

Characterization and Genomic Analysis of Phage ascc ϕ 28, a Phage of the Family *Podoviridae* Infecting *Lactococcus lactis*[∇]

Steven E. Kotsonis,¹† Ian B. Powell,^{2*} Christopher J. Pillidge,² Gaëtan K. Y. Limsowtin,² Alan J. Hillier,³ and Barrie E. Davidson¹

Department of Biochemistry and Molecular Biology, The University of Melbourne, Melbourne, Victoria 3010, Australia¹; Cultures Division, Dairy Innovation Australia, Ltd., Private Bag 16, Werribee, Victoria 3030, Australia²; and Food Science Australia, Private Bag 16, Werribee, Victoria 3030, Australia³

Received 23 October 2007/Accepted 23 March 2008

Bacteriophage ascc ϕ 28 infects dairy fermentation strains of *Lactococcus lactis*. This report describes characterization of ascc ϕ 28 and its full genome sequence. Phage ascc ϕ 28 has a prolate head, whiskers, and a short tail (C2 morphotype). This morphology and DNA hybridization to *L. lactis* phage P369 DNA showed that ascc ϕ 28 belongs to the P034 phage species, a group rarely encountered in the dairy industry. The burst size of ascc ϕ 28 was found to be 121 ± 18 PFU per infected bacterial cell after a latent period of 44 min. The linear genome (18,762 bp) contains 28 possible open reading frames (ORFs) comprising 90% of the total genome. The ORFs are arranged bidirectionally in recognizable functional modules. The genome contains 577 bp inverted terminal repeats (ITRs) and putatively eight promoters and four terminators. The presence of ITRs, a phage-encoded DNA polymerase, and a terminal protein that binds to the DNA, along with BLAST and morphology data, show that ascc ϕ 28 more closely resembles streptococcal phage Cp-1 and the ϕ 29-like phages that infect *Bacillus subtilis* than it resembles common lactococcal phages. The sequence of this phage is the first published sequence of a P034 species phage genome.

Bacteriophage infection of *Lactococcus lactis* starter cultures used in cheese making and in other dairy fermentations affects product quality and can result in fermentation failure. Because of these worldwide industrial and financial consequences, phages infecting *L. lactis* are among the best-studied and most commonly isolated phages infecting any bacterial species (1, 6, 9, 20). As occurs with several well-studied bacterial species (such as *Escherichia coli* and *Bacillus subtilis*), a diverse range of phage types has been found to infect *L. lactis*.

Taxonomic classification of lactococcal phages continues to be an area of debate. All *L. lactis* phages documented so far are double-stranded DNA phages, and the vast majority have a *Siphoviridae* morphotype (long noncontractile tail, Bradley morphotypes B1 and B2). Using mainly morphology and DNA-DNA hybridization data, Jarvis et al. (20) divided these phages into 11 or 12 species. The status of the T187 group of phages was doubtful because of DNA homology between it and phage BK5-T (34). Later, Proux et al. (33) reported recognizable modular similarities among phages infecting a range of gram-positive bacteria and proposed five genera of lactococcal phages in a phage classification scheme (not confined to lactococcal phages) partially based on the phylogeny of phage genome modules deduced from DNA sequence comparisons. However, this scheme included only phage types for which whole-genome sequence data were available. No sequence information is available for some phage types, so they cannot yet

be classified in this way. Moreover, some reported types (e.g., RZh and P107) are possibly extinct (9, 20), at least in the formal taxonomic sense that reference examples are no longer available.

In the latest review of the biodiversity and classification of lactococcal phages, Deveau et al. (9) revised the species scheme proposed by Jarvis et al. (20) so that it included eight species, and they added two novel species belonging to the family *Siphoviridae* (1706 and Q54). The 936 (small isometric-headed) and c2 (prolate-headed) phage species each comprise only lytic phages. There is a high degree of genetic conservation within each of these two different species. By contrast, the P335 species, which contains both temperate and lytic phages, has greater genome diversity (9, 11, 28, 33). Based on an apparent high degree of genetic exchange between the formerly recognized 1483, T187, BK5-T, and P335 species, Deveau et al. (9) proposed merging these species into a revised P335 species. Recent DNA sequencing has confirmed that phages Q54 (11) and 1706 (14) are new phage species.

Most *L. lactis* phages isolated from dairy fermentations belong to the 936 (most common), c2, or P335 species (9, 19, 20, 23, 31). It is no coincidence, therefore, that most publicly available *L. lactis* phage genome sequences belong to one of these species (9). The reported ratio of phage species isolated in the dairy industry varies, perhaps depending on the geographical region, the types of starter strains used (*L. lactis* subsp. *lactis* or *L. lactis* subsp. *cremoris*), and the criteria employed for strain selection.

Distinct from these common phage types, one of the least-studied species is P034 (family *Podoviridae*, Bradley C2 morphotype), which is characterized by a prolate head and short tail with prominent whiskers emanating from the head-tail interface region. Historically, this species has comprised on

* Corresponding author. Mailing address: Cultures Division, Dairy Innovation Australia, Ltd., Private Bag 16, Werribee, Victoria 3030, Australia. Phone: 61 (3) 9974 8910. Fax: 61 (3) 9742 4233. E-mail: ipowell@dairyinnovation.com.au.

† Present address: Orica Pty Limited, 1 Nicholson Street, Melbourne, Victoria 3000, Australia.

[∇] Published ahead of print on 4 April 2008.

average less than 1% of the phages isolated from dairy fermentations (1, 6). A "dot blot" DNA hybridization survey of Australian phage isolates using probes derived from reference phages detected only 3 P034 species phages, compared with 100 936 species phages, 34 P335 species phages (including the BK5-T and 1483 phage groups), and 13 c2 species phages. A further 49 phages were not classified unambiguously in that study (23).

In this study, we describe phage ascc ϕ 28 and its complete genome sequence and present evidence showing that this phage has genetic and functional similarities to the ϕ 29-like phages of *B. subtilis* (30) and phage Cp-1 of *Streptococcus pneumoniae* (29, 35). This is the first such analysis of an *L. lactis* phage belonging to the P034 species. In structure, genome sequence, and mode of replication this phage species is distinctly different from other phage species infecting *L. lactis*.

