Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Genomics of tailless bacteriophages in a complex lactic acid bacteria starter culture

Svetlana Alexeeva
Wageningen Universiteit
Yue Liu
Wageningen Universiteit
Jingjie Zhu
Wageningen Universiteit
Joanna Kaczorowska
Wageningen Universiteit
Thijs R. H. M. Kouwen
DSM Biotechnology Center Delft
Tjakko Abee
Wageningen Universiteit
Eddy J. Smid (eddy.smid@wur.nl)
Wageningen University https://orcid.org/0000-0002-6687-5083

Research article

Keywords: Prophage, lysogenic bacteria, dairy starter culture, Lactococcus, Siphoviridae, phage defence mechanism, defective phage, tail disruption, comparative genomics

Posted Date: January 21st, 2020

DOI: https://doi.org/10.21203/rs.2.21416/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Version of Record: A version of this preprint was published at International Dairy Journal on March 1st, 2021. See the published version at https://doi.org/10.1016/j.idairyj.2020.104900.

Abstract

Background

Our previous study on a model microbial community originating from artisanal cheese fermentation starter revealed that bacteriophages not only co-exist with bacteria but also are highly abundant. To gain more insight into the potential role of prophages in the microbial community, we analysed the genomic content of 6 phage crops released by different strains in the starter culture, performed comparative genome analysis, and demonstrated their roles in phage defence of respective hosts.

Results

The identified prophages belong to three different subgroups of the Siphoviridae P335 phage group. Remarkably, most analysed prophages show disruptions in different tail encoding genes, resulting in a common tailless phenotype. Furthermore, a number of potentially beneficial features for the host carried by prophages were identified. The prophages carry up to 3 different phage defence systems per genome that are functional in protecting the host from foreign phage infection.

Conclusion

We suggest that the presumably defective prophages are a result of bacteria-phage coevolution and convey advantages to host bacteria; knowledge on the ecological role of such (defective) prophages may contribute to a refreshed look in strain selection criteria in (dairy) industry.

Background

Cells from all domains of life are susceptible to viral infections. As prokaryotes outnumber eukaryotes, their viruses (bacteriophages or simply phages), are estimated to be the most abundant biological entities on Earth. Total number of bacteriophages is estimated to be 10^{31} in the biosphere (Comeau et al., 2008). Phages are ubiquitously distributed in nature and play an important role in the ecology of their bacterial hosts. In complex microbial consortia such as those found in a marine environment, the gastrointestinal tract and in complex food fermentations, bacteriophages can alter the dynamics and diversity of microbial communities (Stern and Sorek, 2011; Spus et al., 2015; Williams, 2013; Shapiro and Kushmaro, 2011; Ohnishi et al., 2001; Erkus et al., 2013; Smid et al., 2014). Additionally, bacteriophages help to drive microbial evolution through phage-mediated gene transfer (Penadés et al., 2015; Canchaya et al., 2003).

Bacteriophages can also occur naturally in food. Lactic acid bacteria have been used for centuries in the production of fermented food products, with for instance Lactococcus lactis as an important player in various dairy fermentations. Bacteriophages infecting L. lactis strains are mostly studied because of their detrimental impact on industrial milk fermentation processes (Mahony and van Sinderen, 2015).

All L. lactis phages described are members of the Caudovirales order and possess a double-stranded DNA genome and a noncontractile tail. The vast majority of lactococcal phages belong to one of the three

main groups within Siphoviridae family: 936, c2, or P335 (Mahony et al., 2017). Groups 936 and c2 consist of only virulent phages, while P335 group consists of both temperate and virulent phages (Chmielewska-Jeznach et al., 2018). P335 phages resemble lambdoid phages, are genetically heterogeneous and have a mosaic genome structure with functional modules exchangeable through homologous recombination (Chopin et al., 2001; Desiere et al., 2002; Moineau et al., 1994). Maintenance of a temperate bacteriophage inside a bacterial chromosome in the form of prophage, also referred to as lysogeny, is a common phenomenon in L. lactis strains (Brondsted and Hammer, 2006; Kelleher et al., 2018). Nevertheless, lysogenic bacterial strains, carrying inducible prophages in the chromosome, usually do not find their way into commercial fermentation practice when it concerns a defined starter composition (Garneau and Moineau, 2011). It was surprizing, therefore, to discover that a naturally evolved complex starter culture (named Ur), featured by stable composition and robustness, is composed of mainly lysogenic strains (Alexeeva et al., 2018). Remarkably, all released bacteriophages from the complex starter culture Ur were found essentially tailless. Up to 10^{10} tailless phage particles per ml of culture are spontaneously produced by cultures of the individual strains without the occurrence of clear signs of cell lysis. Therefore, we postulated that the presence of tailless prophages must have translated into an evolutionary success of the lysogens despite the metabolic burden that the phages impose on host cells.

To identify potential beneficial or adaptive features encoded by phages and to obtain information on genetic diversity, genome organization, and their relevance in the microbial community, we subjected 6 representative phage crops to DNA sequencing followed by annotation and attribution of functions based on in silico approaches. Further, we experimentally demonstrated the functionality of selected prophage-encoded genetic elements.

Results

Genomic organisation and annotation

Whole phage genome sequences were obtained for 6 phage crops released by L. lactis strains TIFN1, TIFN2, TIFN4, TIFN5, TIFN6 and TIFN7 upon mitomycin C (MitC) induction. Only one phage genome per strain was assembled in single scaffold of high coverage (above 1100x). The sequence assembly was checked for correctness by comparison with restriction analysis data obtained by using two restriction enzymes (not shown). The data indicated that the assembly was correct and no additional fragments, that would indicate a presence of an additional released phage, were detectable. The bacteriophages released by TIFN1, TIFN2, TIFN4, TIFN5, TIFN6 and TIFN7 strains were named proΦ1, proΦ2, proΦ4, proΦ5, proΦ6 and proΦ7 respectively. The genome sequence of proΦ1 was found to be identical to proΦ5, and the sequence of proΦ2 was found to be identical to proΦ4.

