Available online at www.sciencedirect.com

SCIENCE DIRECT

Virology 339 (2005) 42 – 55

VIROLOGY

www.elsevier.com/locate/yviro

Complete genomic sequence of the temperate bacteriophage ΦAT3 isolated from *Lactobacillus casei* ATCC 393

Ta-Chun Lo, Tsung-Chieh Shih, Chao-Fen Lin, Hung-Wen Chen, Thy-Hou Lin*

Institute of Molecular Medicine and Department of Life Science, National Tsing Hua University, Hsinchu 30043, Taiwan, ROC

Received 21 November 2004; returned to author for revision 6 May 2005; accepted 16 May 2005
Available online 21 June 2005

Abstract

The complete genomic sequence of a temperate bacteriophage ΦAT3 isolated from *Lactobacillus* (*Lb.*) *casei* ATCC 393 is reported. The phage consists of a linear DNA genome of 39,166 bp, an isometric head of 53 nm in diameter, and a flexible, noncontractile tail of approximately 200 nm in length. The number of potential open reading frames on the phage genome is 53. There are 15 unpaired nucleotides at both 5′ ends of the ΦAT3 genome, indicating that the phage uses a *cos*-site for DNA packaging. The ΦAT3 genome was grouped into five distinct functional clusters: DNA packaging, morphogenesis, lysis, lysogenic/lytic switch, and replication. The amino acid sequences at the NH2-termini of some major proteins were determined. An in vivo integration assay for the ΦAT3 integrase (Int) protein in several lactobacilli was conducted by constructing an integration vector including ΦAT3 *int* and the *att*P (int-*att*P) region. It was found that ΦAT3 integrated at the $tRNA^{Arg}$ gene locus of *Lactobacillus rhamnosus* HN 001, similar to that observed in its native host, *Lb. casei* ATCC 393.

Keywords: Genomic sequence; Bacteriophage; *Lactobacillus casei*; Integration vector

Introduction

Lactic acid bacteria (LAB) are Gram-positive bacteria widely used in the industrial fermentations. A serious problem encountered in the fermentation industry when using LAB as starter cultures is attack by bacteriophages. Cultures of LAB are usually prepared as a mixture of several different microorganisms in the fermentation process. This may lead to intergenomic rearrangements and mutual infections from bacteriophages resulting in genetic polymorphisms and DNA dynamics (Josephsen and Neve, 1998). Consequently, dairy bacteriophages usually share close DNA sequence homology in their genetic modules (Brüssow, 2001). Most of the bacteriophages isolated from lactobacilli thus far are classified as members of the Siphoviridae family, phages with long and noncontractile tails (Ackermann and DuBow, 1987). Differences in their DNA packaging mechanisms allow further division into *pac*-site or *cos*-site phages (Le Marrec et al., 1997). Numerous phages isolated from LAB are now grouped into two newly proposed genera, Sfi11-like or Sfi21-like (Lucchini et al., 1999; Brüssow and Desiere, 2001). While Sfi11-like phages use *pac*-sites, Sfi21-like phages use *cos*-sites. Comparing the genome sequences of eight Sfi21-like Siphoviridae that infect five distinct genera of low GC content Gram-positive bacteria reveals the relatedness of the phages with their host bacteria, suggesting that co-evolution has occurred. With 31–61% amino acid (aa) sequence identity over the DNA packaging, head and tail morphogenesis genes, *Lactobacillus gasseri* phage Φadh ranks third in its relatedness to phage Sfi21 of the reference strain *Streptococcus thermophilus* (Brüssow and Desiere, 2001). Phage A2 of *Lactobacillus casei* (Herrero et al., 1994), which shares the same organization of replication and lysogeny modules with those of *St. thermophilus* phage Sfi21, is recognized as an additional member of the group

* Corresponding author.
E-mail address: thlin@life.nthu.edu.tw (T.-H. Lin).

0042-6822/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2005.05.022

(Proux et al., 2002). The two closest relatives of *St. thermophilus* phage Sfi11 are phig1e and LL-H of *Lactobacillus plantarum* and *Lactobacillus delbrueckii* subsp. *lactis*, respectively, which share up to 58% aa sequence identity over the entire DNA packaging, head and tail modules. Sfi11-like and Sfi21-like phages differ by having two rather than one major head proteins, lacking of proteolytic processing of the major head protein, and the presence of a scaffolding protein (Brüssow and Desiere, 2001). However, genomic similarities over the structural genes of both LAB phage groups have been linked to those of the phages of other branch of Gram-positive bacteria and even to some coliphages (Desiere et al., 1998; Brüssow and Desiere, 2001).

In this report, we present the complete nucleotide sequence of a temperate bacteriophage ΦAT3, which was induced from *Lb. casei* ATCC 393 with 0.2 μg/ml mitomycin-C. ΦAT3 has a linear genome of 39,166 bp, an isometric head, and a noncontractile tail of 200 nm in length. Fifty-three open reading frames (ORFs) were identified on the phage genome. We compared the nucleotide sequences of *att*B, *att*L, *att*R, and *att*P regions and identified a 15-bp core sequence; the attachment site *att*B is located in a putative tRNA^{Arg} gene. We compare sequences of the phage major structural proteins, immunity system, integration protein, and the putative origin of replication with those of known phages. The activity of ΦAT3 Int is also assayed using an integration vector that is based on the sequences of ΦAT3 *int*, *att*P region, and a potential *int* promoter (P_{IN}).

Results and discussion

The complete genomic sequence of ΦAT3 has been determined. The functions of most predicted ORFs are unknown although they share sequence homology with comparable ORFs of *Lb. casei* phage A2 within the replication and downstream regions. The morphology of temperate bacteriophage ΦAT3, isolated from *Lb. casei* ATCC 393, is very similar to that of other Sfi21-like phages. ΦAT3 belongs to the Siphoviridae family (Canchaya et al., 2003); the electron micrograph of the phage reveals an isometric head of 53 nm in diameter and a noncontractile tail of approximately 200 nm in length. A tail fiber from the baseplate can also be seen (Fig. 1A). Some extremely long tails, polytails, were also observed (Fig. 1B), these are not