MATERIALS AND METHODS

Phages, lactococcal strains, and growth media. Phage ascc ϕ 28 was isolated from an Australian cheese factory whey sample, one of numerous whey samples that were surveyed during the period from 1993 to 1996 (23). *L. lactis* subsp. *lactis* ASCC99-319, a cheddar cheese starter culture component and host strain for phage ascc ϕ 28, was obtained from the Australian Starter Culture Research Centre (now the Cultures Division of Dairy Innovation Australia) collection. Phage P369 (20) was kindly provided by H.-W. Ackermann (Félix d'Hérelle Reference Center for Bacterial Viruses, Québec, Canada) as a representative of the P034 species. Phage P034 itself has apparently been lost. *L. lactis* strains were grown at 30°C in M17 broth (Oxoid, Basingstoke, Hampshire, United Kingdom) or on M17 agar supplemented with both 1% glucose and 10 mM CaCl₂ (M17gc) for growth of phages. Phage PFU enumeration was performed in M17gc double-layer agar plates at 30°C. The phage ascc ϕ 28 burst size and latent period were determined as described by Powell and Davidson (32).

Electron microscopy. Phage particles were purified by CsCl density gradient ultracentrifugation, stained with 0.5% uranyl acetate on a Formvar-coated copper grid (Proscitech, Brisbane, Queensland, Australia), and examined using a Hitachi H-300 electron microscope (Hitachi, Tokyo, Japan) calibrated using catalase crystals (Proscitech).

DNA methods. Phage DNA was prepared by solvent extraction and ethanol precipitation of a CsCl-purified phage preparation. Where indicated below, DNA was treated with proteinase K essentially as described by García et al. (12). The DNA was digested with restriction enzymes (Roche) used according to the manufacturer's instructions, electrophoresed on agarose gels, stained with ethidium bromide, and visualized using UV (38). Transfer of DNA fragments in agarose gels to nylon membranes (Hybond-N+; Amersham, Little Chalfont, Buckinghamshire, United Kingdom) for Southern blot (40) analysis was performed as described by Moineau et al. (31). Phage P369 DNA used to probe the blots was ³²P labeled using random priming.

DNA sequencing and computer analyses. Fragments (650 to 950 bp) resulting from partial AluI digestion of ascc ϕ 28 DNA were purified by agarose gel electrophoresis and cloned using the pGEM-T plasmid cloning system with *E. coli* SURE cells (Promega, Madison, WI). The ascc ϕ 28 genome sequence was determined by sequencing these cloned fragments, followed by designing specific primers for directed sequencing from ascc ϕ 28 genomic DNA. A Taq DyeDeoxy terminator cycle sequencing kit (Applied Biosystems, Scoresby, Victoria, Australia) and an Applied Biosystems 373A automated DNA sequencer were used. Sequences were assembled and analyzed using Sequencher 3.0 (Gene Codes, Ann Arbor, MI). The average sequence data redundancy was 4.5. DNA sequence data and deduced protein sequences were analyzed using the University of Wisconsin Genetics Computer Group suite of programs (now Accelrys GCG) and Clone software (version 9; Scientific and Educational Software, Cary, NC). Phage ascc ϕ 28 sequences were compared with sequences publicly available in databases using BLAST software (3) at the NCBI website (<http://www.ncbi.nlm.nih.gov>).

Protein analysis. CsCl-purified phage samples were boiled for 5 min in sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) loading buffer, and then 1 to 5 μ g total phage protein was electrophoresed through a polyacrylamide gel (24) and stained with Coomassie blue or Blu/T (B/T Scientific Technologies, Carlsbad, CA) used according to the manufacturer's instructions.

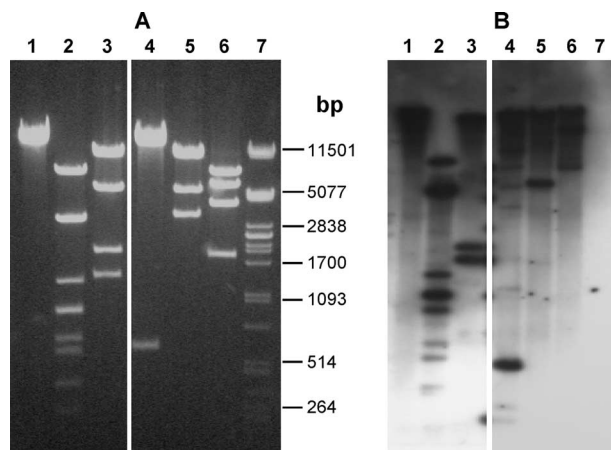


FIG. 1. (A) Restriction-digested ascc ϕ 28 DNA electrophoresed on a 1% agarose gel and stained with ethidium bromide. (B) Southern hybridization of the same fragments probed with DNA from phage P369. The lanes contained the following: lane 1, uncut DNA; lane 2, HindIII restriction enzyme digest; lane 3, AccI restriction enzyme digest; lane 4, ScaI restriction enzyme digest; lane 5, SpeI restriction enzyme digest; lane 6, AatII restriction enzyme digest; lane 7, marker coliphage λ DNA cut with PstI. DNA bands not obvious in the ethidium bromide-stained gel are visible in the Southern hybridization and presumably represent incomplete digestion products. Selected sizes of the marker are indicated between the panels.

For DNA terminal protein (TP) analysis, phage ascc ϕ 28 DNA (not treated with proteinase K) was digested with DNase I prior to boiling and electrophoresis (12). For N-terminal sequencing, proteins in SDS-PAGE gels (unstained) were electroblotted using a Bio-Rad transblot system (Bio-Rad, Hercules, CA) onto a polyvinylidene difluoride membrane (Bio-Rad) as described by Towbin et al. (42) and stained with Ponceau S (37). Bands of interest were excised from blots, destained in 200 μ M NaOH, and rinsed with water. N-terminal sequence analysis was carried out with an ABI Procise 494 protein sequencer using a standard blot program (18) at the Australian Proteome Analysis Facility (Sydney, Australia).

Nucleotide sequence accession number. The genome sequence of phage ascc ϕ 28 has been deposited in the GenBank database under accession number EU438902.

RESULTS

Species classification and physical characterization of ascc ϕ 28. Southern hybridization showed that DNA from phage P369 (P034 species) hybridized strongly to DNA from phage ascc ϕ 28 (Fig. 1). Hybridization was not uniform to all bands, presumably reflecting differences between the two phages. The level of genetic heterogeneity within the P034 phage species is not known. Electron microscopy (Fig. 2) showed that phage ascc ϕ 28 has a prolate capsid (head) (59 by 42 nm) and a short noncontractile tail (21 by 9 nm). A collar with fibrils is located near the head-tail interface, consistent with the conclusion that it is a member of the P034 species (20). SDS-PAGE of phage ascc ϕ 28 proteins (Fig. 3) resulted in a band pattern similar to that previously reported for phage P034 (6).

