varying between 46 (orf9) and 2020 amino acids (orf42). The predominant start codon appears to be AUG (49 out of 52). Of the residual three, two appear to be AUG (orf16 and 19) and one UUG (orf43).

S. thermophilus SK778 could not be cured of its prophage. To find a host for lytic propagation, a set of 16 non-lysogenic S. thermophilus wild-type strains were tested for sensitivity to TP-778L. Only S. thermophilus strain B106, a host strain for propagation of temperate phage 7201 (Proux et al., 2002), all were considered as protein-encoding genes (orf42). The predominant start codon appears to be AUG (49 out of 52). Of the residual three, two appear to be AUG (orf16 and 19) and one UUG (orf43).

**TP-EW**

From an industrial yoghurt, we isolated lysogenic S. thermophilus strain EW carrying a prophage (called TP-EW). Upon induction with mitomycin C, a phage lysate of morphologically intact phage particles was obtained. Using a spot assay, TP-EW was shown to be able to productively infect S. thermophilus J34-6 (not shown). Restriction analysis with HindIII of DNA isolated from CsCl-purified phage particles revealed a pattern basically identical to TP-J34 DNA. Therefore, we consider this phage to be almost identical to TP-J34. However, two differences in the restriction pattern with respect to TP-J34 DNA were noticed (Figure 1A): the two fragments of TP-J34 of 5.0 and 4.4 kb were missing, instead, two new fragments of 1.7 and 6.0 kb were detected.

By Southern hybridization (Figure 1B) and DNA sequencing we could show that TP-EW DNA did not contain the 3 × 912 bp repeats found in the 4.4 kb fragment of TP-J34 DNA, but that it instead contained the fragment of 1.7 kb identical to the one of TP-J34L (Figure 4).

The second differing restriction fragment of ca. 6 kb, when analyzed by additional restriction hydrolyses (not shown), appeared to be altered within the region of the lysin gene (orf54) with respect to TP-J34. A PCR with primers LYSup and LYSdown (Table 1) showed that TP-J34 DNA yielded a product of ca. 1.0 kb, while that of TP-EW DNA was ca. 1 kb larger (not shown). DNA sequencing and comparison with the TP-J34 DNA sequence indicated that the lysin gene of TP-EW contained an insertion of 1016 bp. BlastX analysis of the inserted sequence revealed an open reading frame encoding a protein of 205 amino acids with high homology to homing endonucleases (Lambowitz, 1993), indicating that the inserted sequence
Table 3 | Features of phage TP-J34 orfs and putative functions of their products.

| ORF (gene) | DNA frame | Start | Stop | Size (aa) | SD sequence AAGGAGGT<sup>a</sup> | Predicted function/best match BLASTp result | E-value | Match identity (%) | References, acc. no. |
|------------|-----------|-------|------|----------|-----------------------------------|-----------------------------------------------|---------|---------------------|---------------------|
| 1 (int)    | −1        | 1080  | 1    | 359      | TTGGGGGAttaaataaATG               | Integrase/ S. thermophilus phage Sfi21, Integrase/ 359 | 0.0     | 100                 | Desiere et al., 1998 |
| 2 (ltp)    | −3        | 1612  | 1184 | 142      | ATGGAGGAaattttATG                | Superinfection exclusion lipoprotein/ Streptococcus parasanguinis, prophage superinfection immunity protein 152 | 3e-42   | 51                 | WP_003010598         |
| 3          | −1        | 2061  | 1693 | 122      | AAAGTGAGaattttATG                | Putative metallo-proteinase/ S. thermophilus phage Sfi21, similar to cl-like repressor, metallo-proteinase motif | 1e-53   | 82                 | Desiere et al., 1998 |
| 4 (crh)    | −1        | 2433  | 2068 | 121      | AAGGAGAAsgatATG                  | Putative CI-repressor/ S. thermophilus phage Sfi21, CI-like repressor | 8e-21   | 55                 | Desiere et al., 1998 |
| 5 (cro)    | +1        | 2602  | 2805 | 67       | GAGGAAGacaaaATG                  | Putative Cro protein/ S. thermophilus phage 7201, Orf1, cro-like protein homolog | 4e-28   | 91                 | /                   |
| 6 (antl)   | +2        | 2858  | 3574 | 238      | AAGGATAATacATG                   | Putative antirepressor/ S. thermophilus phage Abc2, antirepressor protein | 2e-129  | 98                 | /                   |
| 7          | +1        | 3595  | 3876 | 93       | ATAGGGGTgaagagactATG             | / S. thermophilus phage Sfi21, Orf180         | 5e-47   | 98                 | Desiere et al., 1998 |
| 8          | +3        | 3936  | 4199 | 87       | AAGGAATTaaaATG                   | / S. thermophilus phage Sfi21, Orf187         | 3e-44   | 100                | /                   |
| 9          | +2        | 4217  | 4357 | 46       | GAGGAGAacaaaATG                  | / S. pyogenes phage315.5, hypothetical protein SpyM3_1347 | 4.4     | 41                 | Beres et al., 2002 |
| 10         | +1        | 4630  | 5517 | 295      | GGGTGAGTctaaaATG                 | / S. thermophilus phage 5093, putative primosome component | 1e-142  | 99                 | NC_012753           |
| 11         | +2        | 5529  | 6311 | 260      | AAAGGGGTgactATG                  | / S. thermophilus phage 5093, DnaC-like protein | 5e-136  | 93                 | NC_012753           |
| 12         | +3        | 6308  | 6490 | 60       | CAAGAGGAtgatgctATG               | / S. thermophilus phage 5093, hypothetical protein | 6e-27   | 100                | NC_012753           |
| 13         | +2        | 6615  | 7277 | 220      | AAGGAGAataaaATG                  | / S. thermophilus phage 5093, putative Erf protein | 1e-122  | 98                 | NC_012753           |

(Continued)
Table 3 | Continued

| ORF (gene) | DNA frame | Start | Stop | Size (aa) | SD sequence AAGGAGGT^a | Predicted function/best match BLASTp result | E-value | Match identity (%) | References, acc. no. |
|-----------|-----------|-------|------|-----------|-------------------------|-----------------------------------------------|---------|-------------------|---------------------|
| 14        | +1        | 7279  | 8238 | 319       | AAGGAGAActgcATG          | / S. thermophilus phage Abc2, hypothetical protein | 7e-146  | 82                | Guglielmotti et al., 2009 |
| 15        | +1        | 8261  | 8710 | 149       | CAGGAGAAaaaaacATG        | / S. thermophilus phage Abc2, single-stranded DNA binding protein | 1e-73  | 90                | Guglielmotti et al., 2009 |
| 16        | +1        | 8719  | 9180 | 153       | AAGGAAActATG             | / S. thermophilus phage Abc2, hypothetical protein | 1e-82  | 97                | Guglielmotti et al., 2009 |
| 17        | +3        | 9177  | 9413 | 78        | AAGGAGCTggaATG           | / S. thermophilus temperate phage O1205, hypothetical protein | 3e-31  | 83                | Stanley et al., 1997 |
| 18        | +2        | 9404  | 9574 | 56        | ATGGAGGActATG            | / S. thermophilus phage Abc2, hypothetical protein | 6e-19  | 85                | Guglielmotti et al., 2009 |
| 19        | +1        | 9571  | 9726 | 51        | AAGGAGATgattgaatATG      | / S. thermophilus phage Sfi21, hypothetical protein | 2e-17  | 87                | Desiere et al., 1998 |
| 20        | +3        | 9822  | 10028| 68        | AAGAGGTTaatataATG        | / Streptococcus pneumoniae, hypothetical protein | 6e-11  | 62                | ZP_01829218          |
| 21        | +3        | 10029 | 10670| 213       | AAGAGGTTgaatagATG        | / S. pyogenes phage 2096.1, phage protein | 8e-91  | 70                | Beres et al., 2002 |
| 22        | +1        | 10672 | 11217| 181       | TTGGAGAAataaaATG         | / S. thermophilus phage Sfi21, Orf178 | 5e-86  | 88                | Lucchini et al., 1999 |
| 23        | +3        | 11220 | 11732| 170       | AAGAGGTTgataaTTG         | / S. thermophilus phage 858, DNA binding protein (170aa) | 4e-80  | 86                | Deveau et al., 2008 |
| 24        | +1        | 11701 | 12018| 105       | AGGGAGATagtaaATG         | / S. thermophilus phage Sfi18, gp99 | 1e-43  | 94                | Lucchini et al., 1999 |
| 25        | +2        | 12020 | 12463| 147       | GTAGAGTaattagaATG        | / S. thermophilus phage Sfi11, hypothetical protein | 1e-64  | 99                | Lucchini et al., 1999 |
| 26        | +1        | 12469 | 13179| 236       | GGCTAGGAttcATG           | / S. thermophilus phage 858, Orf146 | 3e-117 | 86                | Deveau et al., 2008 |
| 27        | +1        | 13615 | 14028| 137       | AGAGAGGTagtacaATG        | / S. thermophilus phage Sfi11, gp137, ArpU phage transcriptional regulator | 5e-72  | 95                | Lucchini et al., 1999 |