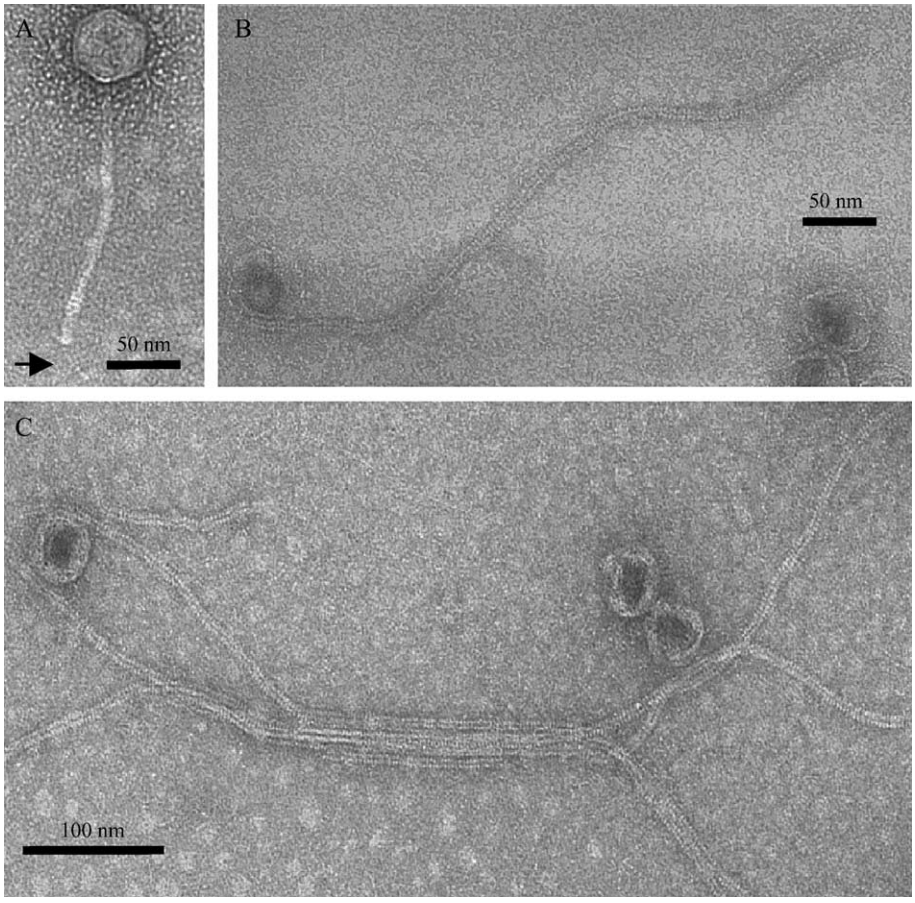

Fig. 1. (A) Electron micrograph of ΦAT3. The phage has an isometric head of 53 nm diameter and a noncontractile tail of approximately 200 nm. A tail fiber extruded from the baseplate is highlighted with an arrow. (B and C) The two electron micrographs, also showing some extra long tails, are associated with empty or defective heads.

uniform in size and their length can extend more than 1 μm (Fig. 1C). These extra long tails may be abnormal ΦAT3 tails, tails of an unrelated induced prophage, or perhaps bacterial flagella or pili.

After equilibrium centrifugation in CsCl, ΦAT3 virion proteins were separated by a 12% SDS–PAGE and at least eight protein species were detected (Fig. 2). Their sizes range from 25 to 100 kDa with a dominant species of 31 kDa. The N-termini of some proteins were isolated from the gel and sequenced. The data were compared with those deduced from the ΦAT3 genome sequence. Five proteins were annotated as parts of the capsid or tail virions. Truncations of 2, 911, 195, and 476 residues from the N-termini of four tail proteins encoded, respectively, by ORF11, ORF13, ORF14, and ORF15 were apparent (Fig. 2, bands 1, 3, 4 and 6). In addition, the major head protein (ORF6), which is predicted by the genome sequence to contain 394 residues, is truncated by 118 residues. Bands 5, 7, and 8 were corresponded to proteins poorly transferred to the membrane and were not analyzed (Fig. 2). The cleavage sites of the portal and major head protein precursors have the common sequence Pol–Pol–Glu↓ (where Pol stands for the polar amino acids and "↓" stands for the cut). This processing is comparable to T4, where a phage-encoded protease recognizes Pho–Pho–Glu↓ (where Pho stands for hydrophobic amino acids) (Black and Showe, 1983). ORF13 is recognized as a functional equivalent of coliphage lambda gpH, the tape measure protein. Post-translational processing removes the C-terminal portion of gpH (Tsui and Hendrix, 1983; Katsura, 1990; Katsura and Hendrix, 1984); similarly, the C-terminal one-third of ΦAT3 ORF13 is cleaved and remains present in the mature particle.

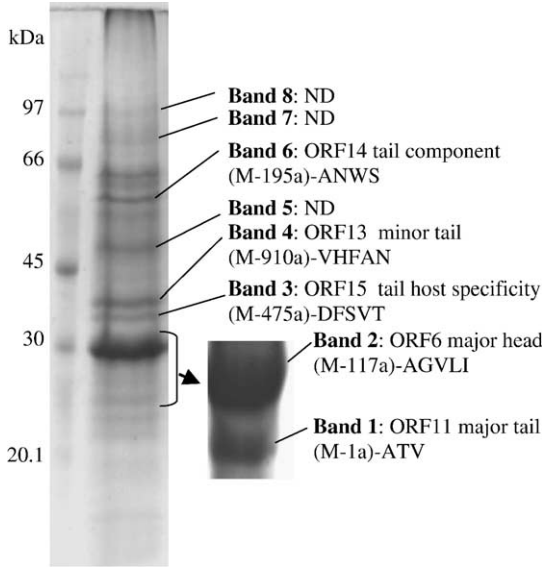

Fig. 2. Proteins of purified ΦAT3 virions were separated by 12% SDS–PAGE and then stained with Coomassie brilliant blue. Bands 1 and 2 are also shown with 10-fold more sample loaded. Bands 5, 7, and 8 correspond to proteins that poorly transferred to the membrane and were not analyzed (ND). Protein standards are indicated in kilodalton.

Morphology of the temperate bacteriophage ΦAT3

The bacteriophage ΦAT3 has a cos-site

To determine the end structure of the ΦAT3 genome, DNA was extracted from phage and digested with *Not*I. Comparing samples incubated or not for 10 min at 70 °C showed that heating results in two bands although the genome contains a unique *Not*I restriction site. This clearly indicates that there is a *cos*-site on the ΦAT3 genome. To determine the *cos*-site sequence of ΦAT3, the sequences between ORF53 of ΦAT3 and ORF175 of *St. thermophilus* phage Sfi21 (Proux et al., 2002), and ORF63 of *Staphylococcus aureus* phage PVL (Kaneko et al., 1998), and those between ORF1 of ΦAT3 and ORF4 of *Lb. casei* phage A2 (Ladero et al., 1999) were compared. A 135-bp noncoding region between ORF1 and ORF53 of ΦAT3 genome, likely containing the *cos*-site, was identified. A sequence run-off experiment using primers complementary to each 5′ end of the 135-bp noncoding region identified the *cos*-site of ΦAT3 genome as containing 15 unpaired nucleotides: 5′-ACCTCACACACCAGT-3′ (nucleotides 693–707 bp). The entire ΦAT3 genome sequence contains 39,166 bp. The GC contents of both ΦAT3 (44.6%) and *Lb. casei* phage A2 (44.8%, Garcia et al., 2003) genomes were closely matched with the chromosome of their common host *Lactobacillus* spp. (Pouwels and Leunissen, 1994). The lowest GC content of 35% lies in the region between ORF10 and ORF11 (Fig. 3). Analysis of the DNA sequence showed that there are 50 putative ORFs coded on one strand while Integrase (ORF20), Excisionase (ORF21), and *c*I repressor (ORF22) proteins are coded on the complementary strand (Table 1). Except for ORF12 and ORF43, all ORFs are preceded by a Shine–Dalgarno sequence complementary to the 3′ end of the 16S rRNA of *Lactobacillus* spp. (3′-UCCUCCAA-5′) (Schouler et al., 1994; Matern et al., 1994) 6–15 bp from one of the commonly used translation initiation codons (ATG, TTG, GTG) (Table 1). The ΦAT3 genome can be organized into five distinct functional clusters: packaging, morphogenesis, lysis, lysogenic/lytic switch, and replication (Fig. 3). Similar to several other phages, the locations of these clusters in the ΦAT3 genome are such that packaging lies on the far left, morphogenesis in the middle, and replication, including some ORFs of unknown function, on the far right. Most of the ORFs appeared to be densely coded on the ΦAT3 genome and were only separated by a few nucleotides, and eight ORFS overlapped (Table 1). No tRNA gene was detected by tRNAscan-SE (Lowe and Eddy, 1997) and FAStRNA (el-Mabrouk and Lisacek, 1996) programs.