Virulence of ascc ϕ 28. The virulence of phage ascc ϕ 28 was assessed by performing a single-step growth curve analysis in M17gc broth using the host strain *L. lactis* subsp. *lactis* ASCC99-319 to determine the burst size and latent period. The burst size was found to be 121 ± 18 PFU per infected bacterial cell after a latent period of 44 min. Host range anal-

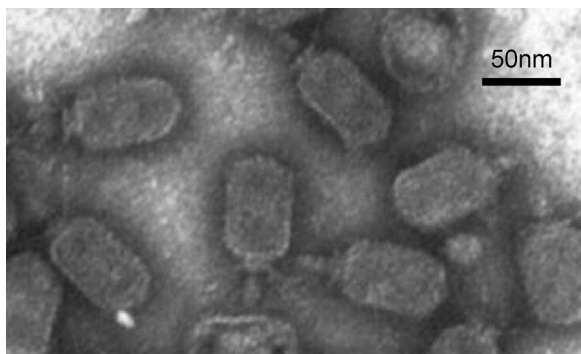


FIG. 2. Electron micrograph of uranyl acetate-stained preparation of CsCl-purified ascc ϕ 28. Bar = 50 nm.

ysis using a plaque assay showed that ascc ϕ 28 infected 2 of 64 dairy lactococcal strains tested, both of which are strains of *L. lactis* subsp. *lactis* (data not shown).

Key features of the ascc ϕ 28 genome. Sequence analysis showed a double-stranded DNA genome consisting of 18,762 bp. A similar genome size (approximately 18.1 kbp) has been estimated for phage P034 (20), suggesting that this small genome is typical of the P034 species. The G+C content of the ascc ϕ 28 genome is 33.7%. The DNA of phage ascc ϕ 28 does not have cohesive ends or terminal redundancy (features observed in other lactococcal phages), but inverted terminal repeat (ITR) sequences are present. A total of 28 open reading frames (ORFs) longer than 40 codons and with putative ribosomal binding sites were predicted (Table 1). Most of the stop codons are TAA. This is the smallest genome and lowest number of ORFs of the *L. lactis* phage genomes sequenced to date. The 28 ORFs comprise 90% of the total genome or just over 96% of the genome excluding the ITRs. Such genome compactness is common in *L. lactis* phages, which generally utilize 93 to 95% of their DNA as coding sequences (2). Predicted ORFs and transcriptional control signals inferred from the sequence are shown in Fig. 4. The ORFs occur on both strands, converging on bp 7709/7710. Many features of the ascc ϕ 28 genome and its ORFs (described below) show strong similarities to features of the genomes and ORFs of *S. pneumoniae* phage Cp-1 (29, 35) and the ϕ 29 family of phages that infect *B. subtilis*, including phage ϕ 29 itself and GA-1 (30).

Terminal repeats and DNA replication. A distinctive feature of the ascc ϕ 28 genome is the presence at both ends of 577-bp ITRs. A long ITR (236 bp) is present in phage Cp-1 (12, 29), whereas a short ITR (6 to 8 bp) is present in the ϕ 29-like phages (30). These phages require ITRs and a terminal replication protein (TP) that binds to their ITRs for DNA replication. Neither of these features has been observed in any other sequenced *L. lactis* phage, including the recently sequenced phage KSY1 belonging to the *Podoviridae* (8).

Phage ascc ϕ 28 was examined for a TP using methods similar to those described by García et al. (12). The terminal ScaI DNA fragments of phage ascc ϕ 28 were seen only after proteinase K treatment of a sample prior to loading on an agarose gel (Fig. 5A). A smear of retarded DNA was seen for the control (untreated) sample. This result is consistent with removal of a DNA-bound protein by proteinase treatment, en-

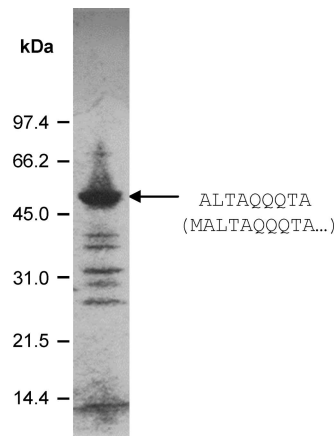


FIG. 3. SDS-PAGE of phage ascc ϕ 28 proteins electrophoresed on a 10 to 20% (top to bottom) polyacrylamide gradient gel. The N-terminal sequence of the dominant protein (arrow) was determined and is shown on the right. Underneath this sequence, the first 10 amino acids of the protein sequence predicted from ORF24 (encoding the MCP) are shown in parentheses.

abling the terminal ScaI fragments to migrate normally and be seen as a clear band in the gel. Similarly, SDS-PAGE of ascc ϕ 28 DNA showed a protein band (estimated molecular mass, 22 kDa) that appeared only after DNase I treatment of the sample (Fig. 5B), consistent with release of a DNA-bound protein by DNase I treatment. Matching the N-terminal sequence of the protein isolated from the gel (Fig. 5B) showed that it is encoded by ORF6.

By analogy with the *Bacillus* ϕ 29 phages and streptococcal phage Cp-1, these data suggest that phage ascc ϕ 28 replicates via a covalently linked replication protein attached to the termini of its genome and that the ITR of phage ascc ϕ 28 contains the origin of phage replication.

Functional assignment of ORFs. Where possible based on homologies detected by BLAST or other sequence analyses or on experimental data, functions were inferred for proteins predicted to be expressed from ORFs (Table 1). Genes with related functions appear to be grouped in discrete modules.

The predicted proteins encoded by ORF1 to ORF5, ORF8, and ORF9 (Fig. 4) do not show any significant BLAST homology with previously sequenced or predicted proteins in public databases. ORF6 encodes the TP involved in the replication of phage ascc ϕ 28 (Fig. 5). The phage DNA polymerase is encoded by ORF7 and shares significant amino acid homology with the protein-primed, DNA-dependent type B DNA polymerases of phage Cp-1 and the ϕ 29-like phages. The tandem arrangements of the DNA polymerase and TP genes in ascc ϕ 28, Cp-1, and the ϕ 29-like phages are identical (Fig. 4).