(Continued)
### Table 3 | Continued

| ORF (gene) | DNA frame | Start | Stop | Size (aa) | SD sequence AAGGAGGT<sup>a</sup> | Predicted function/best match BLASTp result | E-value | Match identity (%) | References, acc. no. |
|------------|-----------|-------|------|-----------|----------------------------------|---------------------------------------------|---------|---------------------|----------------------|
| 28 (terS)  | +1        | 14177 | 14671| 165       | AAGGAGGTgattgATT                  | Putative terminase small subunit/ \( S. \) thermophilus phage Sfi11, gp172, putative terS product | 2e-111  | 98                  | Lucchini et al., 1999 |
| 29 (terL)  | +2        | 14658 | 15893| 411       | AAGGAGGTttaacaATG                 | Putative terminase large subunit/ \( S. \) thermophilus phage Sfi11, gp411, putative terL product | 0       | 98                  | Lucchini et al., 1999 |
| 30         | +1        | 15899 | 17407| 502       | TAGGAGGaatgATG                    | Putative portal protein/ \( S. \) thermophilus phage Sfi11, gp502, portal protein | 0       | 99                  | Lucchini et al., 1999 |
| 31         | +2        | 17404 | 18297| 297       | GAGGAGGTtatgATG                  | / \( S. \) thermophilus phage Sfi11, gp284, putative minor head protein; / | 5e-144  | 92                  | Lucchini et al., 1999 |
| 32         | +2        | 18486 | 19067| 193       | TAGGAAAtaaatgATG                 | / \( S. \) thermophilus phage Sfi11, gp193, putative scaffold protein | 2e-105  | 99                  | Lucchini et al., 1999 |
| 33         | +3        | 19087 | 19446| 119       | AAGGATTttaaATG                  | / \( S. \) thermophilus temperate phage O1205, Orf30, putative structural protein | 6e-57   | 94                  | Stanley et al., 1997 |
| 34         | +3        | 19465 | 20511| 348       | GAGGAGGAatattaaacATG             | Putative major head protein/ \( S. \) thermophilus phage Sfi11, gp348, major head protein | 0       | 97                  | Lucchini et al., 1999 |
| 35         | +2        | 20523 | 20684| 53        | GAGGTGCTactATG                  | / \( S. \) thermophilus phage Sfi11, gp53 | 3e-22   | 100                 | Lucchini et al., 1999 |
| 36         | +3        | 20696 | 21037| 113       | AGCGAGGTtgcgcATG                 | / \( S. \) thermophilus temperate phage O1205, hypothetisches Protein | 4e-57   | 96                  | Stanley et al., 1997 |
| 37         | +2        | 21034 | 21348| 104       | GGTGAGGTgcctttctATG              | / \( S. \) thermophilus phage Sfi11, gp104 | 6e-54   | 100                 | Lucchini et al., 1999 |
| 38         | +2        | 21348 | 21692| 114       | AAGGTGTTagataATG                 | / \( S. \) thermophilus phage Sfi11, gp114 | 7e-60   | 100                 | Lucchini et al., 1999 |
| 39         | +1        | 21689 | 22075| 128       | TGGGATGAaacATG                   | / \( S. \) thermophilus phage Sfi11, gp128 | 3e-71   | 100                 | Lucchini et al., 1999 |
| 40         | +3        | 22088 | 22594| 168       | TAGGAGGaaaaaATG                  | Putative major tail protein/ \( S. \) thermophilus temperate phage O1205, Orf37, major tail protein | 7e-90   | 99                  | Stanley et al., 1997 |

(Continued)
| ORF (gene) | DNA frame | Start | Stop | Size (aa) | SD sequence AAGGAGGT<sup>a</sup> | Predicted function/best match BLASTp result | E-value | Match identity (%) | References, acc. no. |
|-----------|-----------|-------|------|----------|-------------------------------|-----------------------------------------------|---------|------------------|----------------------|
| 41        | +2        | 22669 | 23022| 117      | TAGGAGTAaacaacaATG            | / S. thermophilus phage Sfi11, gp117          | 2e-61   | 100              | Lucchini et al., 1999|
| 42        | +2        | 23085 | 23402| 105      | TACGAGGAaattactacgagactATG    | / S. thermophilus phage Sfi11, gp105          | 1e-51   | 100              | Lucchini et al., 1999|
| 43        | +2        | 23392 | 27945| 1517     | AGAGAGGGgcttgcagATG           | Putative tape measure protein/ S. thermophilus phage Sfi11, gp1510, putative minor tail protein | 0       | 95               | Lucchini et al., 1999|
| tmp       | +2        | 27945 | 29483| 512      | TGAGAGGtcattaATG              | / S. thermophilus phage Sfi11, gp512, putative minor tail protein | 0       | 94               | Lucchini et al., 1999|
| 45        | +1        | 29483 | 32485| 1000     | AAGGTGGATtaATG               | / S. thermophilus phage Sfi11, gp1000, putative minor tail protein (Lysozyme and Chap domain) | 0       | 97               | Lucchini et al., 1999|
| 46        | +1        | 32501 | 33622| 373      | TAGGAGGAAattaATG              | / S. thermophilus phage Sfi11, gp373          | 0       | 98               | Lucchini et al., 1999|
| 47        | +3        | 33622 | 33795| 57       | TGTAGGGAaatcaataATG           | / S. thermophilus phage Sfi11, gp57          | 7e-24   | 94               | Lucchini et al., 1999|
| 48        | +1        | 33773 | 38716| 1647     | GCAGAGGTagtgaATG              | Putative host specificity protein / S. thermophilus phage DT2, host specificity protein | 0       | 72               | Duplessis and Moineau, 2001|
| 49        | +2        | 38718 | 40727| 669      | TAGGAGAGattaacaATG            | / S. thermophilus phage Sfi11, gp669, putative minor structural protein | 0       | 96               | Lucchini et al., 1999|
| 50        | +1        | 40691 | 41092| 133      | AAAATGATG                    | / S. thermophilus phage Sfi11, gp149          | 1e-59   | 76               | Lucchini et al., 1999|
| 51        | +2        | 41112 | 41258| 48       | AAAGAGGAAaaagatATG            | / S. thermophilus phage Sfi21, hypothetical protein | 9e-12   | 75               | Desiere et al., 1998|
| 52        | +2        | 41276 | 41599| 107      | AGGGATGTagtATG               | / S. thermophilus phage DT1, Or123            | 3e-53   | 95               | Lamoth et al., 2005|
| hol       | +3        | 41605 | 41847| 80       | TGAGAGGAtaagacaATG            | Putative holin/ S. thermophilus temperate phage O1205, putative holin | 4e-35   | 93               | Stanley et al., 1997|