Deduced proteins in the five functional clusters of ΦAT3 genome

Packaging

A comparison of sequences of the packaging, head and tail assembly functional domain of ΦAT3, and *Lb. casei* phage

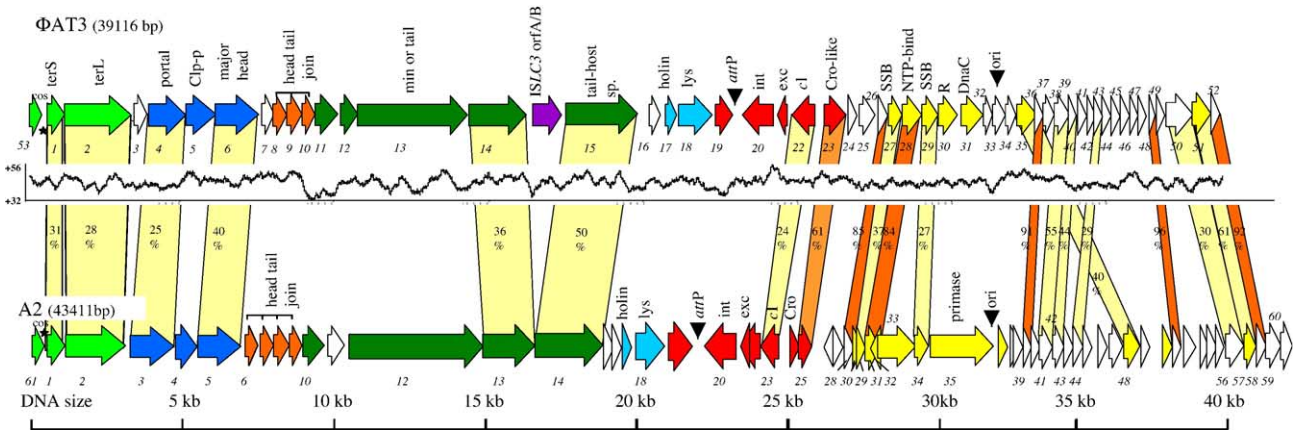

Fig. 3. Alignment of ΦAT3 and *Lb. casei* A2 genomes. The phage genomes are conventionally depicted with the DNA packaging genes at the left. ORFs are organized into five distinct functional clusters. The colors for each functional cluster are as follows: light green for DNA packaging; blue for head morphogenesis; orange for head-to-tail joining; green for tail morphogenesis; light blue for lysis; red for lysogenic-lytic switch; and yellow for DNA replication and transcription regulation genes. ORFs lacking a match are colored with white. The percentage of amino acid identity is indicated. The *att*P site or replication origin (ori.) is represented by an arrow. Levels of amino acid identity (>80, >60, >50, or ≤50%) are marked with red shading. The GC% of ΦAT3 DNA is shown along the entire phage genomes (GC%, 300 base windows).

A2 (Garcia et al., 1997) is shown in Fig. 3. With moderate homology over the four phages compared including *St. thermophilus* phages DT1 (Tremblay and Moineau, 1999) by the BLASP analysis, the most conserved gene detected was ORF1, which lies immediately downstream of the *cos*-site (Table 1). ORF1 corresponds to the putative small subunit of terminase of ΦAT3, which is most similar to ORF4 of *Lb. casei* phage A2. However, ORF2, the putative large terminase subunit, shares sequence identity of 39% and 38% with ORF623 and ORF624 of *St. thermophilus* phage Sfi21 and *Lb. gasseri* phage Φadh (Engel et al., 1998), respectively.

Morphogenesis

Sequence identity between ORF4 of ΦAT3 and ORF397 of *Lb. gasseri* phage Φadh is 29%, while that estimated between ORF4 and ORF384 of *St. thermophilus* phage Sfi21 is 27% (Table 1). A multi-subunit ring structure is formed by these putative portal proteins to serve as an entry point for the translocation of phage DNA into its head (Bazinet and King, 1988). ΦAT3 ORF5 shares sequence homology with a series of endopeptidase ClpP-like proteins, which are widespread among bacteria and eukaryotes and are recognized as a family of serine protease. The ClpP protein is a proteolytic subunit of a multi-component ATP-dependent protease complex that was initially described in *Escherichia coli* as a heat-shock-inducible system. Two highly conserved catalytic residues, Ser^{85} and His^{108}, have been identified in the active site of the C1pP-like proteins of some Gram-positive bacteria (Maurizi et al., 1990; Desiere et al., 1999). An IS element (IS*LC3*, GenBank accession No. AF445084) is inserted into ORF14 such that a premature stop code is introduced. The insertion causes the loss of 36 amino acids from the C-terminus of the original protein. Electron microscopy shows that some ΦAT3 particles might have extra long tails (Figs. 1B and C).

The arrangement of the putative tail morphogenesis genes ORF8 to ORF15 is very similar to those of *St. thermophilus* phages Sfi21 and BK5-T or *Lb. casei* phage A2. We suspect that the extra long tails may be caused by the insertion of IS*LC3* in ORF14. We propose that ORF5 is the processing enzyme for the ΦAT3 morphogenetic proteins. A multiple sequence alignment, with sequence identities ranging from 30% to 51%, reveals strong sequence similarity between ΦAT3 ORF5 and 20 different members of the large C1pP protease family. The catalytic domain of the ΦAT3 C1pP-like protein contains two conserved serine and histidine (Maurizi et al., 1990) flanked by two conserved glycine residues. The bacteriophage-encoded C1pP-like proteins have been suggested to be a new class of proteins because horizontal gene transfer from their hosts is rare (Desiere et al., 1999; Hendrix et al., 1999). The ORFs flanking ORF5, ORF4 and ORF6, are homologous with the corresponding ORF3 and ORF5 of *Lb. casei* phage A2 though ORF5 itself has no sequence similarity with the corresponding *Lb. casei* phage A2 protein (Fig. 3). There is also no significant sequence similarity between *Lb. casei* phage A2 protein and other ClpP family members compared (data not shown). However, it has been shown that *Lb. casei* phage A2 is more closely related to *Staphylococcus* phage PVL than to *St. thermophilus* phage Sfi21 over their structural gene regions (Proux et al., 2002).

Lysis

Holin (ORF17) and lysin (ORF18) of ΦAT3 are closely related to the lysis cassettes of other Gram-positive bacteriophages. ΦAT3 lysin is 70–80 amino acids shorter at its C-terminus than other lysins. A catalytic domain near the N-terminus and a target recognition domain are found that are comparable to lysins of the Siphoviridae (Loessner et al., 1995). It has been reported that a large deletion (295 out of 467 aa) at the C-terminus of *S. aureus* phage pvl lysin does not impair activity (Coleman et al., 1989). The N-