Kyte-Doolittle analysis of the predicted amino acid sequence encoded by ORF10 (not shown) indicated the presence of two transmembrane domains and a highly charged C-terminal region. These features are common in phage holins (26, 46). Along with the position of ORF10 upstream from the putative lysin gene (ORF12) (Table 1), this suggests that ORF10 encodes the holin protein of ascc ϕ 28. Holins of different *L. lactis* phages do not share any sequence similarity (26), and no sequence similarity was observed for the putative holin

TABLE 1. Characteristics, closest GenBank homologs, and predicted functions of ORFs in the ascϕ28 genome

ORF	Ribosomal binding site and start codon sequence ^a	Genome position ^b	Stop codon	Predicted protein			
				No. of amino acids	Mol wt (10 ³)	Closest homolog (GenBank accession no.; amino acid identity; total length) ^c	Predicted function ^d
ORF1	<u>GAAAGT</u> GAGAGAAAATG	614–871	TAA	85	10.3	NS	
ORF2	<u>AAAAAGGAGA</u> AATAAA GATG	861–1061	TAG	66	8.1	NS	
ORF3	<u>ATAGGAGATAAAATG</u>	1069–1281	TAA	70	8.1	NS	
ORF4	<u>GGAGAATAAATATG</u>	1283–1564	TAG	93	10.8	NS	
ORF5	<u>GAGGATATTATG</u>	1596–1742	TAA	48	6.0	NS	DNA binding? ^e
ORF6	<u>AAAAAGGATT</u> TATAACA AATG	1736–2239	TAA	167	19.7	NS	Terminal protein
ORF7	<u>GGAACGTT</u> CAGAAATG	2264–4432	TGA	722	85.8	<i>Bacillus</i> phage GA-1 DNA polymerase (NP_073685; 135/511 [26%]; 578 amino acids)	DNA polymerase
ORF8	<u>GAGGTCAA</u> ACTATG	4429–4869	TAA	146	17.3	NS	
ORF9	<u>AAGGAGTTT</u> AAAAATG	4871–5041	TAA	56	6.8	NS	
ORF10	<u>AGGAGAAAAA</u> ACATT TATG	5081–5467	TAA	128	14.3	NS	Holin
ORF11	<u>GGGGTGCT</u> GATG	5451–6551	TAA	366	43.5	<i>Streptococcus</i> phage Cp-1 encapsidation protein (NP_044835; 75/262 [28%]; 360 amino acids)	Encapsidation protein
ORF12	<u>AAAGGAGAAA</u> ATTATG	6563–7690	TAA	375	41.8	<i>Staphylococcus</i> phage PH15 hydrolase (YP_950686; 32/97 [32%]; 633 amino acids) ^f	Lysin
ORF13	<u>AGATAGAGATAGTTACT</u> CTATG	8203–7730	TAA	157	18.4	<i>Streptococcus</i> phage Cp-1 tail protein (NP_044833; 33/133 [24%]; 586 amino acids)	Tail
ORF14	<u>GAAAGGAA</u> ATTATG	9777–8089	TAA	562	62.6	<i>Bacillus</i> phage ϕ29 tail protein (NP_040727; 79/282 [28%]; 599 amino acids)	Tail
ORF15	<u>ATAGGAGGCTATGAGT</u> CATG	10527–9820	TAA	235	26.6	<i>Streptococcus</i> phage Cp-1 collar protein (NP_044824; 33/154 [21%]; 194 amino acids)	Structural
ORF16	<u>AGAAAGGAA</u> ATTTAA GATG	11282–10524	TGA	252	29.1	<i>L. lactis</i> phage bIL170 putative accessory fiber (AAR26444; 22/41 [53%]; 605 amino acids) ^g	Structural
ORF17	<u>GAAAGATATG</u>	12118–11297	TAG	273	30.2	<i>Streptococcus</i> phage Cp-1 hypothetical protein p18 (NP_044830; 57/222 [25%]; 253 amino acids) ^h	Structural
ORF18	<u>AAGGAGATTT</u> AAAAATG	13081–12128	TAG	317	35.2	<i>Burkholderia cenocepacia</i> phage tail protein (EAY61823; 31/88 [35%]; 1,301 amino acids)	Tail
ORF19	<u>AAGGAAATA</u> AAAAATG	13322–13176	TAG	48	5.5	NS	
ORF20	<u>GAGGAAGAT</u> TAAATG	14085–13093	TAA	330	36.1	<i>L. lactis</i> phage KSY1 protein gp73 (ABG21616; 99/153 [64%]; 347 amino acids) ⁱ	Tail
ORF21	<u>AAGAAATGATG</u>	14856–14086	TAA	256	29.0	<i>Streptococcus</i> phage Cp-1 connector protein (NP_044823; 60/217 [27%]; 337 amino acids)	Upper collar or connector
ORF22	<u>AAAGGGTTT</u> GTATG	15056–14856	TGA	66	8.0	NS	
ORF23	<u>GAAAGTAGATG</u>	15594–15244	TAA	116	14.1	NS	
ORF24	<u>AAGGAGAA</u> TAAAAATG	16570–15110	TAG	486	53.1	<i>L. lactis</i> subsp. <i>cremoris</i> MG1363 cell wall surface anchor family protein (YP_001032441; 38/85 [44%]; 1,349 amino acids) ^j	Major capsid protein
ORF25	<u>AGGAGTAA</u> ATAATG	16914–16573	TAA	113	12.8	NS	
ORF26	<u>AAAGGAAT</u> TAAAAAT TATG	17082–16918	TAA	54	6.3	NS	
ORF27	<u>AAAGGAAA</u> ACACAAA AATG	17682–17191	TAA	163	18.6	<i>Bacillus</i> phage GA-1 SSB (NP_073688; 53/163 [32%]; 170 amino acids)	SSB
ORF28	<u>ATAAAGACAG</u> CAAATG	18168–17896	TGA	90	10.8	NS	

^a The sequence complementary to the 3' end of the *L. lactis* 16S rRNA (43) is underlined, and the start codon is indicated by bold type.

^b ORF13 to ORF28 (inclusive) are on the complementary strand.

^c The amino acid identity for the region shown and the number of amino acids (total length) for each homolog are indicated. NS, no significant homology found.

^d Based on a BLAST search of databases or other experimental data described in the text.

^e The ORF5 predicted protein contains a zinc finger motif which suggests a DNA binding function (27).

^f The highest sequence similarity occurs in the N-terminal part of the predicted protein, the region likely to be responsible for lytic activity (10, 13, 39).

^g There was also homology to structural or tail antireceptor (host-binding) proteins in other *L. lactis* phages, including bIL170 (GenBank accession no. AAR26444), bIL41 (AAB41477), rIt (NP_695075), TP901-1 (NP_112714), and 4268 (NP_839936), as well as to *Streptococcus thermophilus* phage antireceptor proteins. This homology occurred in the same region consisting of up to 41 amino acids centrally located in the ORF16 predicted protein.