General characteristics of the released phages such as the genome size, the GC content and the number of open reading frames (ORFs) are summarized in Table 1. The initiation codon AUG is present in most

protein-coding genes, but GUG and UUG are also found in all phage genomes albeit with lower frequency (Table 1).

Analysis of phage genomes revealed that all released phage genomes have mosaic structure, organized in two clusters of divergently transcribed genes, typical for temperate lactococcal phages. The cluster transcribed leftwards mainly comprises genes encoding functions for integration and maintenance of lysogeny. Genes encoding proteins involved in DNA replication, transcriptional regulation, packaging, structural proteins and phage release are mostly transcribed rightwards.

Detailed functional annotation with predicted protein functions are presented for each genome in Additional file 1, Tables S1-S4.

*Genome sequences of proΦ1 and proΦ5 are identical to each other, proΦ2 and proΦ4 are identical to each other.

Table 1
General characteristics of the 6 phage genomes

Phage	Genome size (bp)	% GC	ORFs total	The frequency of initiation codon usage, % AUG/GUG/UUG
proΦ1 & proΦ5*	41249	35.92	62	89/5/6
proΦ2 & proΦ4*	36976	35.61	49	90/4/6
proΦ6	37410	35.4	55	89/7/4
proΦ7	38158	35.24	56	95/2/4

Taxonomy and comparative genomics

To establish the degree of diversity between the newly described phages and their relatedness to other lactococcal phages, a comparative genome analysis was carried out. In total, 16 published P335 phage genomes, along with proΦ1 & proΦ5 (sequence identical), proΦ2 & proΦ4 (sequence identical), proΦ6 and proΦ7 were used to construct a phylogenetic tree (Fig. 1). The 20 phage genomes fall into 3 of the 4 known sub-groups described for P335 phages earlier (Kelly et al., 2013; Ventura et al., 2007). The phages proΦ1 & proΦ5 are most closely related to Tuc2009, TP901-1 and SK11-2. The genome sequence of proΦ6 showed most similarity to that of prophage SK11-3 and phage μl36. All of these phages fall into subgroup 2 family P335 bacteriophages, all of which seem to share pac-type packaging mechanism (Labrie et al., 2008).

The sequences of proΦ2 & proΦ4 clustered into subgroup 3 and proΦ7 is found among subgroup 4 bacteriophages, with bIL286 and r1t as most closely related bacteriophages respectively. The members of

these subgroups share, in contrast to the subgroup 2 members, cos-type packaging mechanism (Labrie et al., 2008).

Tail disruptions

Thorough sequence analysis of genes encoding structural tail elements (Table 2) provided an explanation for the tailless phenotype of the released phage particles (Alexeeva et al. 2018; additional EM pictures in file 1, Fig. S1). In proΦ1/proΦ5, ORF48 and ORF51 resemble most the N- and C terminal parts of a structural tail protein in prophage SK11-2 (98 and 94% amino acid (aa) identity, respectively) (Fig. 2A). ORF51 is 90% identical to 475 aa C terminal part of tail length tape measure protein (TMP, ORF45) of phage TP901-1, a tail protein determining tail length (Pedersen et al., 2000; Vegge et al., 2005). ORF48 shares 36% identity to N terminal part of TMP (ORF45) of phage TP901-1. Whereas SK11-2 and TP901-1 both encode a complete protein of 874 and 937 amino acids long respectively, in proΦ1/proΦ5 the two open reading frames are separated by insertion of two mobile elements encoded on the opposite strand (Fig. 2A). These observations suggest that ORF48 and ORF51 of proΦ1/proΦ5 encode the putative TMP that is disrupted by insertions, resulting in the tailless phenotype of the released phages.

In proΦ6 a different tail element is the target of disruption. ORF39 of proΦ6, 75 aa, is 97% identical to first 75 aa of ORF40 in Tuc2009 (102 aa) and 95% to ORF39 in TP901-1 (103 aa). This protein seems to be highly conserved, identical proteins are also present and intact in phages μl36, P335 and in proΦ1/proΦ5 (Fig. 2B). It functions at head-tail interface and has been described earlier as putative head to tail joining protein (Brøndsted et al., 2001). In proΦ6, however, the C-terminal sequence is separated from its N-terminus by insertion of a mobile element (transposase, ORF40) encoded on the opposite strand. This fits with the observed phenotype of proΦ6: TIFN6 is the only strain that released separated tails next to the phage heads (Alexeeva et al., 2018; additional EM pictures in Additional file 1, Fig. S1) indicating that the head to tail joining function is indeed impaired.

The 3450 bp-region encompassing ORFs 44, 45, 46 of proΦ7 shares 90% nucleotide identity to ORF40, 41 and 42 of phage r1t (Fig. 2C). ORF 44 and 46 in proΦ7 and ORF 40 and 42 in phage r1t encode the N and C-terminus of a TMP. It has been suggested that r1t ORF41, identical to proΦ7 ORF45 and separating N and C-terminus of the TMP, belongs to HNH homing endonuclease or a group I introns (van Sinderen et al., 1996; Mc Grath et al., 2006). However, r1t has been shown to possess a tail (Lowrie, 1974). On the other hand, phage phiLC3 encodes 843 aa long TMP, identical to r1t and proΦ7, but not interrupted by an insertion (Fig. 2C). It is therefore unclear, whether or not the insertion between N and C-terminus of proΦ7 TMP results in the tailless phenotype of proΦ7. In addition, no obvious disruptive elements could be identified in the tail module of proΦ2/proΦ4 despite the observed tailless phage morphology.