Table 1
ORFs and genetic features of bacteriophage ΦAT3

ORF order	Predicted start and stop site	ORF order a.a. (kDa)	Predicted function	Putative RBS[a] and start codon 3′-UCCUCCAA-5′	Significant match[b]	Gene	Identical/overall (%)	*E* value[c]	Accession no.
53	219–662	148 (17.5)	Unknown	atgcaAGGATgtaatcccg**tATG**	*S. aureus* phage PVL	ORF63	39/113 (34%)	1e−07	BAB95775
					St. thermophilu s phage Sfi21	ORF175	43/151 (28%)	1e−04	AAD41027
1	805–1269	155 (17.4)	Terminase small subunit	aaAGGAGGTgtgttttt**TTG**	*Le. mesenteroides* ATCC 8293				
					Putative prophage	Lmes0515	49/152 (32%)	2e−16	ZP_00063034
					Lb. casei phage A2	ORF1	40/127 (31%)	4e−07	CAD43892
					L. lactis phage BK5-T	ORF1	36/131 (27%)	7e−06	AAK56801
2	1259–3145	629 (72.3)	Terminase large subunit	tgaAGGAGGTTggctttg**GTG**	*Le. mesenteroides* ATCC 8293				
					Putative prophage	Lmes0514	398/626 (63%)	0	ZP_00063033
					St. thermophilus hage Sfi21	ORF623	233/584 (39%)	1e−113	AAD44086
					Lb. gasseri phage Φadh	ORF624	239/617 (38%)	1e−109	CAB52518
					L. lactis phage BK5-T	ORF2	232/602 (38%)	1e−108	CAC80143
					Lb. casei phage A2	ORF2	161/567 (28%)	1e−46	CAD43893
3	3148–3318	57 (6.1)	Unknown	atgtttGGAGGTggtaaata**ATG**	*Le. mesenteroides* ATCC 8293				
					Putative prophage	Lmes0513	27/56 (48%)	0.003	ZP_00063032
4	3328–4482	385 (42.3)	Portal protein	agaaAGGgGGTgaagttaa**ATG**	*Le. mesenteroides* ATCC 8293				
					Putative prophage	Lmes0512	199/385 (51%)	3e−96	ZP_00063031
					Lb. gasseri phage Φadh	ORF397	111/382 (29%)	2e−37	CAB52519
					St. thermophilus phage DT1	ORF6	103/372 (27%)	8e−28	AAD21882
					St. thermophilus phage Sfi21	ORF384	103/373 (27%)	2e−27	AAD44088
					Lb. casei phage A2	ORF3	77/299 (25%)	5e−10	CAD43894
5	4475–5158	228 (24.3)	ClpP-like protease	tgaAGGgAGGTgatgaag**GTG**	*Le. mesenteroides* ATCC 8293				
					Putative prophage	Lmes0511	117/227 (51%)	9e−56	ZP_00063030
					Lb. gasseri phage Φadh	ORF242	91/238 (38%)	9e−30	CAB52520
					St. thermophilus phage Sfi21	ORF221	88/227 (38%)	9e−28	AAD44089
6	5161–6342	394 (42.1)	Major head protein	caaaaAGGAGGTccaata**ATG**	*Le. mesenteroides* ATCC 8293				
					Putative prophage	Lmes0510	233/389 (59%)	1e−112	ZP_00063029
					Listeria innocua phage Bil285	lin2576	146/289 (50%)	6e−61	CAC97803
					Lb. casei phage A2	ORF5	138/363 (38%)	4e−52	CAD43896
					Lb. gasseri phage Φadh	ORF395	100/416 (24%)	4e−09	CAB52521
					L. lactis phage BK5-T	ORF7	90/405 (22%)	2e−08	CAC80148
					St. thermophilus phage Sfi21	ORF397	76/361 (21%)	4e−53	AAD44090
7	6416–6757	114 (12.4)	Unknown	gctgAGGAGGgctgaat**ATG**	*Le. mesenteroides* ATCC 8293				
					Putative prophage	Lmes0509	58/107 (54%)	2e−21	ZP_00063028

8	6750–7091	113 (12.7)	Head-tail joining	gcagcAAAGGctggtgggaa**ATG**	*L. lactis* phage bIL309	ORF45	39/113 (34%)	2e−08	AAK08393
					St. thermophilus phage DT1	ORF10	27/116 (25%)	0.002	AAD21886
					St. thermophilus phage Sfi21	ORF116	24/103 (23%)	0.003	AAC39276
					L. lactis phage BK5-T	ORF9	28/117 (23%)	0.036	CAC80150
9	7097–7513	139 (14.8)	Tail component	aactAAgGGggcataacc**GTG**	*St. thermophilus* phage 7201	ORF29	66/137 (48%)	2e−21	AAF43522
					St. thermophilus phage DT1	ORF11	64/137 (46%)	1e−20	AAD21887
					St. thermophilus phage Sfi21	ORF141	53/114 (46%)	9e−17	AAC39277
					Le. oenos phage L10	ORFA	50/151 (33%)	2e−10	AAA66332
10	7513–7881	123 (13.8)	Tail component	tGGcGGTgacagcaa**GTG**	*Le. oenos* phage L10	ORFI	43/123 (34%)	8e−16	AAA66333
					St. thermophilus phage DT1	ORF12	40/100 (40%)	5e−15	AAD21888
					St. thermophilus phage Sfi21	ORF123	40/105 (38%)	7e−15	AAC39278
11	7896–8525	210 (22.4)	Major tail	ttAGGAGGcattcaa**ATG**	*St. thermophilus* phage DT1	ORF13	100/197 (50%)	3e−43	AAD21889
					St. thermophilus phage 7201	mps-7201	99/197 (50%)	1e−42	AAB71820
					St. thermophilus phage Sfi21	ORF202	99/197 (50%)	1e−38	AAC39279
					Le. oenos phage L10	ORFE	75/194 (38%)	1e−26	AAA66334
					L. lactis phage BK5-T	ORF12	79/194 (40%)	5e−26	CAC80153
12	8686–9033	116 (13.3)	Tail component	tacaaAGGatggtattacca**ATG**	*St. thermophilus* phage DT1	ORF14	33/102 (32%)	3e−05	AAD21890
					St. thermophilus phage Sfi21	ORF117	30/102 (29%)	8e−04	AAC39280
13	9250–13,146	1299 (139.9)	Minor tail	agAGAAAGGgtgattaaa**GTG**	*Lb. rhamnosus* phage Lc-Nu	ORFC	206/215 (95%)	1e−112	AAQ03087
					St. thermophilus phage Sfi21	ORF1560	21/1212 (26%)	9e−83	AAC39280
					St. thermophilus phage DT1	ORF15	193/671 (28%)	1e−52	AAD21891
					Lb. gasseri phage Φadh	ORF1487	122/287 (42%)	2e−49	CAB52529
					L. lactis phage BK5-T	ORF15	74/232 (31%)	3e−24	CAC80156
14	13,150–15,129	660 (73.6)	Tail component	aaAGGAGGaaccgtttaa**TTG**	*Lb. rhamnosus* phage Lc-Nu	ORF644	311/346 (89%)	1e−108	AAQ03088
					Lb. casei phage A2	ORF13	143/394 (36%)	4e−47	CAD43904
					L. lactis phage BK5-T	ORF16	42/138 (30%)	7e−06	CAC80157
IS*LC3*	15,119–16,469	381 (44.2)	IS*LC3 (IS3* family IS*LC3* transposase	aaAGGAAGgaatttttac**ATG** A/B (15,200–16,344 bp)	*Lb. sanfranciscensis* IS*153* L. lactis IS*1076*	ORFAB ORF1	106/108 (98%) 127/383 (33%)	5e−54 1e−45	CAB63123 CAA37193
15	16,584–19,001	806 (88.2)	tail-host specificity	gaAGGAGGcgtggttatag**GTG**	*Lb. casei* phage A2	ORF14	349/689 (50%)	1e−172	CAD43905
					Lb. rhamnosus phage Lc-Nu	ORFD	184/192 (95%)	1e−100	AAQ03089
					Prophage Lambda Sa1	SAG0598	53/118 (44%)	2e−19	AAM99497
					St. thermophilus phage DT2	ORF18	50/129 (38%)	1e−16	AAK83245
					St. thermophilus phage Sfi21	ORF1276	51/117 (43%)	6e−16	AAC39283
16	19,505–19,885	129 (15.0)	Unknown	ggaAGGAAGtgatgaca**ATG**	*Lb. johnsonii* prophage Lj928	LJ1422	28/67 (41%)	6e−10	AAS09188
17	20,076–20,462	129 (13.7)	Holin	tcAGGAaGGaaaacaaatc**ATG**	*Le. oenos* phage 10MC	P163	52/115 (45%)	1e−13	AAD02488
					Lb. delbrueckii phage LL-H	hol	27/106 (25%)	1.8	AAC00556
18	20,550–21,728	393 (43.0)	Lysin	aatAGGAGGacacc**ATG**	*L. lactis* phage ul36	ORF429	187/401 (46%)	2e−77	AAM75805
19	21,849–22,346	166 (19.6)	Unknown	tgAtGAGGTgataatatc**ATG**	*L. lactis* phage bIL312	ORF24	59/164 (35%)	2e−17	AAK08477