^h The gene encoding hypothetical protein p18 in phage Cp-1 is located between two other phage tail genes.

ⁱ There was also homology to the lysins (amidases) of *Lactobacillus casei* phages A2 (GenBank accession no. NP_680500) and PL-1 (BAA96749) (41% identity for 299 of 350 amino acids for both proteins).

^j There was also homology to the major capsid or other major structural proteins of various *L. lactis* phages, including F4-1 (GenBank accession no. P26596), p2 (AAF85633), sk1 (NP_044957), and bIL170 (NP_047126) (37 to 44% identity for 81 to 83 amino acids located at the C-terminal end of the predicted protein encoded by ORF24), and to the major capsid protein of *Streptococcus* phage Cp-1 (NP_044821) (20% identity for 267 of 360 amino acids).

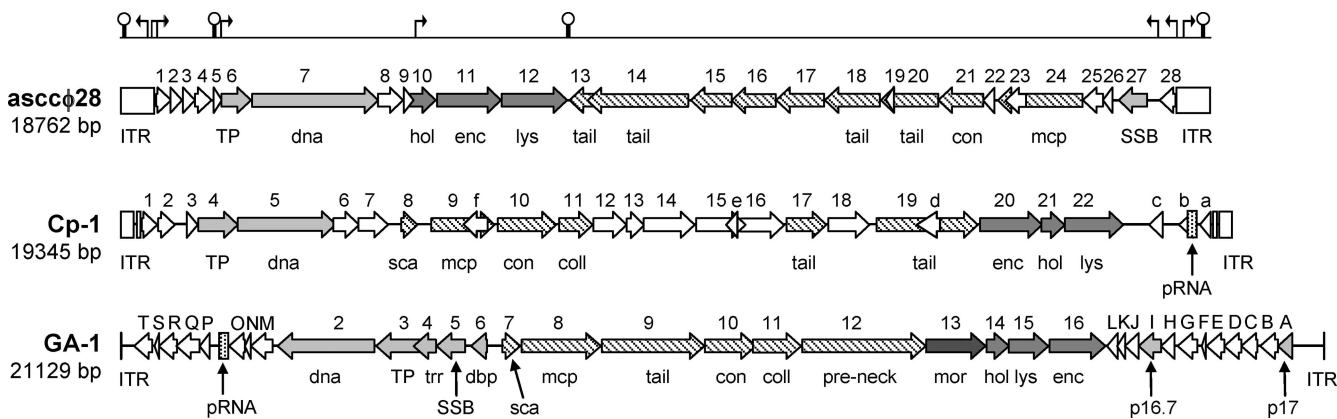


FIG. 4. Deduced genetic map of phage asccφ28 compared with maps for phages Cp-1 (29) and GA-1, a φ29-like phage (30). The maps for Cp-1 and GA-1 are derived from data in the GenBank database (accession numbers Z47794 and X96987, respectively). Gene functions are abbreviated as follows: TP, terminal protein; dna, DNA polymerase; hol, holin; enc, encapsidation protein; lys, lysin; tail, tail protein; coll, collar protein; mcp, major capsid (head) protein; con, collar or connector protein; SSB, single-stranded DNA-binding protein; sca, scaffolding protein; trr, transcriptional regulator; dbp, DNA-binding protein; pre-neck, preneck appendage protein; mor, morphogenesis protein. The genes encoding DNA replication proteins p16.7 and p17 in GA-1 are indicated. Light gray shading indicates genes involved in DNA replication or transcriptional regulation; medium gray shading indicates genes involved in cell lysis; diagonal lines indicate genes encoding structural proteins; and dark gray shading in the GA-1 map indicates the gene encoding the morphogenesis protein. Proteins with no known function are indicated by open arrows. Packaging RNAs and inverted terminal repeat regions are indicated by pRNA and ITR, respectively. Above the asccφ28 map, putative promoters and rho-independent transcriptional terminators (see the GenBank accession number EU438902 sequence for details) are indicated by arrows and hairpin structures, respectively.

of asccφ28 and any previously sequenced phage holin. ORF11 encodes a predicted protein homologous to the encapsidation (DNA packaging) protein of phage Cp-1 and the φ29-like phages. This protein functions to encapsidate phage DNA into the mature phage capsid by an ATP-driven process at the end of the phage replicative cycle (16, 30).

ORF13 to ORF28 are on the complementary DNA strand

(Fig. 4) and are discussed below in the presumed order of gene transcription. ORF28 encodes a predicted protein of unknown function. The ORF27 predicted protein has homology to a single-stranded DNA-binding protein (SSB) of *Bacillus* phage GA-1 (15). ORF26 encodes a small protein containing seven direct repeats of the tetrapeptide sequence VDEK or VNEK in its C-terminal region (amino acids 22 to 54), plus an imperfect eighth repeat at the C terminus. The significance of this is unknown.

Most of the remaining ORFs in asccφ28 appear to be structural genes. The predicted protein encoded by ORF25 is responsible for serological cross-reactivity between phages c2 and asccφ28 (22). However, it showed no BLAST homology to proteins from phage c2 or to any other proteins. ORF24 putatively encodes the major capsid protein (MCP), and its product corresponds to the major protein band shown in Fig. 3. The predicted asccφ28 MCP (53.1 kDa) is almost exactly the same size as that of *B. subtilis* phage GA-1 (53.0 kDa), but it is larger than that of phage Cp-1 (42 kDa). N-terminal sequence data (Fig. 3) indicate that there is no posttranslational cleavage, in contrast to the MCP of Cp-1 (29). The MCP gene in phage Cp-1 is followed downstream by a gene encoding an upper collar (or connector) protein (29). Genes encoding similar proteins occur downstream of the MCP gene in the φ29-like phages (30). In phage asccφ28 this gene order is conserved, since the predicted protein encoded by ORF21 shows homology to the upper collar proteins from both phage Cp-1 and the φ29-like phages. With the exception of ORF19 (no BLAST homology was found for the predicted protein), ORF20 to ORF13 also appear to encode structural proteins (Table 1). ORF20 encodes a predicted protein with significant homology to predicted protein gp73 in *L. lactis* phage KSY1 (8) and to *Lactobacillus* phage amidases (Table 1).