Potentially beneficial prophage encoded features

Next to the obvious phage-related features, e.g. integration, the regulation of lytic/lysogenic conversion and structural proteins, the prophages also encode proteins with potential benefits to the host. Phage-defence protein encoding genes are among the most frequently observed ones (Table 2). ORF2 in

proΦ1/proΦ5 is 98% identical to Sie family SieIL409 protein of Lactococcus phage 409, and SieIL 409 has been shown to mediate phage resistance by a DNA injection blocking mechanism (McGrath et al., 2002; Mahony et al., 2008). ORF14 of proΦ1/proΦ5 as well as ORF14 of proΦ2/proΦ4 encode adenine-specific DNA-methyltransferase (99% identical to ORF18 of phage Tuc2009) that may be a part of a restriction modification system (RMS) protecting the host against invading foreign DNA and reduce the efficiency of plaque formation (Labrie et al., 2010). The terminal ORF62 in proΦ1/proΦ5 also resembles a restriction/modification DNA-methylase and is possibly part of an RMS system. ORF61 of proΦ1/proΦ5 codes for an abortive infection AbiD/AbiF-like protein, another feature potentially involved in a phage defence mechanism: Abi systems allow phage absorption and phage DNA injection but interfere with further phage development, so that the death of infected cells occurs but no viral progeny is released (Labrie et al., 2010). Also proΦ7 carries two genes (ORF55 & ORF56) encoding Abi-like proteins near the attR terminus.

Table 2

Phage-defence protein encoding genes identified in Ur prophages.

Prophage	ORF	Putative product	Defence mechanism
proΦ1 / proΦ5	ORF2	Sie protein	superinfection-exclusion
	ORF14	Site-specific DNA-methyltransferase	restriction modification
	ORF61	Abi-like protein	abortive infection
	ORF62	N-4/N-6 DNA methylase	restriction modification
proΦ2 / proΦ4	ORF14	Adenine-specific methyltransferase	restriction modification
proΦ7	ORF55	abortive phage resistance protein	abortive infection
	ORF56	abortive phage resistance protein	abortive infection

Furthermore, proΦ6 carries a gene for a membrane protein related to a metallopeptidase (ORF49, with 80% identity to ORF53 in proΦ1/proΦ5). The product of ORF33 in proΦ2/proΦ4 is a putative protease (ATP-dependent serine endopeptidase, ClpP) and finally proΦ6 possesses prepillin peptidase (ORF13) encoding gene. Whether products of these genes could offer competitive advantages to the hosts in a microbial community by inhibiting other species remains to be elucidated.

Prophages provide phage resistance phenotype

The variety of phage resistance genes carried by prophage genomes suggests that the prophages likely contribute to host's ability to counteract superinfections. To assess this contribution we challenged lysogenic wild-type strains TIFN1 in parallel with its phage-cured derivative TI1c with 16 lactococcal phages (Additional file 1, Table S5).

Six phages from the phage collection were lysis-positive for TIFN1 and TI1c as revealed by spot tests, and clear differences in sensitivity of the two strains were already visualized towards some of the phages (example Fig. 3A). Further, phage sensitivity towards the 6 lysis-positive phages was quantified for the two strains (Fig. 3B). TIFN1 showed significantly ($p < 0.1$) lower phage sensitivity towards phages DSMΦ4, DSMΦ5 and DSMΦ6 compared to TI1c. Moreover, it was also noticed that the plaques were smaller and more opaque on TIFN1 than TI1c for these 3 phages (Additional file 1, Figure S2).

Discussion

In our previous study we described morphologically tailless bacteriophages, abundantly and continuously released by all analysed L. lactis strains originating from a complex dairy starter culture Ur without showing obvious cell lysis (Alexeeva et al., 2018). Because of their distinct morphology, behaviour and no apparent impact on host cell integrity, we hypothesized that these bacteriophages belong to a separate (novel) group of temperate lactococcal phages. However, detailed genome analysis of the released phages presented here, revealed that the bacteriophages possess a typical lactococcal P335 group Siphoviridae family genome structure, and that the phages fall under three different known subgroups of P335 phages (Ventura et al., 2007).

Temperate lactococcal phages of Siphoviridae family, belonging to the P335 group, are usually characterised by a long non-contractile tail - a structure responsible for host recognition, adsorption and the initiation of phage infection by envelope penetration and DNA ejection (Mc Grath et al., 2006). Despite the fact that all inducible prophages found in strains from the complex starter culture Ur so far are essentially tailless, phage sequencing revealed the presence of genes encoding most of the tail structural elements: head-tail connector, major tail protein (MTP), tail length tape-measure protein (TMP), distal tail protein, tail associated lysin, upper and lower base plate protein (Vegge et al., 2005; Veesler and Cambillau, 2011; Veesler et al., 2012; Stockdale et al., 2015). However, detailed sequence analysis identified the presence of insertions of mobile genetic elements in the tail module in most of the phage genomes analysed in this study. Furthermore, the insertions occurred at different sites of the tail module: proΦ1, proΦ5 and proΦ7 contained (different) insertions in the TMP while proΦ6 in the head-tail joining protein. TMP determines the length of the phage tail (Pedersen et al., 2000; Katsura, 1987) and serves as a component of the precursor complex, involved in the initiation of polymerization of MTP. However, it has been shown for λ phage that in the absence of TMP, MTP polymerization may be initiated but the formation of tail-tube related structure is abolished (Katsura, 1976). The tailless morphology of proΦ2/proΦ4 could not be explained by this sequence analysis as no obvious disruptive elements could be identified in the tail modules. It is postulated that the absence of tails for the phage proΦ2/proΦ4 belonging to these particular L. lactis strains is caused by modifications at transcriptional and translational level of the phage genes, or minor mutations in gene sequence.

Bacteria-phage coevolution has been regarded as an important driver of evolutionary processes and an essential player in shaping of microbial communities (Koskella and Brockhurst, 2014). This is also reflected in the results of the phage genomics analysis performed in this study. In the Ur strains, most of

the inducible prophages have mutational insertions in different tail encoding genes, resulting in tailless phage particles that are likely to be defective. In fact, defective prophages are commonly observed in bacterial genomes. Out of more than 200 prophages from 83 bacterial genomes analysed in a study (Casjens, 2003), only 9 prophages were experimentally shown to be fully functional. All other prophages were found to have experienced different levels of mutational decay. Moreover, it is also acknowledged that many genes in the defective prophages remain functional and contribute to various traits of the hosts, and that the prophage functions are a result of purifying selection in the bacterial chromosome (Bobay et al., 2014; Kelleher et al., 2018). Prophage genes encoding core phage-related functions, e.g. tail and lysis proteins, were found to be under stronger purifying selections (Bobay et al., 2014), presumably due to the critical functions in phage spreading or host integrity carried by these genes. This finding coincides with our observation that prophages in bacterial community members of the dairy starter culture Ur, showed disruptions in the tail protein encoding genes. This could be explained, as the loss of the key structure to re-infect the same bacterial species was likely advantageous for the species to be maintained in the microbial community. In addition, the genetic analysis also clearly identified prophage-encoding phage resistance systems in this study, and the phage resistance phenotype of the host was confirmed experimentally; this could be also a result of purifying selection.