(continued on next page)

Table 1 (*continued*)

ORF order	Predicted start and stop site	ORF order aa (kDa)	Predicted function	Putative RBS[a] and start codon 3′-UCCUCCAA-5′	Significant match[b]	Gene	Identical/overall (%)	*E* value[c]	Accession no.
20 (c)[d]	22,760–23,926	389 (44.5)	Integrase	aaAGGAGGaagattgc**ATG**	*L. lactis* phage bIL309	int	125/380 (32%)	7e−53	AAK08349
					Bacillus halodurans C-125	BH3551	114/392 (29%)	1e−40	BAB07270
					St. thermophilus phage Sfi21	ORF359	112/384 (29%)	2e−39	AAD44095
					L. lactis phage BK5-T	ORF32	117/394 (29%)	6e−37	CAC80173
					L. lactis phage ΦLC3	int	117/394 (29%)	1e−36	AAA32254
21 (c)[d]	24,103–24,462	120 (12.6)	Putative excisionase	ttGGAGGgattcatt**ATG**					
22 (c)[d]	24,523–25,293	257 (28.7)	*c*I repressor	acAGGAGGTgccacat**ATG**	*S. aureus* prophage ΦPV83	ORF5	76/248 (30%)	3e−23	BAA97812
					St. pyogenes MGAS8232 prophage	spyM18 _0718	83/265 (31%)	1e−19	AAL97389
					Clostridium perfringens Φ3626	ORF24	28/64 (43%)	2e−08	AAL96794
					L. lactis phage r1t	rro	74/284 (26%)	3e−08	AAB18678
					Lb. casei phage A2	ORF23	62/253 (24%)	1e−04	CAB63660
23	25,699–26,460	254 (28.0)	Cro-like protein	cctAAGGgatgacggt**ATG**	*Lb. johnsonii* prophage Lj928	LJ1455	125/249 (50%)	1e−61	AAS09223
					S. aureus prophage ΦPV83	ORF9	118/264 (44%)	1e−54	BAA97816
					Lb. casei phage A2	ORF25	97/160 (60%)	3e−42	CAB63662
					St. aureus phage PVL	ORF34	84/144 (58%)	2e−39	BAA31909
24	26,460–26,633	58 (0.6)	Unknown	caAAAGGgatgacggt**ATG**					
25	27,088–27,441	118 (13.9)	Unknown	gaaAGGAGGaaatgcc**ATG**	*L. lactis* phage: domain of unknown function		29/79 (37%)	6e−04	pfam05595
26	27,686–27,886	67 (0.7)	Unknown	ttAAGGgatgacatg**ATG**	*Lb. casei* phage A2	ORF28	57/67 (85%)	5e−26	CAB63665
27	27,908–28,390	161 (17.9)	Putative SSB	atAAGGaaaatattat**ATG**	*Lb. casei* phage A2	ORF29	57/153 (37%)	5e−16	CAB63666
					St. thermophilus phage Sfi21	ORF157	54/163 (33%)	1e−12	AAD44102
28	28,394–29,098	235 (26.3)	NTP. binding	tcGGActgggggtgcggtag**ATG**	*Lb. casei* phage A2	ORF31	190/225 (84%)	1e−104	CAB63668
					St. thermophilus phage phi O1205	ORF9	123/223 (55%)	4e−66	AAC79525
					St. thermophilus phage Sfi21	ORF233	124/223 (55%)	4e−66	AAF44103
					L. lactis phage BK5-T	ORF46	119/233 (51%)	8e−60	CAC80187
					Lb. gasseri phage Φadh	ORF223	85/214 (39%)	4e−35	CAB52496
29	29,105–29,659	185 (20.9)	SSB	acaAGGAGGactaaaac**ATG**	*L. lactis* phage mi7-9	gp18C	58/144 (40%)	1e−19	AAB22891
					St. thermophilus phage Sfi19	ORF151	52/135 (38%)	5e−14	AAD44072
					St. thermophilus phage DT1	ORF34	52/135 (38%)	8e−14	AAD21910
					St. thermophilus phage Sfi21	ORF124	52/135 (38%)	6e−13	AAC72436
					Lb. casei phage A2	ORF34	55/199 (27%)	1e−07	CAB63671
30	29,677–30,471	265 (31.1)	DNA replication	tttGAGGTgatcac**ATG**	*E. coli* O157:H7 phage Stx2 ΦII	ORFC59	38/113 (33%)	2e−09	BAC78107
					E. coli phage lambda	protein P	34/105 (32%)	6e−09	ORBPL
					Listeria innocua phage A118	gp49	34/112 (30%)	0.021	CAC97639

31	30,461–31,240	260 (29.6)	DNA replication protein (DnaC)	caAAGGAGaactggggat**ATG**	*Le. mesenteroides ATCC 8293*				
					Putative prophage	Lmes0523	83/231 (35%)	1e−30	ZP_00063042
					L. lactis phage bIL286	ORF17	57/188 (30%)	5e−11	AAK08304
					L. lactis prophage pi3	pi346	59/188 (31%)	9e−13	AAK05517
					Lb. delbrueckii phage LL-H	ORF267	47/155 (30%)	1e−11	AAL77547
32	31,240–31,569	110 (12.2)	Unknown	gtAAGGgggaagagact**GTG**					
33	31,571–32,015	149 (16.9)	Unknown	AGGgGGcagaatt**ATG**					
34	32,021–32,401	127 (14.8)	Unknown	gtGAGtcgttagcc**ATG**	*L. lactis* phage bIL286	ORF19	50/128 (39%)	7e−14	AAK08306
35	32,424–33,146	241 (27.2)	Endonuclease	ctAAGGAGaaaaaatc**ATG**	*Lb. casei* phage A2	ORF48	67/167 (40%)	2e−17	CAD43918
36	33,161–33,343	61 (0.7)	Unknown	atgAGAGGctaacaa**ATG**	*Lb. casei* phage A2	ORF39	54/59 (91%)	8e−23	CAB63676
37	33,343–33,747	135 (15.7)	Unknown	gtAGAGGacgaaaa**ATG**					
38	33,747–34,271	175 (19.8)	Unknown	ttGGAGGacgaaaa**ATG**	*Lb. casei* phage A2	ORF41	100/181 (55%)	1e−34	CAD43911
39	34,264–34,629	122 (13.9)	Unknown	aaGAGGTgactgacg**ATG**	*Lb. casei* phage A2	ORF42	63/141 (44%)	4e−17	CAD43912
40	34,625–34,894	90 (10.8)	Unknown	ttGGAGGcggatc**ATG**					
41	34,894–35,160	89 (10.4)	Unknown	ttGGAGGaagaaaa**ATG**					
42	35,160–35,492	111 (12.9)	Unknown	atGAGGTggagaa**ATG**	*Lb. casei* phage A2	ORF44	24/81 (29%)	0.16	AD43914
43	35,508–35,657	50 (0.5)	Unknown	aatcggcattaatcggttcatgtc**ATG**					
44	35,657–35,926	90 (10.5)	Unknown	tcAGGGAGGcggagaa**ATG**	*Lb.* phage Φgle	Rorf115	32/84 (38%)	3e−04	CAA66774
45	35,926–36,120	65 (7.6)	Unknown	ttGGAGatcctgatcag**ATG**					
46	36,139–36,342	68 (7.7)	Unknown	ctggAGGAAGaatcaa**ATG**					
47	36,454–36,759	102 (11.7)	Unknown	ctcAGAAGtcccgtg**ATG**	*Lb. casei* phage A2	ORF50	55/100 (56%)	5e−25	CAD43920
48	36,802–37,011	70 (7.5)	Unknown	gcttttccgtcatatGAtggtatt**GTG**	*Lb. casei* phage A2	ORF51	49/51 (96%)	1e−15	CAD43921
49	37,041–37,364	108 (12.0)	Unknown	ccAAGGTcgcgg**GTG**					
50	37,367–38,332	322 (37.1)	Unknown	tcaaacAGtcacggcctta**ATG**	*Lb. casei* phage A2	ORF56	51/168 (30%)	2e−09	CAD43926
51	38,187–38,852	222 (24.9)	Endo-deoxyribonuclease	tcatAGcGAGcggtttaaag**ATG**	*Lb. casei* phage A2	ORF57	104/168 (61%)	2e−52	CAD43927
					L. lactis phage Φ31	ORF6	59/168 (35%)	3e−17	CAC04164
52	38,843–39,163	107 (12.6)	Unknown	tttagtgGGAGGTgtaagct**ATG**	*Lb. casei* phage A2	ORF58	24/26 (92%)	2e−06	CAD43928