(Continued)
### Table 3 | Continued

| ORF (gene) | DNA frame | Start (bp) | Stop (bp) | Size (aa) | SD sequence AAGGAGGT<sup>a</sup> | Predicted function/best match BLASTp result | E-value | Match identity (%) | References, acc. no. |
|------------|-----------|------------|-----------|-----------|-------------------------------|---------------------------------------------|--------|-------------------|---------------------|
| 54 lys     | +1        | 41849      | 42694     | 281       | AAGGAGGAaaaatat ATG          | Putative lysin/ S. thermophilus phage S38, putative lysin | 7e-141 | 90                | Foley et al., 2000 |
| 55         | +2        | 42858      | 43256     | 132       | AAGAAAAAggctattgac CTG      | / S. pneumoniae, hypothetical protein (trans-membrane region) | 5e-12  | 42                | NZ_ABAA01000017     |
| 56         | +1        | 43557      | 43892     | 111       | AAGAGGAAatgaa ATG           | / S. thermophilus phage Sfi19, gp111          | 8e-55  | 100               | Desiere et al., 1998 |
| 57         | +1        | 43914      | 44465     | 183       | AAGAGGAataaaaa ATG          | / S. thermophilus phage Sfi11, gp183         | 5e-104 | 100               | Lucchini et al., 1999 |
| 58         | +2        | 44491      | 44742     | 83        | AACGAGGTgaaaca ATG          | / S. thermophilus phage Sfi11, gp83          | 6e-40  | 100               | Lucchini et al., 1999 |
| 59         | +1        | 44768      | 44947     | 59        | AAGCTTAAactgat ATG          | / S. thermophilus phage 5093, hypothetical protein | 1e-25  | 93                | NC_012753           |
| 60         | +1        | 45006      | 45428     | 140       | GAGAAGGTAatgaa ATG          | / S. thermophilus phage 5093, hypothetical protein | 4e-63  | 85                | NC_012753           |

<sup>a</sup>da Silva Oliveira et al., 2004.

---

**FIGURE 4** Comparison of the genetic structure of the TP-J34 DNA region containing the triple repeat sequences R1–R3 with that of TP-J34L and TP-EW, respectively. The bp numbers indicate the first bp of a repeat. “a” and “b” denote the regions with similarities to sequences within the repeats (marked as “a” and “b”). Sequences exclusively found within the three repeats are indicated as “int.” HindIII restriction sites flanking the 4.4 and 1.7 kb fragment of TP-J34 and TP-J34L/TP-EW, respectively are shown. Gene 48 start and stop are marked by solid triangles.

is a group I intron. Such introns have frequently been found in *S. thermophilus* phages to be located within the lysin gene (Foley et al., 2000). Comparison of the putative splice sites indicated high homology between *S. thermophilus* phages containing an intron in that position (Figure 7). Comparison of the DNA sequences flanking the insertion site of the intron with TP-J34 DNA sequence of that region revealed many deviations from TP-J34 sequence in the close vicinity, while the DNA sequences of TP-EW and TP-J34 were identical when they were more than a few hundred nucleotides apart from the insertion site.

Finally, for sequencing the *ltp*<sub>TP-EW</sub> gene, we amplified a DNA region comprising the *ltp* gene plus the flanking regions by means of primers targeting sequences of TP-J34 genes int and orf3, respectively. The ca. 900 bp of nucleotide sequence obtained were 100% identical to those of TP-J34.
S. thermophilus DSM20617 was obtained from the German type culture collection. It had been included in a screening for lysogenic S. thermophilus strains carrying ltp-expressing prophages (Sun, 2002). The DNA region of lysogenic strain S. thermophilus DSM20617 comprising orf1 (integrase) through orf6 (antirepressor) and defined by primers primer4 (left) and Mz12.R (right) was sequenced by primer walking. The sequence of ca. 3.7 kb was more than 99% identical to that of prophage TP-778 residing in S. thermophilus SK778. Only one base within orf1 (int), one base within orf2 (ltp), and two bases within orf5 (ant) turned out to be different. Restriction analyses of DNA isolated from the phage lysate obtained by induction of the prophage did not reveal any similarities to restriction patterns of DNA isolated from TP-34L and TP-778L, respectively (Figure S1A). Also, comparison of the HindIII and EcoRI patterns of TP-DSM20617 DNA with λ DNA. Other lanes (from left to right): J34-RL2; J34-6-RL2a; J34-6-RL2b; J34-6-RL2c; J34-6-RL2d; J34-6-RL2e; J34-6-RL2f; J34-6-RL4a; J34-6-RL4b; J34-6-RL4; J34-6-RL4c. The sizes of the λ DNA bands are indicated in the right margin.

FIGURE 5 | Southern blot with DIG-labeled 1 kb probe of HindIII-cleaved phage and chromosomal DNA of eleven S. thermophilus strains relysogenized with TP-J34. Lane a: TP-J34; lane b: TP-J34L; lane c: J34; lane d: J34-6 (no prophage, negative control); lane M: DIG-labeled, HindIII-cleaved λ DNA. Other lanes (from left to right): J34-RL2; J34-6-RL2a; J34-6-RL2b; J34-6-RL2c; J34-6-RL2d; J34-6-RL2e; J34-6-RL2f; J34-6-RL4a; J34-6-RL4b; J34-6-RL4c. The sizes of the λ DNA bands are indicated in the right margin.

TP-DSM20617

S. thermophilus DSM20617 was obtained from the German type culture collection. It had been included in a screening for lysogenic S. thermophilus strains carrying ltp-expressing prophages (Sun, 2002). The DNA region of lysogenic strain S. thermophilus DSM20617 comprising orf1 (integrase) through orf6 (antirepressor) and defined by primers primer4 (left) and Mz12.R (right) was sequenced by primer walking. The sequence of ca. 3.7 kb was more than 99% identical to that of prophage TP-778 residing in S. thermophilus SK778. Only one base within orf1 (int), one base within orf2 (ltp), and two bases within orf5 (ant) turned out to be different. Restriction analyses of DNA isolated from the phage lysate obtained by induction of the prophage did not reveal any similarities to restriction patterns of DNA isolated from TP-34L and TP-778L, respectively (Figure S1A). Also, comparison of the HindIII and EcoRI patterns of TP-DSM20617 DNA with λ DNA. Other lanes (from left to right): J34-RL2; J34-6-RL2a; J34-6-RL2b; J34-6-RL2c; J34-6-RL2d; J34-6-RL2e; J34-6-RL2f; J34-6-RL4a; J34-6-RL4b; J34-6-RL4c. The sizes of the λ DNA bands are indicated in the right margin.

FIGURE 5 | Southern blot with DIG-labeled 1 kb probe of HindIII-cleaved phage and chromosomal DNA of eleven S. thermophilus strains relysogenized with TP-J34. Lane a: TP-J34; lane b: TP-J34L; lane c: J34; lane d: J34-6 (no prophage, negative control); lane M: DIG-labeled, HindIII-cleaved λ DNA. Other lanes (from left to right): J34-RL2; J34-6-RL2a; J34-6-RL2b; J34-6-RL2c; J34-6-RL2d; J34-6-RL2e; J34-6-RL2f; J34-6-RL4a; J34-6-RL4b; J34-6-RL4c. The sizes of the λ DNA bands are indicated in the right margin.