Another uncommon phenotype of the studied Ur phages was the spontaneous, continuous release of phage particles, even when no stress or prophage-inducing condition was applied (Alexeeva et al., 2018). For lambda phage the spontaneous excision rates are approximately 10^{-6} per cell division (Gottesman and Yarmolinsky, 1968; Bobay et al., 2014), while up to 10^{10} phage particles are released by Ur strains spontaneously (Alexeeva et al., 2018). When examining the genome sequences, we did not identify obvious disruptions in prophage repressor or translational regulator genes. However, it is plausible that the phage-host coevolution has resulted in (minor) mutations in sequence, or modifications in the transcriptional or translational levels in the regulatory elements, that allow continuous assembly of these phage particles. Notably, the fitness of the host does not seem to be compromised by the continuous phage release, when comparing the growth performance of TIFN1 to its proΦ1-cured derivative based on OD measurement (unpublished data). This suggests that instead of conferring disadvantageous metabolic burdens to the host, the continuous phage producing phenotype may inhibit successful infection of other incoming phages by competing and interfering with the assembly process and therefore provide advantage to the bacterial host in a microbial community.

All these could serve as the explanation of the phenomenon that in the bacterial community of the dairy starter Ur, presumably after a long-term selection, predominantly strains with a prophage sequence were maintained. In conclusion, the analysis of the genomic content of 6 phage crops released by different strains in the starter culture Ur provided indications for bacteria-phage coevolution and selection for phage defence systems, and this may also provide new leads in for future research and implications in practise, for example in defining strain selection criteria in (dairy) industry, where traits like phage resistance are desired.

Methods

L. lactis strains and bacteriophages

Representative strains of Lactococcus lactis TIFN1-TIFN7 were used throughout this study. These strains represent isolates from different genetic lineages originally isolated from single colonies from a complex starter culture Ur and the genome sequence has been determined (Erkus et al., 2013). L. lactis TI1c, TI2c and TI4c are phage cured derivatives of TIFN1, TIFN2 and TIFN4 described earlier (Alexeeva et al., 2018). The strains were maintained as 15% glycerol stocks at -80 °C and routinely grown in M17 broth (OXOID) with 0.5% (wt/vol) glucose or lactose addition (OXOID). Bacteriophages used in the phage sensitivity screen are listed in Additional file 1, Table S5.

Phage preparation and DNA isolation

Induction of the prophages was performed using mitomycin C (the stock 0.5 mg/ml solution in 0.1M $MgSO_4$ was stored at 4°C). Overnight cultures were grown in M17 supplemented with 0.5% glucose at 30°C. The cells were diluted in fresh medium to OD_{600} = 0.2 and incubated 1–2 hours, until middle/late exponential phase was reached (OD_{600} = 0.4–0.5). At this point mitomycin C solution was added to final concentration 1 µg/ml. The cultures were further incubated for 6 hours, as OD_{600} was monitored hourly. Cultures were centrifuged for 15 minutes at 6000 x g at 4 °C and the supernatant was filter-sterilized using 0.2 µm polyethersulfone (PES) filters. The phage particles in the filtrate were precipitated using 1 volume of 20% polyethylene glycol 8000 and 2.5M NaCl to 4 volumes crude bacteriophage suspension. The samples were incubated overnight at 4 °C and then centrifuged at 11000 x g for 60 min at 4 °C. Supernatants were discarded and the pellets were left to dry on clean absorbent paper.

The pellets were directly used for DNA extraction using Promega Wizard® Genomic DNA Purification Kit. The phage pellets were re-suspended in 600 µl of Nuclei Lysis Solution and transferred into clean 1.5 ml micro-centrifuge tubes. The suspension was incubated at 80 °C for 5 min and then cooled to room temperature. Three microliter of RNase solution (provided with the kit) was added to each sample. The samples were mixed by inversion and incubated at 37 °C for about 50 min. Proteinase K (20 mg/ml) was added to a final concentration 200 mg/l and then samples were incubated at 50 °C for 30 min. The samples were cooled to room temperature and 200 µl of Protein Precipitation Solution was added to the tubes. The tubes were vortexed vigorously for 20 seconds and incubated on ice for 5 min. After this step the samples were centrifuged (17000 x g, 3 min, room temperature). The supernatant was transferred in a new 1.5 ml micro-centrifuge tube containing 600 µl of isopropanol kept at room temperature. Next the samples were mixed by inversion until strands of DNA were visible, centrifuged (17000 x g, 2 min, room temperature) and the supernatant was poured off. Six hundred microliter of 70% ethanol (kept at room temperature) was added to the tubes. Tubes were inverted several times to wash the DNA pellet and then centrifuged (17000 x g, 2 min, room temperature). Ethanol was allowed to evaporate and the tubes were dried on clean absorbent paper. The tubes were left open for 15 min to air dry the pellet. 100 µl of DNA

Rehydratation Solution was added and DNA was rehydrated at 65 °C for 1 h. Alternatively, the DNA was rehydrated by incubating the solution overnight at 4 °C.

The isolated and rehydrated DNA was additionally purified from low molecular weight DNA species using Amicon® Ultracel 100K columns (cut-off 100 kDa or double-stranded nucleotide cut-off > 600 bp). DNA solution was brought to 0.5 ml using 5 mM Tris buffer (pH = 8) and added to the column. The samples were centrifuged at 14 000 x g and washed 3 times with 0.5 ml buffer. The quality of DNA before and after micro-column purification was examined by agarose gel electrophoresis.