Note. Streptococcus (*St.*); *Staphylococcus* (*S.*); *Lactobacillus* (*Lb.*); *Lactococcus* (*L.*); *Leuconostoc* (*Le.*).

[a] Nucleotides complementary to the 3′ end of 16S rRNA of *Lactobacillus* ssp. (3′-UCCUCCAA-5′) (Schouler et al., 1994; Matern et al., 1994).

[b] The protein sequences in the databases which show homology with the respective ΦAT3 ORFs are shown. Empty cells indicate that no significant homologies were observed.

[c] *E*, expect value.

[d] Location on complementary strand is indicated by (c).

terminal region of ΦAT3 lysin is more conserved than the C-terminal. Although the lysin genes of other *St. thermophilus* phages (Desiere et al., 1998) are preceded by two holins, only one ORF with holin characteristics is found in ΦAT3.

Lysogenic/lytic switch

A comparison of the amino acid sequences of ΦAT3 ORF22 and ORF23, the putative repressor and Cro-like proteins, with those of ORF5 of *S. aureus* phage ΦPV83 and ORF25 of *Lb. casei* phage A2 (Ladero et al., 1998) gives identities of 30% and 60%, respectively (Table 1). The N-terminal sequence of ΦAT3 ORF23 was also compared with those of lambda and P22 and *E. coli Lac*I repressor and a DNA-binding domain and helix–turn–helix motif (Brennan and Brian, 1989) were identified. The Cro-like proteins of ΦAT3 and A2 are sufficiently similar to suggest that the phages have exchanged genetic material in the same *Lb. casei* host. It is known that lateral gene transfer between phages is a common event in the course of phage evolution (Hendrix et al., 1999). For example, the *St. thermophilus* phage Sfi21 shares a DNA sequence identity of 60% over the DNA packaging, head morphogenesis, and replication genes with those of the *L. lactis* phage BK5-T.

Replication

DNA sequence identity between ΦAT3 and *Lb. casei* phage A2 in the region of NTP-binding protein is high (Table 1). Homology with the putative single-stranded DNA-binding (SSB) proteins of *St. thermophilus* phages Sfi19, DT1, and *Lb. casei* phage A2 is found for ΦAT3 ORF29. SSB proteins bind to single-stranded DNA and play important roles in DNA repair, replication, and recombination (Bruttin et al., 2002). They are ubiquitously found in prokaryotic and eukaryotic organisms. The N-terminal region of ΦAT3 ORF29 is rich in α-helices and β-sheets while the C-terminal region consists of an acidic domain containing 10 acidic residues with four dispersed aspartates. This acidic domain (IDVSDDDLPF) is highly conserved among prokaryotic SSB proteins (Learn et al., 1997) but its function is not well defined. *E. coli* SSB is nonfunctional if its acidic domain is truncated (Lohman and Ferrari, 1994). The N-termini of *E. coli* and mitochondrial SSB proteins bind DNA. A putative helix–turn–helix motif is identified at the N-terminus of ΦAT3 ORF30, suggesting that the protein is also a DNA-binding protein. Homology with a replication protein O157:H7 of *E. coli* phage Stx2 was also found for the N-terminal region of ΦAT3 ORF30 (Table 1). The C-terminus of ORF30 was found to be 32% identical to lambda replication protein P (Table 1), suggesting that the region may be near the ΦAT3 origin of replication. It has been hypothesized that while the N-terminus of lambda P interacts with O and the replication origin, the C-terminus mediates protein–protein interactions (Casjens, 2003; Zylicz et al., 1984). ΦAT3 ORF31 could be a second type of replication protein as it contains 30% identity with the DnaC-like proteins of *L. lactis* phage bIL286 or *Lb. delbrueckii* phage LL-H. *E. coli* DnaC and lambda P protein are known to recruit the bacterial DnaB replicative helicase to the initiator complexes assembled at the origin of replication and to promote the formation of the replication fork (Stephens and McMacken, 1997). An AT-rich sequence, comprised of five short direct repeats, is present in two or three copies in ΦAT3 ORF31 (the potential DnaC protein) and ORF32. An inverted repeat capable of forming a stem-loop structure was also identified in this DNA segment. An AT-rich region clustered with several direct and inverted repeats are characteristics of phage replication origins (Schnos et al., 1989). ORF35 is a putative endonuclease that is related to ORF48 of *Lb. casei* phage A2. However, the relative positions of the genes are different, implying that they may have undergone different recombination processes after a horizontal gene transfer from their respective ancestral phages.

Integrase and the attachment sites

Sequence homology was found between ΦAT3 Int (ORF20) and the corresponding proteins of Gram-positive phages, including *L. lactis* phage bIL309 and *St. thermophilus* phage Sfi21 (Table 1). Sequence alignment of the C-terminal part of ΦAT3 Int against those of several phage Int proteins identified four highly conserved residues, Arg^{217}, His^{332}, Arg^{335}, and Tyr^{367}, which may be located near the catalytic site. A physical map of the *att*B region of *Lb. casei* ATCC 393, a natural host of ΦAT3, was determined by inverse PCR. The *att*B region is about 1 kb in length and by comparison of the *att*L and *att*R junctions with the *att*P and *att*B regions contains a common core region of 15 bp (ΦAT3 nucleotides 22,617–22,631 bp) (Fig. 4A). ΦAT3 was integrated in a $tRNA^{Arg}$ gene at the 3′ end of the *att*B region of *Lb. casei* ATCC 393 and the integration does not affect the integrity of the tRNA. The phage attachment site *att*P of ΦAT3 is located close to the 3′ end of the putative *int*, a typical organization found in a temperate phage. Both phage A2 and ΦAT3 infect strains of *Lb. casei*, but *att*B for phage A2 is located at the 3′ end of *Lb. casei* ATCC 393 $tRNA^{Leu}$ (Alvarez et al., 1998). An ORF, 319 bp downstream from the phage ΦAT3 common core sequence in the *Lb. casei* ATCC 393 chromosome, shares 98% sequence identity with *Lb. plantarum* WCFS1 α-galactosidases (Kleerebezem et al., 2003; Silvestroni et al., 2002; Ventura et al., 2003) (Fig. 4A).