DISCUSSION

Our screening for Ltp-expressing prophages in S. thermophilus yielded just four different phages, three of which (TP-J34, TP-EW, TP-778) can be assigned to the Sfi11 sub-species species of S. thermophilus phages (Proux et al., 2002; Quiberoni et al., 2010), since they are pac-type phages and their genome sequences show high similarities to phages Sfi11 and O1205. The fourth phage, TP-DSM20617 cannot be classified due to lack of information on its genome. The three phages, TP-J34 and TP-EW on one hand and TP-778 on the other, appear to represent two different lines within the Sfi11 sub-species, with the major difference between the two types being lack of homology between the genes within the “replication” module. Other minor differences are seen within the modules of “DNA-packaging,” “tail morphogenesis,” and “lysogenic conversion.” The exchange of entire functional modules appears to be the general mechanism of recombination between bacteriophages (Lucchini et al., 1998). Such exchange is easily accomplished without impairing functionality of the phage, especially when interaction with proteins of other modules does not occur. This is the case with the proteins of the “replication” as well as the “lysogenic conversion” module. The “DNA packaging” module consists of two proteins only,
**Table 4 | Features of phage TP-778L orfs and putative functions of their products.**

| Orf (gene) | DNA-frame | Start | Stop | Size (aa) | SD sequence AAGGAGGA | Predicted function/best match BLASTp result | E-value | Match identity (%) | References, acc. no. |
|------------|-----------|-------|------|-----------|----------------------|---------------------------------------------|---------|---------------------|---------------------|
| 1 (int)    | −3        | 420   | 1    | 139       | TTGGGGGAattaaaATG    | Putative integrase/ *Streptococcus thermophilus* phage Sfi21, integrase / 359 | 2e-90   | 99                  | Desiere et al., 1998, NP_049990 |
| 2 (ltp)    | −1        | 952   | 524  | 142       | TGGTAGGAatatATG      | Putative superinfection exclusion lipoprotein/ *Streptococcus thermophilus* phage TP-J34/ 142 | 3e-79   | 93                  | Neve et al., 1998, AAC03455 |
| 3          | −2        | 1400  | 1032 | 122       | AAGGAAAtgtgaaATG     | Putative metallo-proteinase motif/ *Streptococcus thermophilus* phage Sfi21, cl- like repressor/ 122 | 4e-81   | 96                  | Desiere et al., 1998, NP_049992 |
| 4 (crh)    | −2        | 1772  | 1407 | 121       | AAGGAAAtagATG        | Putative CI-repressor/ *Streptococcus thermophilus* phage TP-J34, putative cl-repressor homolog/ 121 | 1e-80   | 100                 | Neve et al., 1998, AAC03457 |
| 5 (cro)    | +3        | 1941  | 2144 | 67        | GAGGAAAcAAAATG       | Putative Cro protein/ *Streptococcus thermophilus* phage TP-J34, Cro-like regulatory protein/ 67 | 3e-41   | 99                  | Neve et al., 1998, AAC03458 |
| 6 (ant)    | +1        | 2197  | 2913 | 238       | AGAAAGGAtatacATG     | Putative antirepressor/ *Streptococcus thermophilus* phage TP-J34, P1-antirepressor homolog / 238 | 2e-175  | 99                  | Neve et al., 1998, AAC03459 |
| 7          | +3        | 2934  | 3215 | 93        | ATAGGGGTagaaaaagactATG | -/ *Streptococcus thermophilus* phage TP-J34, hypothetical protein/ 93 | 1e-61   | 100                 | Neve et al., 1998, AAC03460 |
| 8          | +2        | 3275  | 3538 | 87        | AAGGAATTaaaATG       | -/ *Streptococcus thermophilus* phage Sfi21 Orf87, hypothetical protein Sfi21p33/ 87 | 6e-57   | 100                 | Desiere et al., 1998, NP_597801 |
| 9          | +1        | 3556  | 3693 | 46        | AAGGAAGGagaaacaaaATG | -/ *Streptococcus thermophilus* phage TP-J34, hypothetical protein/ 46 | 4e-26   | 100                 | This study                                                                 |
| 10         | +2        | 3932  | 4405 | 157       | AAGGAAGTaattaaaaATG  | -/ *Streptococcus thermophilus* phage ALQ13.2, hypothetical protein/ 157 | 4e-88   | 84                  | Guglielmotti et al., 2009, YP_003344879 |

(Continued)
| Orf (gene) | DNA-frame | Start | Stop | Size (aa) | SD sequence AAGGAGGT<sup>a</sup> | Predicted function/best match BLASTp result | E-value | Match identity (%) | References, acc. no. |
|-----------|----------|-------|------|----------|-------------------------------|---------------------------------------------|---------|--------------------|----------------------|
| 11        | +1       | 4402  | 5103 | 233      | AAGGAGAAacctaacataagATG       | / Streptococcus thermophilus phage, putative replication initiation protein/ 233 | 3e-168  | 99                 | Levesque et al., 2005, YP_238517 |
| 12        | +2       | 5060  | 6472 | 470      | AAAAGGTTagaagtagATG           | / Streptococcus thermophilus phage 858 Orf 37, putative helicase/ 470 | 0.0     | 99                 | Deveau et al., 2008, YP_001686831 |
| 13        | +2       | 6479  | 6952 | 157      | TGGGATaaaacacATG              | / Streptococcus thermophilus phage 858 Orf 38/ 157 | 3e-108  | 97                 | Deveau et al., 2008, YP_001686832 |
| 14        | +3       | 6957  | 7772 | 271      | TTTGCATcctaagactATG           | / Streptococcus thermophilus phage 858 Orf 39, primase-polymerase domain/ 271 | 0.0     | 99                 | Deveau et al., 2008, YP_001686833 |
| 15        | +1       | 741   | 9297 | 518      | AAGGAGTtagataaaacATG          | Putative primase/ Streptococcus phage YMC-2011, putative primase/ 519 | 0.0     | 92                 | Geng et al., 2011, YP_006561246 |
| 16        | +1       | 9541  | 9861 | 106      | AGAAAGGTTaatttaaiGTG          | / Streptococcus thermophilus phage 858 Orf 41, VRR_NUC domain/ 106 | 1e-65   | 92                 | Deveau et al., 2008, YP_001686835 |
| 17        | +2       | 9845  | 10081| 78       | AAGGAAGCttggatatagtaaATG      | / Streptococcus thermophilus phage Abc2, hypothetical protein/ 78 | 7e-42   | 87                 | Guglielmotti et al., 2009, YP_003347446 |
| 18        | +3       | 10098 | 10253| 51       | AAGATGTTagagttATG             | / Streptococcus thermophilus phage Sfi19 Orf 51; hypothetical protein Sfi19p40/ 51 | 1e-23   | 84                 | Desiere et al., 1998, NP_049960 |
| 19        | +3       | 10254 | 10835| 193      | GAGGGTGAtaaATG                 | / Streptococcus thermophilus phage Abc2, hypothetical protein/ 166 | 9e-58   | 69                 | Guglielmotti et al., 2009, YP_003347451 |
| 20        | +3       | 10836 | 11348| 170      | GAAGGAGGTgaataaATG            | Putative DNA-binding protein/ Streptococcus thermophilus phage 5093, DNA binding protein, HTH_XRE/ 170 | 6e-111  | 91                 | Mills et al., 2011, YP_002925093 |