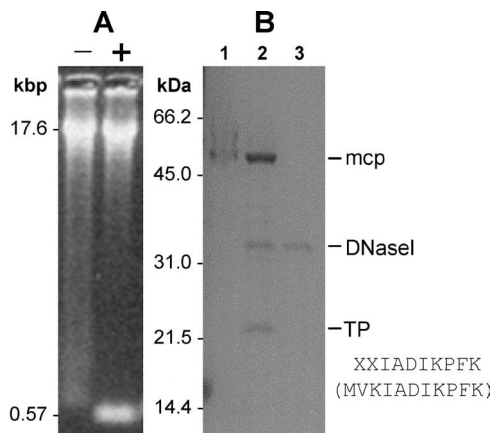


FIG. 5. TP detection. (A) Phage asccφ28 DNA was restriction digested with ScaI (each ITR has a ScaI site close to its inner end), resulting in two 573-bp terminal DNA fragments. Digested DNA was treated (lane +) or not treated (lane -) with proteinase K and then electrophoresed on a 1.2% agarose gel. (B) Phage asccφ28 DNA (not treated with proteinase K) was electrophoresed on a 12% polyacrylamide gel. Lane 1, asccφ28 DNA carrying the bound TP; lane 2, asccφ28 DNA digested with 1 U of DNase I, which released the TP; lane 3, 1 U of DNase I (control). Lanes 1 and 2 also show that there was residual MCP in the DNA preparation. The N-terminal sequence of the TP was determined and is shown along with the first 10 amino acids of the protein predicted from ORF6 (in parentheses).

Transcription. Putative transcript initiation and termination signals in the ascc ϕ 28 genome (as detailed in the GenBank accession no. EU438902 sequence) are shown in Fig. 4. No transcript mapping or analysis of temporal control of gene expression was done in this study, so any transcript model remains speculative.

DISCUSSION

In this study we present the first complete genome sequence of an *L. lactis* bacteriophage belonging to the P034 species of phages (family *Podoviridae*). Phages belonging to this group have a C2 morphotype encountered only rarely in the dairy industry. The reason for this rarity is not clear. Phage ascc ϕ 28 has a narrow host range, but narrow host ranges have been observed for many lactococcal phages (data not shown). Its virulence (measured by latent period and burst size) is comparable to that of other lactococcal phages (21).

Morphologically and in gene arrangement, mode of replication, and protein sequence homology, phage ascc ϕ 28 more closely resembles phage Cp-1 that infects *S. pneumoniae* (29) and the ϕ 29 family of phages that infect *B. subtilis* (30) than it resembles other sequenced *L. lactis* phages. Phage Cp-1 and the ϕ 29-like phage GA-1 are the phages most similar to ascc ϕ 28, as determined by homology at the protein level.

Phage ascc ϕ 28 presumably evolved from a common ancestor of the *Streptococcus* Cp-1 and *Bacillus* ϕ 29-like phages, but it is not clear at what point in its evolution it acquired the ability to infect *L. lactis*. In the G+C content of its DNA, phage ascc ϕ 28 more closely resembles *L. lactis* than either *S. pneumoniae* or *B. subtilis*. However, only three ORFs (ORF16, ORF20, and ORF24) in ascc ϕ 28 encode proteins with recognizable homology to known proteins from *L. lactis* and its phages (Table 1), presumably due to recombination that occurred at some stage during evolution. Some of the homologies may be associated with host specificity. Two of these ORFs (ORF16 and ORF20) encode tail proteins that could be directly involved in phage-host surface interaction.

Further evidence for the relatedness of ascc ϕ 28 to phage Cp-1 and the ϕ 29-like phages is in the mode of DNA replication. All these phages have a phage-encoded DNA polymerase, TP, and ITRs, indicating that they replicate in similar ways (not previously documented for an *L. lactis* phage). The known mechanism of DNA replication for phage Cp-1 and the ϕ 29-like phages requires formation of a stable complex between the TP, the phage-encoded DNA polymerase, and the phage DNA ITR regions, with the TP itself initially priming the DNA polymerase. During replication no multiple-length DNA concatemers are formed (they are formed in other lactococcal phages), and single-stranded DNA is generated (15, 17, 29, 30, 36). Consistent with this, the ascc ϕ 28 genome appears to encode an SSB protein (Table 1). The predicted DNA polymerase of ascc ϕ 28 is large (722 amino acids) compared with the corresponding proteins of phages Cp-1 (568 amino acids), GA-1 (578 amino acids), and ϕ 29 (575 amino acids). A ClustalW alignment (not shown) of these DNA polymerases revealed significant differences for the ascc ϕ 28 DNA polymerase in the nine semi-conserved domains described by Meier et al. (30), although the three aspartic acid residues needed for catalysis at the active site remain conserved.

Critical components of the active process of packaging DNA into mature phage heads in the Cp-1 and ϕ 29-like phages include a phage-encoded packaging (or prohead) RNA (pRNA) that is transcribed from a specific phage promoter, as well as an encapsidation protein (4, 30). The pRNAs have been described for Cp-1 and ϕ 29-like phages (4, 17, 29, 30) and are 160 to 180 nucleotides long. Although there is little sequence homology between these pRNAs, their secondary structures are remarkably similar (4, 29). Noncoding regions of the ascc ϕ 28 genome were examined for potential transcripts exhibiting secondary structure consistent with the properties of these pRNAs, but none was found. The location of the presumed pRNA gene of ascc ϕ 28 remains unknown. However, BLAST homologies indicate the presence of an encapsidation protein in ascc ϕ 28, encoded by ORF11 (Table 1).

A comparison of complete gene maps for phages ascc ϕ 28, Cp-1, and GA-1 (Fig. 4) shows that genes with similar functions are arranged in modules (operon-like structures) and that the order of these modules in the genome varies, consistent with the modular evolution of phages (5, 33). Despite rearrangements that change the gene order, the essential nature of each functional module is preserved. Considering likely transcripts and ORF functions (Table 1 and Fig. 4), four discrete types of modules can be hypothesized for the ascc ϕ 28 genome: (i) ITR regions, (ii) DNA replication modules, (iii) a structural protein module, and (iv) a cell lysis module. It is suggested that, as in phage Cp-1 (29) and the ϕ 29-like phages (30), the outer regions of the ascc ϕ 28 genome encode proteins produced early in the phage life cycle, while the center region of the ascc ϕ 28 genome encodes proteins needed later in the life cycle (Fig. 4). The deduced promoter and terminator positions are consistent with this modular arrangement and imply that there is modular control of gene expression.