Sequencing and assembly of sequences

Library preparation using TruSeq DNA sample kit and genome sequencing was performed by BaseClear BV (Leiden, The Netherlands). A paired-end DNA library with a mean gap length size between 230 and 360 bp was constructed and the sequencing was performed using HiSeq 2500 Illumina technology (Illumina Inc, Hayward, CA) using a 50-cycle or 100-cycle (proΦ7) paired-end protocol. The sequencing yielded on average 1.350.000 reads (135 MB, for 50-cycle) and 800.000 reads (155 MB, for 100-cycle).

Next-generation assembly was performed using SeqMan NGen de novo algorithm (v12, DNAStar, USA). In 5 of 6 cases 68–88% of all reads assembled in a single contig (see Table 3 for assembly statistics).

Table 3
Phage genome assembly statistics.

	N reads	Contig	Contig length raw	Contig coverage	Background average coverage	% reads in contig	x covarage above background
proΦ1	10,000,000	Contig 1	41886	7211	59	68	121
proΦ2	3,336,448	Contig 1	37935	3175	27	75	119
proΦ4	2,927,954	Contig 1	38129	3106	28	88	112
proΦ5	2,288,894	Contig 3	41186	2292	40	88	58
proΦ6	2,248,210	Contig 1	38549	2189	38	82	58
proΦ7	1,547,972	Contig 15	40045	1146	47	29	24

The sequences were initially subjected to automated annotation using MyRAST (Rapid Annotation Subsystem Technology (RAST) server (Aziz et al., 2008). All predicted protein-coding genes were screened

using BLASTP and Psi-BLAST algorithms against the non-redundant protein database at NCBI and for conserved motifs using InterProScan 5 (Jones et al., 2014).

All contigs > 600 nt long with high coverage or > 25000 nt long obtained from SeqMan NGen analyses were annotated using MyRAST server (Aziz et al., 2008) and screened for phage-related features.

The sequence assembly was checked for correctness by comparison with restriction data obtained by using several restriction enzymes (not shown). The data indicated that the assembly was correct.

Phylogenetic analysis and classification

Alignment of sequences was run on MAFFT server. The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura-Nei model (Tamura and Nei, 1993). The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Joining and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site (next to the branches). The analysis involved 10 and 22 nucleotide sequences. All positions containing gaps and missing data were eliminated. Evolutionary analyses were conducted in MEGA6 (Tamura et al., 2013).

Nucleotide and protein sequences accession numbers

The sequence data reported in the present study have been deposited in GenBank database under accession no. MN534315 - MN534320. The complete genomic sequences of the P335 group phages analysed in this study are available under the following GenBank accession numbers: phage name (accession number):, (P335 (DQ838728), 4268 (AF489521), bIL285 (AF323668), bIL286 (AF323669), bIL309 (AF323670), BK5-T (AF176025), phiLC3 (AF242738), r1t (U38906), TP901-1 (AF304433), Tuc2009 (AF109874), and µl36 (AF349457). Prophage sequences of L. lactis SK11 (SK11-1, SK11-2, SK11-3, SK11-4 and SK11-5) were derived from (NC_008527) (Ventura et al., 2007).

Bacteriophage sensitivity tests

Plaque assays were conducted to quantify the phage sensitivities of wild type strains and their cured derivatives. One hundred microliter of phage suspensions in 3–4 dilution series were mixed with 100 µl overnight culture of the target bacterium. The mixtures were incubated for 10 min prior to adding to tubes containing 2.8 ml soft agar. After 24-h incubation at 30 °C, plaques were counted and plaque-forming units per milliliter values were calculated.

Abbreviations

MitC: mitomycin C

ORF: open reading frame

RMS: restriction modification system

Abi: abortive infection

MOI: multiplicity of infection

RT: room temperature

TMP: tail length tape-measure protein

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare no competing interests.

Funding

This study was financed by Top Institute Food and Nutrition (TIFN) in Wageningen, the Netherlands.

Authors' contributions

SA and EJS conceived the study. SA, EJS and YL designed the experiments. SA and YL executed the experiments and carried out the data analysis and interpretation. SA, YL, TA and EJS wrote the manuscript. JZ performed phage sensitivity tests, JK participated in purification of phage DNA, TK co-designed phage sensitivity test. All authors read and approved the final manuscript.

Acknowledgements

The authors are grateful to Laurens Hanemaaijer of DSM for technical support in phage sensitivity tests, acquiring and sharing of phages DSMF1-DSMF10 suspension stocks, Michiel Wels for support with

sequence assemblies and Anne de Jong for support with sequence annotations. We also thank students of Wageningen University for their contribution: Venera Proveva for purification of proF4, and proF5, and proF6 DNA; Yixin Ge for purification of proF7 DNA. We thank Dr Jennifer Mahony of University College Cork, Cork, Ireland for kindly providing us bacteriophages sk1, p2 and jj50. We also gratefully acknowledge Jan van lent (Wageningen Electron Microscopy Centre (WEMC)) for his help with electron microscopy.

References

Alexeeva S, Guerra Martínez JA, Spus M, et al. Spontaneously induced prophages are abundant in a naturally evolved bacterial starter culture and deliver competitive advantage to the host. BMC Microbiol. 2018;18:120. doi:10.1186/s12866-018-1229-1

Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, et al: The RAST Server: rapid annotations using subsystems technology. BMC genomics. 2008;9:75.

Bobay L, Touchon M, Rocha, EP. Pervasive domestication of defective prophages by bacteria. PNAS. 2014;111:12127-32. doi:10.1073/pnas.1405336111

Brøndsted L, Hammer K. Phages of Lactococcus lactis. In: Calendar R, editor. The Bacteriophages. New York: Oxford University Press, Inc.; 2006. P. 572-592.

Brøndsted L, Østergaard S, Pedersen M, Hammer K, Vogensen FK. Analysis of the complete DNA sequence of the temperate bacteriophage TP901-1: evolution, structure, and genome organization of lactococcal bacteriophages. Virology. 2001;283:93-109.