(Continued)
### Table 4 | Continued

| Orf (gene) | DNA-frame | Start (aa) | Stop (aa) | Size (aa) | SD sequence | Predicted function/best match BLASTp result | E-value | Match identity (%) | References, acc. no. |
|------------|-----------|------------|-----------|-----------|-------------|-----------------------------------------------|--------|---------------------|----------------------|
| 21         | +1        | 11317      | 11634     | 105       | AGGGAGAGtagaATG | -/Streptococcus thermophilus phage TP-J34, hypothetical protein/105 | 1e-65  | 95                  | This study            |
| 22         | +2        | 11636      | 12079     | 147       | GTAGGAGGattaAgATG | -/Streptococcus thermophilus phage TP-J34, hypothetical /147 | 2e-103 | 99                  | This study            |
| 23         | +1        | 12085      | 12795     | 236       | GTGGGGGCtaggattcATG | /Streptococcus thermophilus phage 7201 Orf 18/235 | 7e-161 | 94                  | Stanley et al., 2000, NP_038319 |
| 24         | +2        | 13232      | 13645     | 137       | AGAGAGGCagaaaaATG | Putative transcriptional regulator/Streptococcus thermophilus phage TP-J34 Orf17, transcriptional regulator ArpU family/137 | 2e-94  | 99                  | This study            |
| 25         | +2        | 13766      | 14278     | 170       | TTTGAGGtgtttaggattgaaATG | Putative terminase small subunit/Streptococcus thermophilus phage 2972, terminase small subunit/150 | 7e-85  | 86                  | Levesque et al., 2005, YP_001686797 |
| 26         | +3        | 14265      | 15500     | 411       | AAGGAGCTgtagcATG | Putative terminase large subunit/Streptococcus thermophilus phage TP-J34 Orf19, putative terminase large subunit /411 | 0.0    | 97                  | This study            |
| 27         | +2        | 15506      | 17014     | 502       | TAGGAGAtgATG | Putative portal protein/Streptococcus thermophilus phage 858 orf6, putative portal protein/502 | 0.0    | 97                  | Deveau et al., 2008, YP_001686800 |
| 28         | +1        | 17011      | 17904     | 297       | GAGGAGGTtagaATG | Putative head protein/Streptococcus thermophilus phage 2972, head protein/297 | 0.0    | 96                  | Levesque et al., 2005, YP_238489 |
| 29         | +3        | 18096      | 18677     | 193       | TAGGAGAcagaaATG | Putative scaffold protein/Streptococcus thermophilus phage 2972, scaffold protein/193 | 7e-130 | 96                  | Levesque et al., 2005, YP_238490 |
| 30         | +1        | 18697      | 19056     | 119       | AAGGAAATtagaATG | Putative head protein/Streptococcus thermophilus phage 2972, head protein/119 | 6e-75  | 97                  | Levesque et al., 2005, YP_238491 |

(Continued)
Table 4 | Continued

| Orf (gene) | DNA-frame | Start (aa) | Stop (aa) | Size (aa) | SD sequence AAGGAGG<sup>a</sup> | Predicted function/best match BLASTp result | E-value | Match identity (%) | References, acc. no. |
|------------|-----------|------------|-----------|-----------|-------------------------------|----------------------------------------|---------|-------------------|----------------------|
| 31         | +1        | 19075      | 20121     | 348       | GAGGAGGAacattaaaacATG         | Putative capsid protein/ *Streptococcus thermophilus* phage ALQ13, capsid/348 | 0.0     | 98                | Guglielmotti et al., 2009, YP_003344853 |
| 32         | +3        | 20133      | 20294     | 53        | TAAGGAGTactgatATG             | -/ *Streptococcus thermophilus* phage ALQ13, hypothetical protein/53 | 5e-19   | 98                | Guglielmotti et al., 2009, YP_003344854 |
| 33         | +2        | 20309      | 20647     | 112       | AGTGAGGTagtgcgtgATG           | -/ *Streptococcus thermophilus* phage 01205 Orf 33, hypothetical protein/122 | 3e-70   | 94                | Stanley et al., 1997, NP_695111 |
| 34         | +1        | 20644      | 20968     | 104       | GTGAGGTctattttctATG           | -/ *Streptococcus thermophilus* phage 2972, hypothetical protein/104 | 7e-62   | 94                | Levesque et al., 2005, YP_238495 |
| 35         | +2        | 20960      | 21304     | 114       | AAGGTGTaTgaataacATG           | -/ *Streptococcus thermophilus* phage 5f11 Orf 114, hypothetical protein/114 | 4e-72   | 94                | Lucchini et al., 1999, NP_056684 |
| 36         | +1        | 21289      | 21687     | 132       | GAAGAGATggcgaaATG             | -/ *Streptococcus thermophilus* phage ALQ13, hypothetical protein/128 | 1e-84   | 95                | Guglielmotti et al., 2009, YP_003344858 |
| 37         | +2        | 21701      | 22210     | 169       | AATAGGAGGAaaaATG              | Putative tail protein/ *Streptococcus thermophilus* phage 2972, tail protein/169 | 9e-115  | 98                | Levesque et al., 2005, YP_238498 |
| 38         | +3        | 22287      | 22640     | 117       | TAGAGGTaaacaacaaATG           | -/ *Streptococcus thermophilus* phage 2972, hypothetical protein/117 | 6e-78   | 99                | Levesque et al., 2005, YP_238499 |
| 39         | +2        | 22703      | 23020     | 105       | GAGGAGTTaatcactaatgccATG      | -/ *Streptococcus thermophilus* phage 2972, hypothetical protein/105 | 2e-65   | 99                | Levesque et al., 2005, YP_238500 |
| 40         | +3        | 23010      | 27563     | 1517      | AGAGAGGGtgtgctagATG           | Putative tape measure protein/ *Streptococcus thermophilus* phage O1205, putative tail protein/1517 | 0.0     | 90                | Stanley et al., 2000 NP_695118 |

(Continued)
| Orf (gene) | DNA-frame | Start (aa) | Stop (aa) | Size (aa) | SD sequence AAGGAGGT<sup>a</sup> | Predicted function/best match BLASTp result | E-value | Match identity (%) | References, acc. no. |
|-----------|-----------|------------|-----------|-----------|---------------------------------|---------------------------------------------|---------|-------------------|---------------------|
| 41        | +2        | 27563      | 29101     | 512       | TCCGGAGTccaaattt<sup>b</sup>ATG | Putative tail protein/ *Streptococcus thermophilus* phage Sfi11, putative minor structural protein/ 511 | 0.0     | 89                | Lucchini et al., 1999, NP_0566980 |
| 42        | +1        | 29101      | 35163     | 2020      | AAGGGTTGAttt<sup>b</sup>ATG | Putative host specificity protein/ *Streptococcus thermophilus* phage 858 Orf21, prophage tail protein/ 1006 | 0.0     | 88                | Deveau et al., 2008, YP_001686815 |
| 43        | +1        | 35164      | 37185     | 673       | TAGGGTTtttt<sup>b</sup>TTG | Putative tail protein/ *Streptococcus thermophilus* phage 858 Orf22/ 673 | 0.0     | 89                | Deveau et al., 2008, YP_001686816 |
| 44        | +1        | 37201      | 37548     | 115       | AAGAGAGAAattt<sup>b</sup>ATG | -/ *Streptococcus thermophilus* phage TP-J34, hypothetical protein/ 133 | 6e-73   | 96                | This study          |
| 45        | +2        | 37568      | 37714     | 48        | AAGAGAGAAaaagg<sup>b</sup>ATG | -/ *Streptococcus thermophilus* phage TP-J34, hypothetical protein/ 48 | 2e-24   | 100               | This study          |
| 46        | +1        | 37732      | 38055     | 107       | TAGGGTTgtgtg<sup>b</sup>ATG | -/ *Streptococcus thermophilus* phage TP-J34, hypothetical protein/ 107 | 2e-71   | 100               | This study          |
| 47        | +3        | 38064      | 38306     | 80        | TAGAGGAtaaataacaa<sup>b</sup>ATG | Putative holin/ *Streptococcus thermophilus* phage Abc2, holin/ 80 | 7e-44   | 93                | Guglielmotti et al., 2009, YP_003347430 |
| 48        | +1        | 38308      | 39153     | 281       | AAGAGAGAAaaatag<sup>b</sup>ATG | Putative lyisin/ *Streptococcus thermophilus* phage TP-J34, putative lyisin/ 281 | 8e-181  | 90                | This study          |
| 49        | +1        | 39499      | 39720     | 73        | AAGATTGAAacaaactagacga<sup>b</sup>ATG | -/ *Streptococcus thermophilus* phage TP-J34/ 73 | 5e-46   | 100               | Neve et al., 1998, AAC03448 |
| 50        | +3        | 40419      | 40715     | 98        | AGAGGTTaaaagaa<sup>b</sup>ATG | -/ *Streptococcus* sp. F0441, hypothetical protein/ 101 | 2e-38   | 66                | WP_009730541        |