The hypothesized "early" modules located near the ends of the ascc ϕ 28 genome contain ORFs encoding proteins involved DNA replication (TP, DNA polymerase, SSB). However, several predicted proteins encoded in these regions show no recognizable BLAST homologies (Table 1), making functional assignments on this basis impossible. Presumably, some early genes are involved in interactions with host elements or in the regulation of transcription or translation. Curiously, the functions of many early genes of other *L. lactis* phages also cannot be ascertained by sequence homology (7, 26).

The structural protein genes in ascc ϕ 28, comprising mostly the MCP, collar, and tail protein genes, are contiguous in the genome (Fig. 4). Unlike the MCP of Cp-1, but similar to that of ϕ 29, the MCP in ascc ϕ 28 is not subject to N-terminal post-translational cleavage during the packaging process. Coding regions for two of the putative tail proteins (ORF13 and ORF14) overlap extensively, suggesting that translation of these two ORFs may be regulated or coupled in some way. This has been reported for assembly of tail genes in a number of other phages (25, 41).

The cell lysis module comprises the holin (ORF10), encapsidation protein (ORF11), and lysin (ORF12) genes. The gene order within this module is different in phages Cp-1 and the ϕ 29-like phages (Fig. 4). Holins mediate the semispecific transport across the plasma membrane of the lysin and, when expressed, lead to rapid death of the producing cell (44, 45). ORF10 was not represented in our plasmid library, consistent

with lethality of the ORF10 protein. Production of holins and lysins must be tightly controlled, since premature production is deleterious to the efficiency of phage replication. In most phages a holin-lysin cassette is found near the end of a late RNA transcript, since these genes are not expressed until they are needed late in the phage infection cycle (45).

In contrast to ORF12 (which putatively encodes the phage asccφ28 endolysin), the similarity of the ORF20 predicted protein to lactococcal phage KSY1 protein gp73 and to *Lactobacillus* phage amidases (Table 1) occurs toward the C-terminal end. For most lysins/autolysins this region controls substrate binding specificity (10, 13, 39). A possible explanation is that asccφ28 has a protein in its tail structure to facilitate binding to the cell wall, followed by DNA injection, like in some other phages. Phage KSY1 (8) also has two putative endolysins, the gp53a and gp73 predicted proteins (the latter shows similarity to the ORF20 predicted protein of asccφ28). However, for both of these lysins in phage KSY1, putative holin genes are immediately upstream, suggesting that they function from within the host cell.

In conclusion, the complete 18,762-bp genome sequence of phage asccφ28 belonging to a rare *L. lactis* phage species (P034) has been determined. Phage asccφ28 resembles the φ29-like phages of *B. subtilis* and phage Cp-1 of *S. pneumoniae* more closely than it resembles other *L. lactis* phage species. Sequence analysis revealed 28 ORFs, 13 of which encode predicted proteins with observable BLAST similarity to other proteins or predicted proteins in the GenBank database. The relative rarity and distinctive characteristics of this phage type raise the question of whether future adaptations might increase the infectivity, potential host range, or prevalence of such phages in dairy industry fermentations.

ACKNOWLEDGMENTS

This work was supported by the Australian Research Council through an Australian Postgraduate Award (Industry) to S.E.K., by the Dairy Research and Development Corporation (now Dairy Australia), and by member companies of the Australian Starter Culture Research Centre (now the Cultures Division of Dairy Innovation Australia).

We are grateful for the assistance of Michelle Taylor and Andrea Merrall with whey sample testing that led to the isolation of phage asccφ28, with phage stock maintenance, and with host range analysis.