Canchaya C, Fournous G, Chibani-Chennoufi S, Dillmann M-L, Brüssow H. Phage as agents of lateral gene transfer. Curr Opin Microbiol. 2003;6:417-424.

Casjens S. Prophages and bacterial genomics: what have we learned so far? Mol Microbiol. 2003;49:277-300. doi:10.1046/j.1365-2958.2003.03580.x

Chandry PS, Moore SC, Boyce JD, Davidson BE, Hillier AJ: Analysis of the DNA sequence, gene expression, origin of replication and modular structure of the Lactococcus lactis lytic bacteriophage sk1. Mol Microbiol. 1997;26:49-64.

Chmielewska-Jeznach M, Bardowski JK, Szczepankowska AK. Molecular, physiological and phylogenetic traits of Lactococcus 936-type phages from distinct dairy environments. Sci Rep. 2018;8:12540. doi:10.1038/s41598-018-30371-3

Chopin A, Bolotin A, Sorokin A, Ehrlich SD, Chopin M. Analysis of six prophages in Lactococcus lactis IL1403: different genetic structure of temperate and virulent phage populations. Nucleic Acids Res. 2001;29:644-651.

Comeau AM, Hatfull GF, Krisch HM, Lindell D, Mann NH, Prangishvili D. Exploring the prokaryotic virosphere. Res Microbiol. 2008;159:306-313.

Desiere F, Lucchini S, Canchaya C, Ventura M, Brussow H. Comparative genomics of phages and prophages in lactic acid bacteria. A Van Leeuw J Microb. 2002;82:73-91.

Erkus O, de Jager VC, Spus M, van Alen-Boerrigter IJ, van Rijswijck IM, Hazelwood L, et al. Multifactorial diversity sustains microbial community stability. ISME J. 2013;4:108.

Garneau JE, Moineau S. Bacteriophages of lactic acid bacteria and their impact on milk fermentations. Microb Cell Fact. 2011;10:S20.

Gottesman ME, Yarmolinsky MB. Integration-negative mutants of bacteriophage lambda. J Mol Biol 1968;31:487–505.

Higgins D, Sanozky-Dawes RB, Klaenhammer TR. Restriction and modification activities from Streptococcus lactis ME2 are encoded by a self-transmissible plasmid, pTN20, that forms cointegrates during mobilization of lactose-fermenting ability. J Bacteriol. 1988;170:3435-3442.

Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, et al. InterProScan 5: genome-scale protein function classification. Bioinformatics. 2014;30:1236-40. doi: 10.1093/bioinformatics/btu031.

Josephsen J, Vogensen FK. Identification of three different plasmid-encoded restriction/modification systems in Streptococcus lactis subsp. cremoris W56. FEMS Microbiol Lett. 1989;59:161-166.

Katsura I. Determination of bacteriophage lambda tail length by a protein ruler. Nature. 1987;327:73-75.

Katsura I. Isolation of λ prophage mutants defective in structural genes: their use for the study of bacteriophage morphogenesis. Mol Gen Genet. 1976;148:31-42.

Kelleher P, Mahony J, Schweinlin K, Neve H, Franz CM, van Sinderen D. Assessing the functionality and genetic diversity of lactococcal prophages. Int J Food Microbiol. 2018;272:29-40. doi: 10.1016/j.ijfoodmicro.2018.02.024.

Kelly WJ, Altermann E, Lambie SC, Leahy SC. Interaction between the genomes of Lactococcus lactis and phages of the P335 species. Front Microbiol 2013;4:257. doi: 10.3389/fmicb.2013.00257

Koskella B, Brockhurst MA. Bacteria-phage coevolution as a driver of ecological and evolutionary processes in microbial communities. FEMS Microbiol Rev. 2014;38:916–931. doi:10.1111/1574-6976.12072

Labrie SJ, Samson JE, Moineau S. Bacteriophage resistance mechanisms. Nat Rev Micro. 2010;8:317-327.

Labrie SJ, Josephsen J, Neve H, Vogensen FK, Moineau S. Morphology, genome sequence, and structural proteome of type phage P335 from Lactococcus lactis. Appl Environ Microb. 2008;74:4636-4644.

Lowrie R. Lysogenic strains of group N lactic streptococci. Appl Microbiol. 1974;27:210-217.

Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res. 2019;47:W636-W641. doi:10.1093/nar/gkz268.

Mahony J, McGrath S, Fitzgerald GF, van Sinderen D. Identification and characterization of lactococcal-prophage-carried superinfection exclusion genes. Appl Environ Microb. 2008;74:6206-6215.

Mahony J, Oliveira J, Collins B, Hanemaaijer L, Lugli GA, Neve H, et al. Genetic and functional characterisation of the lactococcal P335 phage-host interactions. BMC Genomics. 2017;18:146. doi: 10.1186/s12864-017-3537-5.

Mahony J, van Sinderen D. Novel strategies to prevent or exploit phages in fermentations, insights from phage-host interactions. Curr Opin Biotech. 2015;32:8-13.

McGrath S, Fitzgerald GF, van Sinderen D. Identification and characterization of phage-resistance genes in temperate lactococcal bacteriophages. Mol Microbiol. 2002;43:509-520.

Mc Grath S, Neve H, Seegers JF, Eijlander R, Vegge CS, Brøndsted L,et al. Anatomy of a lactococcal phage tail. J Bacteriol. 2006;188:3972-3982.

Moineau S, Pandian S, Klaenhammer TR. Evolution of a Lytic Bacteriophage via DNA Acquisition from the Lactococcus lactis Chromosome. Appl Environ Microb. 1994;60:1832-1841.

Ohnishi M, Kurokawa K, Hayashi T. Diversification of Escherichia coli genomes: are bacteriophages the major contributors? Trends Microbiol 2001;9:481-485.

Penadés JR, Chen J, Quiles-Puchalt N, Carpena N, Novick RP. Bacteriophage-mediated spread of bacterial virulence genes. Curr Opin Biotech. 2015;23:171-178.