(Continued)
Table 4 | Continued

| Orf (gene) | DNA-frame | Start | Stop | Size (aa) | SD sequence AAGGAGG^a | Predicted function/best match BLASTp result | E-value | Match identity (%) | References, acc. no. |
|------------|-----------|-------|------|----------|------------------------|--------------------------------------------|--------|-------------------|---------------------|
| 51         | +2        | 40778 | 41200| 140      | AAGGAAGGTatATG          | -/ Streptococcus thermophilus phage Sfi21, hypothetical protein / 140 | 5e-85  | 90                | Desiere et al., 1998, NP_049988 |
| 52         | +3        | 41202 | 41624| 140      | AAAGATGTaatctaaATG      | -/ Streptococcus thermophilus phage Sfi21, hypothetical protein / 140 | 1e-81  | 87                | Desiere et al., 1998, NP_049989 |

*a da Silva Oliveira et al., 2004.

Figure 6 | Mechanism of excision of TP-778 prophage from its host’s genome to yield phage TP-778L. Prophage and host DNA are shown by black and green line, respectively. Genes are indicated by arrows. Binding sites of primers 4 and 3 (Bruttin et al., 1997 are shown). The region of predicted cross-over is indicated by a cross.

the small (TerS) and the large terminase (TerL) units. The portal protein, encoded by the gene immediately following that of the large terminase, may be considered part of this module, however it also plays a critical role in head assembly (Padilla-Sanchez et al., 2013). The lack of similarity within the “DNA packaging” module only affects the N-terminal and central regions of TerS, which are involved in DNA binding and oligomerization, respectively (Sun et al., 2012). The C-terminal part, which is involved in interaction with the portal protein, is absolutely identical between TP-J34 and TP-778L. Thus, functionality defined as productive interaction with other components of the module is apparently not impaired by the alterations affecting TerS. The fact that both phages are pac-type phages and show high genome similarities to phages Sfi11 and O1205 confirms this finding. The last region of divergence between TP-J34 and TP-778L DNA concerns the “tail morphogenesis” module. Compared to the TP-J34 module, orfs 45 and 48 appear to be fused to form the one large orf42 of TP-778L. The gene product of orf45 is characterized by a Lyz2 (Nambu et al., 1999) and a CHAP-domain (Bateman and Rawlings, 2003), indicating involvement in peptidoglycan hydrolysis during infection following adsorption. The gene product of orf48 appears to be the receptor binding protein, containing a domain which is found in galactose-binding proteins (Gaskell et al., 1995). These three domains are found in the orf42 gene product of TP-778L. It appears that both functions, which are required at the first steps of infection in TP-778, are combined in just one protein. This is not too surprising, since proteins encoded by genes with adjacent positions on the genetic map may also be in close contact within the structures formed. A fact that has been the basis for successful “block cloning” applied for elucidation of tail sub-structures (Campanacci et al., 2010).

The orf48 gene product, containing the three 912 bp repeats, appears to be either physically unstable or inactive in the tail
assembly process. The few intact phage particles found after induction may arise from recombinational loss of the repeats occurring during replication: the few functional copies of Orf48 produced may initiate successful tail assembly. If TP-J34 DNA lacking the 912 bp repeat is packaged into such phage particles, TP-J34L phage particles are produced. The observed very low efficiency of plating for phage lysates resulting from induction of the prophage (Neve et al., 2003), even if they contained just one repeat may be due to phenotypic mixing (Streisinger, 1956), i.e., packaging of DNA into phage particles which are not derived from that DNA.

The 912 bp repeat shows DNA sequence homology to its flanking regions. However, an internal region of ca. 450 bp of the 912 bp repeat does not show homology to the flanking DNA or to other regions of TP-J34 DNA, which may indicate that this DNA region had been introduced by horizontal gene transfer. BlastN analysis revealed 80% sequence identity over the 450 bp to the host specificity gene of S. thermophilus bacteriophage DT2 (Duplessis and Moineau, 2001), and BlastX revealed 75% sequence similarity (E-value 2e-60) over 150 amino acids of the product of that gene. One may speculate that the DNA region has been obtained by horizontal gene transfer from a not yet identified phage with homology to phage DT2 in this genome region.

Horizontal gene transfer is apparently also responsible for the distribution of ltp genes, encoding a sie lipoprotein, among strains and bacteriophages of Gram-positive bacteria (Sun et al., 2006). The members of this family of “host cell surface-exposed lipoproteins” (Marchler-Bauer et al., 2011) are found scattered within annotated genomes of bacteriophage and bacteria (Sun et al., 2006). This would argue for ltp to be a member of the so called “morons,” genes inserted into prophage genomes by horizontal gene transfer which provide some benefit to the host (Cumby et al., 2012). Further additional evidence for the “moron” character of ltp like presence of promoter and terminator will be presented elsewhere (Koberg et al., in preparation). The fact that the few temperate S. thermophilus phage harboring ltp are all very closely related indicates that horizontal transfer of an ltp gene into S. thermophilus phage occurred just once. The genome deviations seen among the three phages TP-J34, TP-778, and TP-EW should therefore have occurred after ltp had been acquired.

The differences in amino acid sequences and activities seen between plasmid-expressed LtpTP-J34 and LtpTP-778 confirm our recent data on LtpTP-J34 structure (Bebeacua et al., 2013), which indicated that the repeat domains are those responsible for super infection exclusion by interaction with the TMP of the super infecting phage and that the negatively charged amino acids in this region are important for interacting with the positively charged C-terminal end region of the P008 TMP. The deviations from LtpTP-J34 seen in the amino acid sequences of the LtpTP-778 repeat domain are mostly conservative. It is intriguing that with one exception the charges are not changed by the deviations. At this point it would just be speculation that the one change from negatively charged Glu to neutral Gly (see Figure 8) would be responsible for the functional differences. Another candidate for this difference could be the amino acid change from His to Pro (see Figure 8). However, this exchange does not affect a helix but just a β-turn within the first repeat domain.

When discussing the potential effects on interaction with TMP of the amino acid exchanges seen between LtpTP-J34 and LtpTP-778, one should bear in mind that no genome sequence is available for lactococcal phage P001, a member of the c2-species. In the available genome sequence of lactococcal phage c2, however, no TMP is annotated (Lubbers et al., 1995). This is apparently due to the fact that phage c2 uses the host “phage infection protein” Pip for adsorption and DNA-injection (Monteville et al., 1994). In phage c2, gene 110 encoding the “tail adsorption protein” should be the TMP of phage c2. This protein would not

FIGURE 7 | Alignment of DNA sequences of S. thermophilus phages TP-EW, S3b (Acc. No. AF148561.1), ST3 (Acc. No. AF148565.1), J1 (Acc. No. AF148566.1), and DT1 (Acc. No. NC_002072) in the region surrounding the group I-intron, present in all phage DNAs. The splice site is indicated by the vertical arrow. Sequence differences are indicated. Two 6-bp inverted repeats are indicated by horizontal arrows above the DNA sequence. The numbers flanking the TP-EW sequence correspond to the nt positions within the ltp gene of this phage.