REFERENCES

- Ackermann, H. W. 2001. Frequency of morphological phage descriptions in the year 2000. *Arch. Virol.* **146**:843–857.
- Allison, G. E., and T. R. Klaenhammer. 1998. Phage resistance mechanisms in lactic acid bacteria. *Int. Dairy J.* **8**:207–226.
- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**:403–410.
- Anderson, D. L., and B. E. Reilly. 1993. Morphogenesis of bacteriophage φ29, p. 859–867. In A. L. Sonenshien, J. A. Hoch, and R. Losick (ed.), *Bacillus subtilis* and other gram-positive bacteria: biochemistry, physiology, and molecular genetics. American Society for Microbiology, Washington, DC.
- Botstein, D. 1980. A theory of modular evolution for bacteriophages. *Ann. N. Y. Acad. Sci.* **354**:484–491.
- Braun, V., S. Hertwig, H. Neve, A. Geis, and M. Teuber. 1989. Taxonomic differentiation of bacteriophages of *Lactococcus lactis* by electron microscopy, DNA-DNA hybridization, and protein profiles. *J. Gen. Microbiol.* **135**:2551–2560.
- Chandry, P. S., S. C. Moore, J. D. Boyce, B. E. Davidson, and A. J. Hillier. 1997. Analysis of the DNA sequence, gene expression, origin of replication and modular structure of the *Lactococcus lactis* lytic bacteriophage ski1. *Mol. Microbiol.* **26**:49–64.
- Chopin, A., H. Deveau, S. D. Ehrlich, S. Moineau, and M.-C. Chopin. 2007. KSY1, a lactococcal phage with a T7-like transcription. *Virology* **365**:1–9.
- Deveau, H., S. J. Labrie, M.-C. Chopin, and S. Moineau. 2006. Biodiversity and classification of lactococcal phages. *Appl. Environ. Microbiol.* **72**:4338–4346.
- Fischetti, V. A. 2004. The use of phage lytic enzymes to control bacterial infections, p. 321–334. In E. Kutter and A. Sulakvelidze (ed.), *Bacteriophages—biology and applications*. CRC Press, Boca Raton, FL.
- Fortier, L.-C., A. Bransi, and S. Moineau. 2006. Genome sequencing and global gene expression of Q54, a new phage species linking the 936 and c2 phage species of *Lactococcus lactis*. *J. Bacteriol.* **188**:6101–6114.
- García, E., A. Gómez, C. Ronda, C. Escarmis, and R. López. 1983. Pneumococcal bacteriophage Cp-1 contains a protein bound to the 5' termini of its DNA. *Virology* **128**:92–104.
- García, P., J. L. García, R. López, and E. García. 2005. Pneumococcal phages, p. 335–361. In M. K. Waldor, D. I. Friedman, and S. L. Adhya (ed.), *Phages: their role in bacterial pathogenesis and biotechnology*. ASM Press, Washington, DC.
- Garneau, J. E., D. M. Tremblay, and S. Moineau. 2008. Characterization of 1706, a virulent phage from *Lactococcus lactis* with similarities to prophages from other Firmicutes. *Virology* **373**:298–309.
- Gascón, I., J. M. Lázaro, and M. Salas. 2000. Differential functional behavior of viral φ29, Nf and GA-1 SSB proteins. *Nucleic Acids Res.* **28**:2034–2042.
- Grimes, S., and D. Anderson. 1997. The bacteriophage φ29 packaging proteins supercoil the DNA ends. *J. Mol. Biol.* **266**:901–914.
- Guo, P., S. Erickson, and D. Anderson. 1987. A small viral RNA is required for *in vitro* packaging of bacteriophage φ29 DNA. *Science* **236**:690–694.
- Hewick, R. M., M. W. Hunkapiller, L. E. Hood, and W. J. Dreyer. 1981. A gas-liquid solid phase peptide and protein sequencer. *J. Biol. Chem.* **256**:7990–7997.
- Jarvis, A. W. 1984. Differentiation of lactic streptococcal phages into phage species by DNA-DNA homology. *Appl. Environ. Microbiol.* **47**:343–349.
- Jarvis, A. W., G. F. Fitzgerald, M. Mata, A. Mercenier, H. Neve, I. B. Powell, C. Ronda, M. Saxelin, and M. Teuber. 1991. Species and type phages of lactococcal bacteriophages. *Intervirology* **32**:2–9.
- Klaenhammer, T. R., and G. F. Fitzgerald. 1994. Bacteriophages and bacteriophage resistance, p. 106–168. In M. J. Gasson and W. M. de Vos (ed.), *Genetics and biotechnology of lactic acid bacteria*. Blackie Academic and Professional, London, United Kingdom.
- Kotsonis, S. E. 1999. Detection and characterization of bacteriophages infecting *Lactococcus lactis*. Ph.D. thesis. University of Melbourne, Melbourne, Australia.
- Kotsonis, S. E., I. B. Powell, M. L. Billingham, A. J. Hillier, G. K. Y. Limsowit, and B. E. Davidson. 1998. Differentiation of 189 bacteriophages infecting *Lactococcus lactis*. *Aust. J. Dairy Technol.* **53**:120.
- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of bacteriophage T4. *Nature* **227**:680.
- Levin, M. E., R. W. Hendrix, and S. R. Casjens. 1993. A programmed translational frameshift is required for the synthesis of a bacteriophage λ tail assembly protein. *J. Mol. Biol.* **234**:124–139.
- Lubbers, M. W., N. R. Waterfield, T. P. J. Beresford, R. W. F. Le Page, and A. W. Jarvis. 1995. Sequencing and analysis of the prolate-headed lactococcal bacteriophage c2 genome and identification of the structural genes. *Appl. Environ. Microbiol.* **61**:4348–4356.
- MacKay, J. P., and M. Crossley. 1998. Zinc fingers are sticking together. *Trends Biochem. Sci.* **23**:1–4.
- Mahony, J., H. Deveau, S. McGrath, M. Ventura, C. Canchaya, S. Moineau, G. F. Fitzgerald, and D. van Sinderen. 2006. Sequence and comparative genomic analysis of lactococcal bacteriophages jjs50, 712 and P008: evolutionary insights into the 936 phage species. *FEMS Microbiol. Lett.* **261**:253–261.
- Martín, A. C., R. López, and P. García. 1996. Analysis of the complete nucleotide sequence and functional organization of the genome of *Streptococcus pneumoniae* bacteriophage Cp-1. *J. Virol.* **70**:3678–3687.
- Meijer, W. J. J., J. A. Horcajadas, and M. Salas. 2001. φ29 family of phages. *Microbiol. Mol. Biol. Rev.* **65**:261–287.
- Moineau, S., J. Fortier, H. W. Ackermann, and S. Pandian. 1992. Characterization of lactococcal bacteriophages from Quebec cheese plants. *Can. J. Microbiol.* **38**:875–882.
- Powell, I. B., and B. E. Davidson. 1985. Characterization of streptococcal bacteriophage c6A. *J. Gen. Virol.* **66**:2737–2741.
- Proux, C., D. van Sinderen, J. Suarez, P. Garcia, V. Ladero, G. F. Fitzgerald, F. Desiere, and H. Brüßow. 2002. The dilemma of phage taxonomy illustrated by comparative genomics of Sh21-like *Siphoviridae* in lactic acid bacteria. *J. Bacteriol.* **184**:6026–6036.
- Relano, P., M. Mata, M. Bonneau, and P. Ritzenthaler. 1987. Molecular characterization and comparison of 38 virulent and temperate bacteriophages of *Streptococcus lactis*. *J. Gen. Microbiol.* **133**:3053–3063.
- Ronda, C., R. López, and E. García. 1981. Isolation and characterization of a new bacteriophage, Cp-1, infecting *Streptococcus pneumoniae*. *J. Virol.* **40**:551–559.
- Salas, M., and F. Rojo. 1993. Replication and transcription of bacteriophage φ29 DNA, p. 843–857. In A. L. Sonenshien, J. A. Hoch, and R. Losick (ed.), *Bacillus subtilis* and other gram-positive bacteria: biochem-

- istry, physiology and molecular genetics. American Society for Microbiology, Washington, DC.
37. **Salinovich, O., and R. C. Montelaro.** 1986. Reversible staining and peptide mapping of proteins transferred to nitrocellulose after separation by SDS-PAGE. *Anal. Biochem.* **156**:341–347.
 38. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
 39. **Sheehan, M. M., J. L. García, R. López, and P. García.** 1997. The lytic enzyme of the pneumococcal phage Dp-1: a chimeric lysin of intergeneric origin. *Mol. Microbiol.* **25**:717–725.
 40. **Southern, E. M.** 1975. Detection of specific binding sequence among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**:503–517.
 41. **Torgov, M. Y., D. M. Janzen, and M. K. Reddy.** 1998. Efficiency and frequency of translational coupling between the bacteriophage T4 clamp loader genes. *J. Bacteriol.* **180**:4339–4343.
 42. **Towbin, H., T. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. USA* **76**:4350–4354.
 43. **van de Guchte, M., J. Kok, and G. Venema.** 1992. Gene expression in *Lactococcus lactis*. *FEMS Microbiol. Rev.* **88**:73–92.
 44. **Wilson, D. B.** 1982. Effect of the lambda S gene product on properties of the *Escherichia coli* inner membrane. *J. Bacteriol.* **151**:1403–1410.
 45. **Young, R.** 1992. Bacteriophage lysis: mechanism and regulation. *Microbiol. Rev.* **56**:430–481.
 46. **Young, R., and U. Bläsi.** 1995. Holins: form and function in bacteriophage lysis. *FEMS Microbiol. Rev.* **17**:191–205.