Construction of a chromosomal integration vector using ΦAT3 int

To assay integration activity of ΦAT3 Int, the integrating vector pSKE-IN (Fig. 4B) was constructed by cloning the *int-att*P region (ΦAT3 nucleotides 22,497–24,052 bp) into plasmid pSKE. pSKE has a ColE1-type replicon and does not replicate in Gram-positive bacteria. The erythromycin resistance gene (Em^r) on the plasmid was used as a selection marker to test whether *int-att*P region was sufficient to

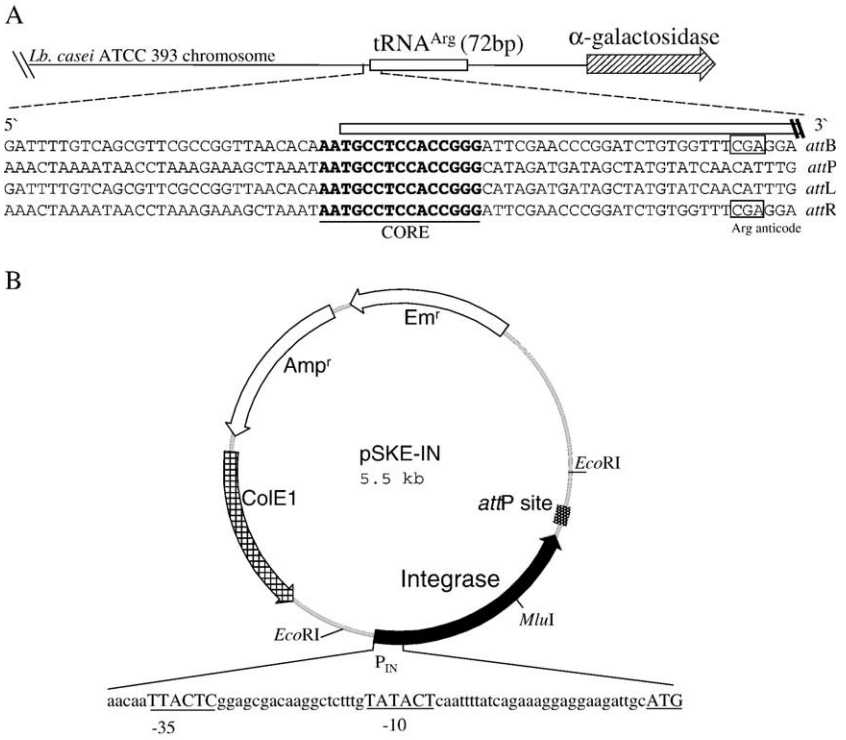

Fig. 4. (A) Schematic representation of the bacterial chromosomal attachment site (*att*B) of *Lb. casei* ATCC 393 containing part of a putative gene. The deduced gene product is homologous to α-galactosidase. Nucleotide positions of the sequenced part of the ORF and that of the 15-bp core are highlighted. The aligned nucleotide sequences of the regions contain *att*P, *att*L, *att*R, and *att*B and the anticodon 5′-CGA-3′ of $tRNA^{Arg}$ are boxed. (B). The integrating vector pSKE-IN constructed for assaying ΦAT3 Int activity. Major genetic elements on the vector are ΦAT3 *int*, *att*P, genes for resistance to ampicillin (Ampr) and erythromycin (Emr), and the ColEI replication origin. The −10 and −35 regions and the ATG start codon of the potential *int* promoter (P_{IN}) are highlighted.

mediate site-specific recombination in vivo. Plasmid pSKE-IN was electroporated into *Lactobacillus rhamnosus* HN 001, *Lactobacillus paracasei* ATCC 27,092, and *Leuconostoc mesenteroides* ATCC 8293 by electroporation. *Lb. paracasei* ATCC 27,092 and *Le. mesenteroides* ATCC 8293 were not transformed but *Lb. rhamnosus* HN 001 yielded 1.0 transformants/μg of supercoiled DNA. DNA of three randomly chosen transformants was digested with *Mun*I, which cleaves pSKE-IN only once within *int* and was analyzed by a Southern blot hybridization using plasmid pSKE as a probe. Two bands would be detected if integration was at a unique site. However, there were four definite and one or two faint bands detected in all three transformants, suggesting that recombination of pSKE-IN in *Lb. rhamnosus* HN 001 occurs at more than one site (Fig. 5). Using inverse PCR followed by DNA sequencing, we found that plasmid pSKE-IN was integrated at the 3′ end of the *Lb. rhamnosus* HN 001 $tRNA^{Arg}$ gene and a sequence homology of 95% between the gene with *Lb. casei* ATCC 393 $tRNA^{Arg}$ gene was detected. It is currently unknown whether the $tRNA^{Arg}$ gene is duplicated in *Lb. rhamnosus* HN 001 chromosome or not. Thus, the two lactobacilli have the same target *att*B sites in their chromosomes. If other lactobacilli also contain an *att*B site for ΦAT3, it could serve as an efficient target for gene addition.

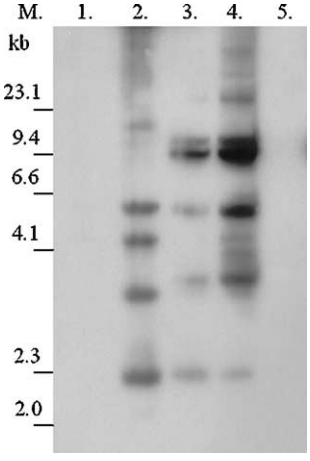

Fig. 5. Integration of pSKE-IN into several lactobacilli chromosomes was confirmed by some Southern hybridization experiments. Plasmid pSKE was used as a probe in the Southern hybridization experiments. Lane M is DNA size marker while lanes 1, 2, 3, 4, and 5 are DNA samples from *Lb. casei* ATCC 393, *Lb. rhamnosus* HN 001 (pSKE-IN)-1, (pSKE-IN)-2, (pSKE-IN)-3, and *Lb. rhamnosus* HN 001, respectively.

Materials and methods

Bacterial strains and culture conditions

Bacterial strains and plasmids used are shown in Table 2. LAB were cultivated at 37 °C or 30 °C in MRS (Difco)

Table 2
Bacterial strains and plasmids

Strains or plasmids	Relevant characteristics	Reference or sources
Strains		
E. coli DH5α	*recA endA1 hsdR17 supE4 gyrA96 relA1* Δ(*lacZYA-argF)U169* (Φ80d*lacZ*Δ M15)	BRL[a]
E. coli TG1	*supE hsd* Δ*5 thi* Δ[*lac-proAB*] F [*traD36 proAB l acI*^{q}*lacZ* ΔM15]	Sambrook[b]
Lb. casei	ATCC 393	ATCC[c]
Lb. paracasei	ATCC 27092	ATCC
Le. mesenteroides	ATCC 8293	ATCC
Lb. rhamnosus	HN 001	Fonterra[d]
Lb. rhamnosus	HN 001(pSKE-IN)-1, -2, -3	This study
Plasmids		
pBluescript SKII+ (pSK+)	Cloning vector, Ap^r, CoEI ori.	Stratagene
pGEM-T Easy	TA Cloning vector, Ap^r, CoEI ori.	Promega
pE194	Include *erm*C gene (Em^r) vector, propagated in *Bacillus subtilis*	ATCC
pSKE	pSK+ include erythromycin-resistant gene from pE194, Ap^r, Em^r, CoEI ori.	This study
pSKE-IN	pSKE include ΦAT3 integrase gene and *att*P site, Ap^r, Em^r, CoEI ori.	This study

[a] BRL, Bethesda Research Laboratories.
[b] Sambrook et al., 1989.
[c] ATCC, American Type Culture Collection.
[d] Fonterra Research Centre (earlier known as New Zealand Dairy Research Institute).

without aeration or on 1.2% MRS agar. *E. coli* DH5α or TG1 were used as a host for cloning and promoter activity assays and were grown in Luria-Bertani (LB) medium (Difco) at 37 °C with vigorous shaking. Antibiotics were added at a concentration of 50 μg/ml for selection purpose. When indicated, X-Gal (20 μg/ml) was spread onto the LB plates. Most chemicals were purchased from Sigma; restriction enzymes, T4 DNA ligase, and Klenow DNA polymerase were obtained from New England Biolabs.