FIGURE 8 | Alignment of amino acid sequences of different Ltp proteins. The cleavage site between signal sequence and mature protein is indicated. The first Cys of the mature TP-J34 lipoprotein is marked as +1. The two repeat regions are underlined. Amino acids identical to those of TP-J34 are indicated by “-.”
A., Foley, S., and Brüssow, H. (1997). The site-specific integration system of c2. At this stage, we can just notice that the C-terminal end of TMP is apparently underlines the peculiar situation of c2-phages with respect to TMP. With UsLtpTP-J34 at hand, we may be able to test whether the “tail adsorption protein” is in fact the TMP of c2. At this stage, we can just notice that the C-terminal end of the c2 “tail adsorption protein” is positively charged, which is in agreement with the proposed binding site of LtpTP-J34 in TMP of P008 (Bebeacu et al., 2013).

To conclude, in this communication we could show that amino acid deviations seen between LtpTP-J34 and LtpTP-778 are apparently responsible for differences seen in the biological activities of both proteins. These deviations provide some clues on how to further study interaction between Ltp and TMP in more detail. Our data also show that phages TP-34, TP-778, and TP-EW belong to the Sfi11 sub-species of S. thermophilus phages. The close relatedness of the three phages argues for acquisition of ltp prior to formation of the three phages from a common ancestor.

ACKNOWLEDGMENTS
We gratefully acknowledge technical assistance by I. Lammertz.

SUPPLEMENTARY MATERIAL
The Supplementary Material for this article can be found online at: http://www.frontiersin.org/journal/10.3389/fmicb.2014.00098/abstract

REFERENCES
Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Bacteriophage phiSfi21.

Bebeacu, C., Lorenzo Fajardo, J. C., Bollmann, S., Neve, H., et al. (2013). X-ray structure of a superinfection exclusion lipoprotein from phage TP-34 and identification of the tape measure protein as its target. Mol. Microbiol. 89, 152–165. doi: 10.1111/mmi.12267

Beres, S. B., Sylvia, G. L., Barbian, K. D., Lei, B., Hoff, J. S., Mammarella, N. D., et al. (2002). Genome sequence of a serotype M3 strain of group A Streptococcus: Phage-encoded toxins, the high-virulence phenotype, and clone emergence. Proc. Natl. Acad. Sci. U.S.A. 99, 10078–10083. doi: 10.1073/pnas.152298499

Beumer, A., and Robinson, J. B. (2005). A broad-host-range, generalized transducing phage (SN-T) acquires 16S rRNA genes from different genera of bacteria. Appl. Environ. Microbiol. 71, 8301–8304. doi: 10.1128/AEM.71.12.8301-8304.2005

Beurnier, A., and Robinson, J. B. (2005). A broad-host-range, generalized transducing phage (Sn-T) acquires 16S rRNA genes from different genera of bacteria. Appl. Environ. Microbiol. 71, 8301–8304. doi: 10.1128/AEM.71.12.8301-8304.2005

Beurrier, V., Mallaret, M., and Beumer, A. (2004). Complete sequence and comparative genome analysis of the dairy bacterium Streptococcus thermophilus. Nat. Biotechnol. 22, 1554–1558. doi: 10.1038/nbt1034

Brüssow, H., and Bruttin, A. (1995). Characterization of a temperate Streptococcus thermophilus bacteriophage and its genetic relationship with lytic phages. Virology 212, 632–640. doi: 10.1006/viro.1995.1521

Bruttin, A., Foley, S., and Brüssow, H. (1997). The site-specific integration system of the temperate Streptococcus thermophilus bacteriophage phiSfi21. Virology 237, 148–158. doi: 10.1006/viro.1997.7869

Campanacci, V., Veesler, D., Chakraborty, M. A., Bollmann, S., Sciara, G., Moineau, S., et al. (2010). Solution and electron microscopy characterization of lactococcal phage baseplates expressed in Escherichia coli J. Struct. Biol. 172, 75–84. doi: 10.1016/j.jsb.2010.02.007

Cuny, N., Davidson, A. R., and Maxwell, K. L. (2012). The moron comes of age. Bacteriophage 2, 225–228. doi: 10.4161/bact.23146

du Silva Oliveira, M. E., Mendes, D. Q., Ferrari, L. I., and Vasconcelos, A. T. R. (2004). Ribosome binding site usage according to the condition of the translation. Genet. Mol. Biol. 27, 644–650.

D’Aquila, R. T., Bechet, L. J., Vidler, J. A., Eron, J. J., Gorczycy, P., and Kaplan, J. C. (1991). Maximizing sensitivity and specificity of PCR by pre-amplification heating. Nucleic Acids Res. 19, 3749.
Desiere, F., Lucchini, S., and Brüssow, H. (1998). Evolution of Streptococcus thermophilus bacteriophage genomes by modular exchanges followed by point mutations and small deletions and insertions. Virology 241, 345–356. doi: 10.1006/viro.1997.8859

Desiere, F., Lucchini, S., Canchaya, C., Ventura, M., and Brüssow, H. (2002). Comparative genomics of phages and prophages in lactic acid bacteria. Antonie Van Leeuwenhoek 82, 73–91. doi: 10.1023/A:10200676823558

Deveau, H., Barrangou, R., Garneau, J. E., Labonte, J., Fremaux, C., Boyaval, P., et al. (2008). Phage response to CRISPR-encoded resistance in Streptococcus thermophilus. J. Bacteriol. 190, 1390–1400. doi: 10.1128/JB.01412-07

Don, R. H., Cox, P. T., Wainwright, B. J., Baker, K., and Mattick, J. S. (1991). “Touchdown” PCR to circumvent spurious priming during gene amplification. Nucleic Acids Res. 19:4008. doi: 10.1093/nar/19.14.4008

Donnelly-Wu, M. K., Jacobs, W. R. Jr., and Hatfull, G. F. (1993). Superinfection exclusion genes. Appl. Environ. Microbiol. 61, 4348–4356.

Lubbock, M. W., Waterfield, N. R., Beresford, T. P. J., Le Page, R. W. F., and Jarvis, A. W. (1995). Sequencing and analysis of the prolate-headed lactococcal bacteriophage c+ genotype and identification of its structural genes. Appl. Environ. Microbiol. 61, 2275–2284.

Lucchini, S., Desiere, F., and Brüssow, H. (1998). The structural gene module in Streptococcus thermophilus bacteriophage phiSfi1 shows a hierarchy of relatedness to Siphoviridae from a wide range of bacterial hosts. Virology 246, 63–73. doi: 10.1006/viro.1998.9190

Lucchini, S., Desiere, F., and Brüssow, H. (1999). Comparative genomics of Streptococcus thermophilus phage species supports a modular evolution theory. J. Virol. 73, 8647–8656.