Shapiro OH, Kushmaro A. Bacteriophage ecology in environmental biotechnology processes. Curr Opin Biotech. 2011;22:449-455.

Smid EJ, Erkus O, Spus M, Wolkers-Rooijackers JC, Alexeeva S, Kleerebezem M. Functional implications of the microbial community structure of undefined mesophilic starter cultures. Microb Cell Fact. 2014;13:S2.

Spus M, Li M, Alexeeva S, Wolkers-Rooijackers J, Zwietering M, Abee T, et al. Strain diversity and phage resistance in complex dairy starter cultures. J Dairy Sci. 2015;98:5173-5182.

Stern A, Sorek R. The phage-host arms race: shaping the evolution of microbes. Bioessays. 2011;33:43-51.

Stockdale SR, Collins B, Spinelli S, Douillard FP, Mahony J, Cambillau C, et al. Structure and Assembly of TP901-1 Virion Unveiled by Mutagenesis. PLoS One 2015;10:e0131676. doi: 10.1371/journal.pone.0131676.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol. 2013;30:2725-2729.

Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol. 1993;10:512-526.

Ventura M, Zomer A, Canchaya C, O'Connell-Motherway M, Kuipers O, Turroni F, et al. Comparative analyses of prophage-like elements present in two Lactococcus lactis strains. Appl Environ Microb. 2007;73:7771-7780.

Williams HT. Phage-induced diversification improves host evolvability. BMC Evol Biol. 2013;13:17.

Pedersen M, Ostergaard S, Bresciani J, Vogensen FK. Mutational analysis of two structural genes of the temperate lactococcal bacteriophage TP901-1 involved in tail length determination and baseplate assembly. Virology. 2000;276:315-328.

van Sinderen D, Karsens H, Kok J, Terpstra P, Ruiters M, Venema G, et al. Sequence analysis and molecular characterization of the temperate lactococcal bacteriophage r1t. Mol Microbiol. 1996;19:1343-1355.

Veesler D, Cambillau C. A common evolutionary origin for tailed-bacteriophage functional modules and bacterial machineries. Microbiol Mol Biol R. 2011;75:423-433.

Veesler D, Spinelli S, Mahony J, Lichiere J, Blangy S, Bricogne G, et al: Structure of the phage TP901-1 1.8 MDa baseplate suggests an alternative host adhesion mechanism. PNAS. 2012;109:8954-8958.

Vegge CS, Brøndsted L, Neve H, Mc Grath S, van Sinderen D, Vogensen FK. Structural Characterization and Assembly of the Distal Tail Structure of the Temperate Lactococcal Bacteriophage TP901-1. J Bacteriol. 2005;187:4187-4197.