Mahony, J., McGrath, S., Fitzgerald, G. F., and van Sinderen, D. (2008). Identification and characterization of lactococcal-prophage-carried superinfection exclusion genes. Appl. Environ. Microbiol. 74, 6206–6215. doi: 10.1128/AEM.01053-08

Maizel, J. V., and Lenk, R. P. (1981). Enhanced graphic matrix analysis of nucleic acid and protein sequences. Proc. Natl. Acad. Sci. U.S.A. 78, 7665–7669. doi: 10.1073/pnas.78.12.7665

Marchler-Bauer, A., Le, S., Anderson, J. B., Chitsaz, F., Derksby, M. K., DeWeese-Scott, C., et al. (2011). CDD: a conserved domain database for the functional annotation of proteins. Nucleic Acids Res. 39, 225–229. doi: 10.1093/nar/gkj1189

McGrath, S., Fitzgerald, G. F., and van Sinderen, D. (2002). Identification and characterization of phage-resistance genes in temperate lactococcal bacteriophages. Mol. Microbiol. 43, 509–520. doi: 10.1046/j.1365-2984.2002.02763.x

Mills, S., Griffin, C., O’Sullivan, O., Coffey, A., McAuliffe, O. E., Meijer, W. C., et al. (2011). A new phage on the ‘Mozzarella’ block: bacteriophage Sfi09 shares a low level of homology with other Streptococcus thermophilus phages. Int. Dairy J. 21, 963–969. doi: 10.1016/j.idairyj.2011.06.003

Montville, M. R., Ardestani, B., and Beller, G. L. (1994). Lactococcal bacteriophages require a host cell wall carbohydrate and a plasma membrane protein for adsorption and ejection of DNA. Appl. Environ. Microbiol. 60, 3204–3211.

Nambo, T., Minamino, T., MacNab, R. M., and Kutsukake, K. (1999). Peptidoglycan-hydrolyzing activity of the Filgp protein, essential for flagellar rod formation in Salmonella typhimurium. J. Bacteriol. 181, 1555–1561.

Neve, I., Blass, J., Fountoulakis, M., and Reidl, J. (1999). Characterization of the major control region of Vibrio cholerae bacteriophage K19: immunity, exclusion, and integration. J. Bacteriol. 181, 2902–2913.

Neve, H., Back, A., and Heller, K. J. (2004). Prophage screening in a Streptococcus thermophilus starter culture collection. Kieler Milchw. Forsch. 56, 265–275.

Neve, H., Freudenberg, W., Dietel-Fieddersen, E., Ehler, R., and Heller, K. J. (2003). Biology of the temperate Streptococcus thermophilus bacteriophage TP-J34 and physical characterization of the phage genome. Virology 315, 184–194. doi: 10.1016/j.viro.2004.02.014

Neve, H., Zenz, K. L., Desiere, F., Koch, A., Heller, K. L., and Brüssow, H. (1998). The lysogenic DNA module from the temperate Streptococcus thermophilus bacteriophage TP-J34: comparison with phage phiS21 and implications for the modular theory of phage evolution. Virology 241, 61–72. doi: 10.1006/viro.1997.8960

Padilla-Sanchez, V., Gao, S., Kim, H. R., Kihara, D., Sun, L., Rossmann, M. G., et al. (2013). Structure-function analysis of the DNA translocating portal of the bacteriophage T1 packaging machine. J. Mol. Biol. 426, 1019–1038. doi: 10.1016/j.jmb.2013.10.011

Pouz, C., van Sinderen, D., Suárez, J., García, P., Ladero, V., Fitzgerald, G. F., et al. (2001). The dilemma of phage taxonomy illustrated by comparative genomics of Sfi1-like Siphoviridae in lactic acid bacteria. J. Bacteriol. 184, 6026–6036. doi: 10.1128/JB.184.21.6026-6036.2002

Qüiberoni, A., Moineau, S., Rousseau, G. M., Reinheimer, J., and Ackermann, H.-W. (2010). Streptococcus thermophilus bacteriophages. Int. Dairy J. 20, 657–664. doi: 10.1016/j.idairyj.2010.03.012

Rutherford., K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M. A., et al. (2000). Artemis: sequence visualization and annotation. Bioinformatics 16, 944–945. doi: 10.1093/bioinformatics/16.10.944

Sambrook, J., and Russel, D. W. (2000). Molecular Cloning: A Laboratory Manual. New York, NY: Cold Spring Harbor Press.

Schultz, J., Milpetic, F., Bork, P., and Ponting, C. P. (1998). SMART, a simple module architecure research tool: identification of signaling domains. Proc. Natl. Acad. Sci. U.S.A. 95, 5857–5864. doi: 10.1073/pnas.95.11.5857
Stanley, E., Fitzgerald, G. F., Le Marrec, C., Fayard, B., and van Sinderen, D. (1997). Sequence analysis and characterization of phi O1205, a temperate bacteriophage infecting Streptococcus thermophilus CNRZ1205. Microbiology 143, 3417–3429. doi: 10.1099/00221287-143-11-3417

Stanley, E., Walsh, L., van der Zwaet, A., Fitzgerald, G. F., and van Sinderen, D. (2000). Identification of four loci isolated from two Streptococcus thermophilus phage genomes responsible for mediating bacteriophage resistance. FEMS Microbiol. Lett. 182, 271–277. doi: 10.1111/j.1574-6968.2000.tb08907.x

Streisinger, G. (1956). Phenotypic mixing of host range and serological specificities in bacteriophages T2 and T4. Virology 2, 388–398. doi: 10.1016/0042-6822(56)90033-2

Sun, S., Gao, S., Kondabagil, K., Xiang, Y., Rossmann, M. G., and Rao, V. B. (2012). Structure and function of the small terminase component of the DNA packaging machine in T4-like bacteriophages. Proc. Natl. Acad. Sci. U.S.A. 109, 817–822. doi: 10.1073/pnas.1110224109

Sun, X. (2002). Molecular and Functional Characterization of a Temperate Streptococcus thermophilus Phage TP-J34 Gene (ltp) Encoding a Membrane-Bound Lipoprotein. Ph.D. thesis, Christian-Albrechts-University, Kiel.

Sun, X., Göhler, A., Heller, K. J., and Neve, H. (2006). The ltp gene of temperate Streptococcus thermophilus phage TP-J34 encodes a lipoprotein which is expressed during lysogeny. Virology 350, 146–157. doi: 10.1016/j.viro.2006.03.001

Sun, Z., Chen, X., Wang, J., Zhao, W., Shao, Y., Wu, L., et al. (2011). Complete genome sequence of Streptococcus thermophilus strain ND03. J. Bacteriol. 193, 793–794. doi: 10.1128/JB.01374-10

Tremblay, D. M., and Moineau, S. (1999). Complete genomic sequence of the lytic bacteriophage DT1 of Streptococcus thermophilus. Virology 255, 63–76. doi: 10.1006/viro.1998.9525

Ventura, M., Bruttin, A., and Brussow, H. (2002). Transcription analysis of Streptococcus thermophilus phages in the lysogenic state. Virology 302, 21–32. doi: 10.1006/viro.2002.1571

Xu, J., Hendrix, R. W., and Duda, R. L. (2004). Conserved translational frameshift in dsDNA bacteriophage tail assembly genes. Mol. Cell 16, 11–21. doi: 10.1016/j.molcel.2004.09.006

Xu, J., Hendrix, R. W., and Duda, R. L. (2013). Chaperone-protein interactions that mediate assembly of the bacteriophage Lambda tail to the correct length. J. Mol. Biol. 426, 1004–1018. doi: 10.1016/j.jmb.2013.06.040

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

Received: 30 December 2013; accepted: 23 February 2014; published online: 13 March 2014.

Citation: Ali Y, Koberg S, Heßner S, Sun X, Rabe B, Back A, Neve H and Heller KJ (2014) Temperate Streptococcus thermophilus phages expressing superinfection exclusion proteins of the Ltp type. Front. Microbiol. 5:98. doi: 10.3389/fmicb.2014.00098

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

Copyright © 2014 Ali, Koberg, Heßner, Sun, Rabe, Back, Neve and Heller. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.