Figures

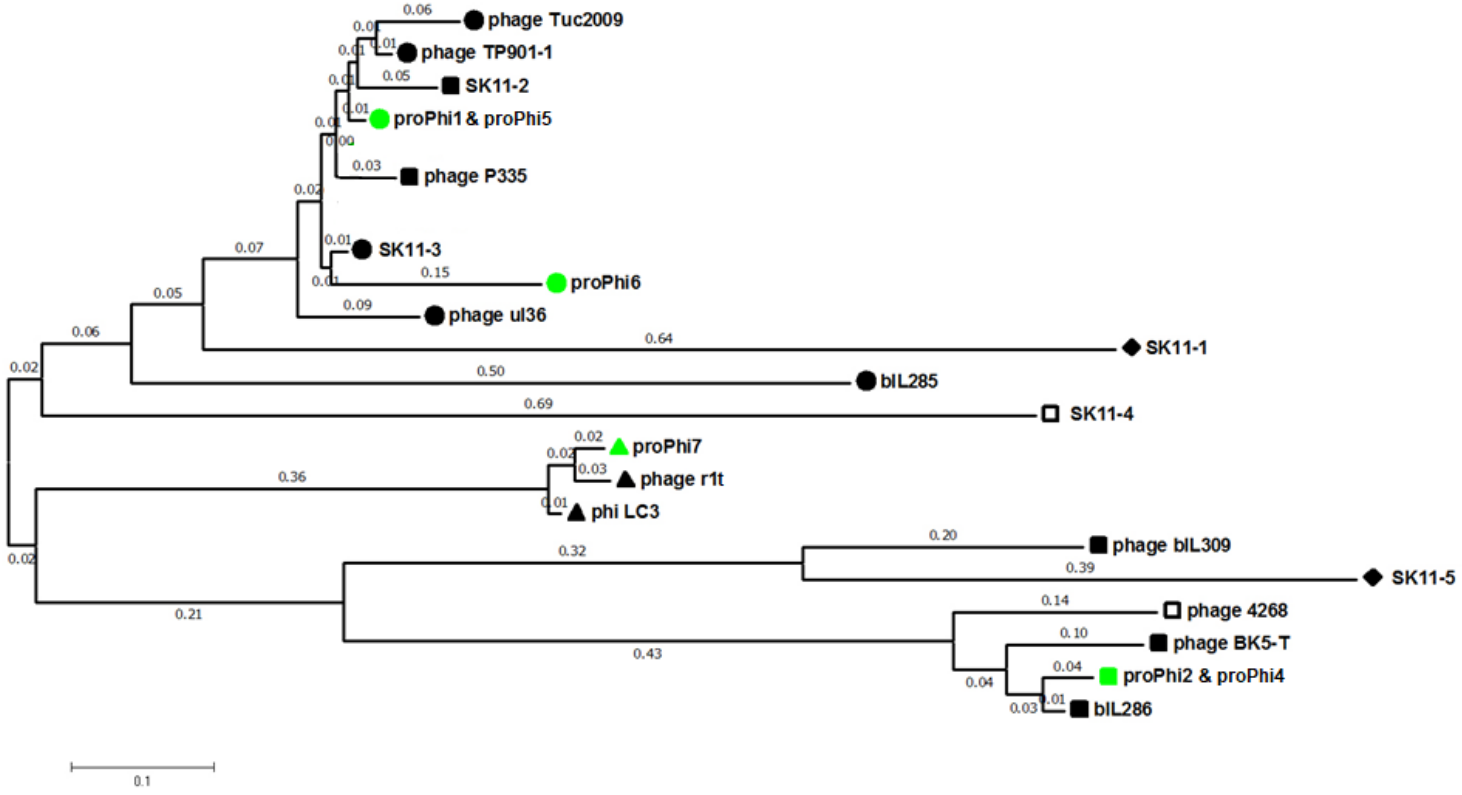

Figure 1

The phylogenetic tree of lactococcal phages including the bacteriophages described in this study (proΦs). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site (next to the branches). The analysis involved 20 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 917 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 (Tamura et al., 2013). Black filled symbols: P335 group, circles - subgroup 2, squares – subgroup 3, triangles – subgroup 4, diamonds – subgroup 1 as classified based on proteomic analysis earlier (Kelly et al., 2013; Ventura et al., 2007). Open symbols – phage 4268 was not analysed in (Kelly et al., 2013; Ventura et al., 2007); phage SK11-4 was unclassified. Green filled symbols – newly classified Ur phages. The numbers indicate branch length.

Figure 2

Alignment of structural tail elements encoding regions of proΦs, disrupted by insertions of mobile elements. The tail elements are aligned to their homologous counterparts of earlier described bacteriophages using MUSCLE (Madeira et al., 2019). Green bar represent homology regions, similarity profile height represents the level of conservation, regions in grey lack detectable homology. Corresponding scale (in base pair) and encoded features are shown under each sequence. The

alignments are shown for A) putative TMP of proΦ1/proΦ5; B) putative head-tail joining protein of proΦ6; C) putative TMP of proΦ7. No disruption in the tail elements was identified in proΦ2/proΦ4 and therefore no information of these phages are shown in this figure.

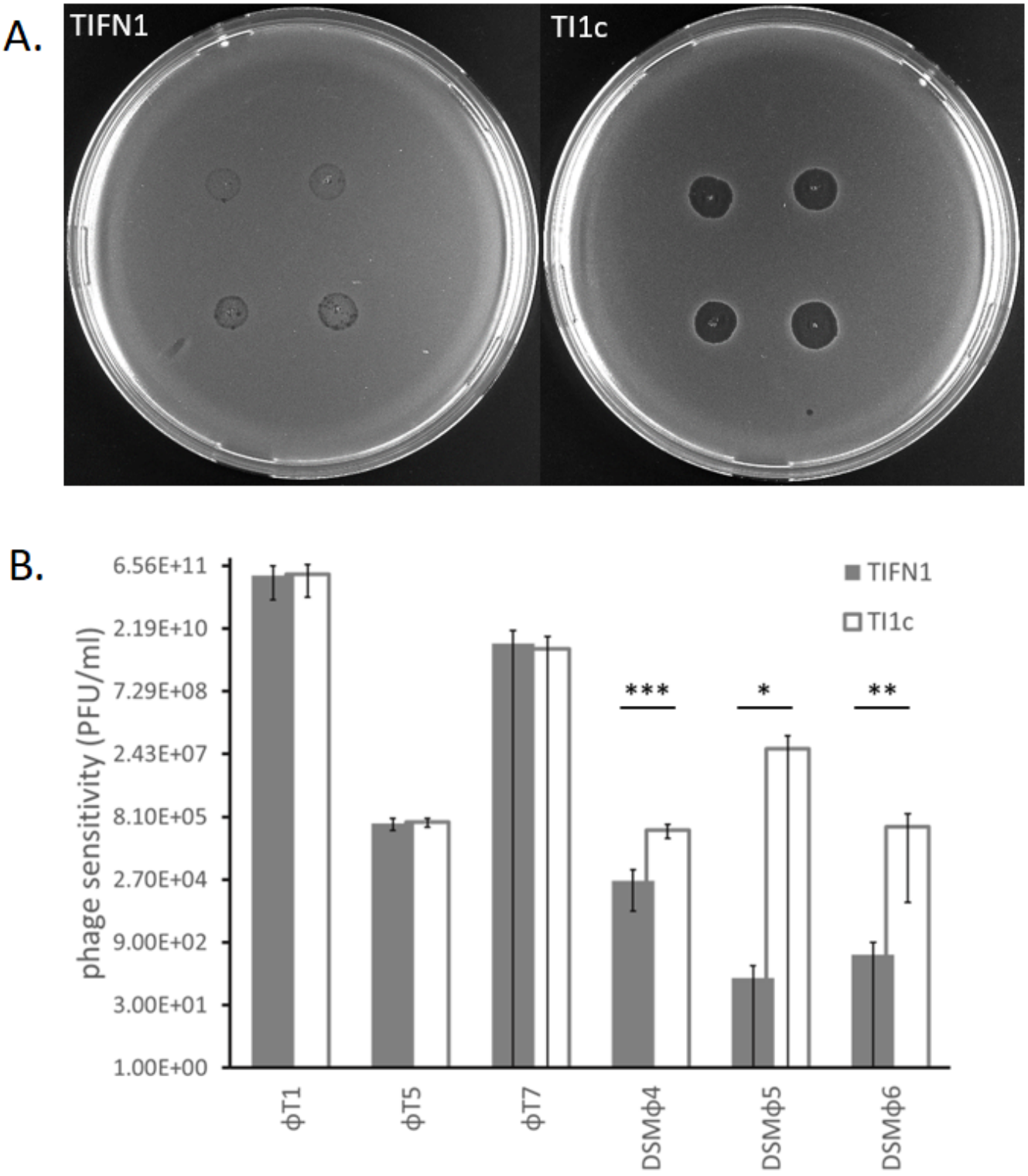

Figure 3

Phage sensitivity in TIFN1 and TI1c. A) Spot clearance and morphology produced on the lawns of TIFN1 (left) and TI1c (right) by DSMΦ5 (upper spots) and DSMΦ6 (lower spots), in duplicate. B) Quantification

of phage sensitivity by plague assay. * indicates p < 0.1; **indicates P < 0.05; *** indicates p < 0.01.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementaryinfoforsubmissionwithauthorsuggestions.docx