Phage preparation

Bacteriophage ΦAT3 was induced from *Lb. casei* ATCC 393 using 0.2 μg/ml mitomycin-C. One liter of phage lysate was treated with RNase (10 μg/ml) and DNase (10 μg/ml) for 30 min at 37 °C to ensure degradation of host nucleic acids. Phage particles were precipitated with a buffer containing 1 M NaCl and 10% (wt/vol) (PEG 8000), incubated at 4 °C for 2 h, and then centrifuged at 8000 rpm for 30 min. The pellet was gently resuspended in SM buffer (100 mM NaCl, 8 mM $MgSO_4$, 50 mM Tris–HCl pH 7.0, and 0.04% gelatin), and insoluble material was removed by low-speed centrifugation. The resulting preparation was centrifuged through a glycerol step gradient by carefully adding 1.0 ml of the aqueous phase onto the top of the gradient. The step gradient was prepared by adding 0.5 ml of 40% glycerol in SM buffer to a centrifuge tube and then freezing at −20 °C; 0.65 ml 5% glycerol in SM buffer was then layered on top and the tube stored at −20 °C until use. Centrifugation was performed at 20,000 rpm in a TH-641 swinging bucket rotor at 4 °C for 2 h. The pellet was gently resuspended in SM buffer and further purified by CsCl density step gradient (1.3, 1.5, and 1.7 g/cm^3) centrifugation for 2.5 h at 30,000 rpm in the same rotor.

Electron microscopy

Purified ΦAT3 particles were fixed on a copper grid before being negatively stained with 1% uranyl acetate. Electron micrographs were taken using a JEOL1200CXII transmission electron microscope (Garcia et al., 2003). Electron microscopy was performed at the Instrument Center of National Chung Hsing University, Taiwan.

DNA preparation

The phage DNA was extracted from the phage particles using a procedure described by Qiagen (Qiagen, GmbH, Hilden, Germany) for coliphage lambda. Alkaline lysis (Sambrook et al., 1989) was used for plasmid purification. For isolation of total genomic DNA from *Lactobacillus* spp., bacterial cells were gently lysed by SDS–proteinase K followed by extraction with phenol–chloroform and then precipitation with ethanol as described (Forsman and Alatossavam, 1994). The QIAquick Gel Extraction Kit (Qiagen) was used to extract and purify DNA from the standard agarose gels.

Plasmids construction and PCR assay

PCR or inverse PCR experiments were performed using *Pfu* DNA polymerase (Promega) according to the manufacturer's instructions and using 30 cycles to amplify the region of interest. The annealing temperatures were 5–7 °C below the lowest *Tm* of each primer pair. To clone PCR products, some appropriate restriction enzyme sites were inserted into the 5′ ends of primers for a direct TA cloning using vector pGEMT-Easy (Promega). Various ΦAT3 DNA fragments were cloned into pBluescript SK+ (Stratagene).

Construction of a vector carrying erythromycin-resistant gene and ΦAT3 int

The erythromycin-resistant gene of pE194 (Horinouchi and Weisblum, 1982) was amplified by PCR using the

forward Ery-A (5′-TTAAACCGTGTGCTCT-3′) and reverse primer Ery-B (5′-GCGCAAAAGACATAAT-3′). The 1.1-kb PCR fragment was inserted into *Ssp*I-digested pBluescript SK+ after addition of dTMP at the 3′-ends to create a T-vector for TA cloning. The vector was designated as pSKE. ΦAT3 *int* and *att*P were amplified using primer ATR (5′-AGAGAAGCCATTGAAACACGATTTAA-TAAC-3′) and *Bam*H-IN (5′-CCGGATCCTGGGGAT-GAATAC-3′). The PCR product was ligated to the pGEMT-Easy vector (Promega) and after transformation the desired product was confirmed by sequencing. A 1.6-kb *Eco*RI fragment excised from this plasmid was then cloned into pSKE to give vector pSKE-IN.

Nucleotide and proteins sequencing

Cloned DNA was treated with exonuclease III (New England Biolabs) and a series of subclones containing deleted fragments were used as the sequencing templates. Sequence gaps were bridged by designing appropriate primers. Plasmid DNA was purified for sequencing by the Qiagen miniprep procedure (Qiagen), followed by DNA sequencing as described in the Applied Biosystems (ABI) *Taq* DyeDeoxy terminator cycle sequencing kit. DNA sequencing was performed by Mission Biotech Ltd. (Taiwan) using ABI model 3700 automated sequencers. The nucleotide sequence across junction points between clones was determined by chromosome walking using synthetic oligonucleotides and purified ΦAT3 DNA as template. Analysis of the nucleotide sequence was conducted using the software Vector NTI Suite 6.0 (InforMax, Bethesda, MD). ORFs were identified using FASTA (GenBank, EMBL, SwissProt, and DDBJ), blastn, tblastn, and ScanProsite (Prosite and SwissProt; http://www.expasy.org/cgi-bin/scanprosite) databases (Altschul et al., 1990, 1997). OrfFinder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) programs were used for ORF prediction. Homologies to sequences in the GenBank database were searched using Blast 2.0 and PSI-BLAST (http://www.ncbi.nlm.nih.gov/BLAST). Molecular weights were calculated by ExPasy Protparam (http://www.expasy.ch/tools/protparam.html). Helix–turn–helix motifs were compiled using the Web server http://npsa-pbil.ibcp.fr/. The sequence reported here is deposited in GenBank with an accession no. AY605066. The N-terminal sequences of phage proteins isolated from 12% SDS–PAGE were obtained by automated Edman degradation using the ABI model Procise® 492 sequencer (Applied Biosystems) by the biochemistry laboratories of National Taiwan University.

Determination of the attachment R/L/P/B and cos-sites

The sequence of *cos*-site of ΦAT3 was determined by a sequence run-off experiment using the oligonucleotide COSA (5′-AGTTGACAGAAGCAAGTGC-3′) and COSB (5′-AACAATCGCCTTCCAAAGACG-3′), which lie approximately 250 bp distal to either end of the linear ΦAT3 genome. To determine the *att* sites, the *Lb. casei* ATCC 393 genomic DNA was partially digested with *Sau*3AI or *Acc*I. Digested genomic DNA fragments were self-ligated to generate circularized substrates for use in an inverse PCR. *att*R and *att*L junctions were amplified using primers atA (5′-TTTGACTGAAAAATCTAGAAGC-3′), atB (5′-GGGGACAAAAAGGGGACA-AG-3′), atC (5′-GATTG-TGTTTGTCCCCTTT-3′), and atD (5′-CAGCTGATCCA-TGAATAACACGG-3′). The corresponding PCR products were cloned and sequenced.

Southern blotting

An equal amount of chromosomal DNA from each strain to be analyzed was digested with *Mun*I and subjected to electrophoresis on a 0.8% agarose gel. DNA fragments were transferred to a Hybond-N+ membrane (Amersham) and cross-linked in a UV cross-linker (Spectronics). The membrane was prehybridized with 0.5 ml of a 1 mg/ml solution of sonicated salmon sperm DNA at 42 for 1 h. The pSKE DNA probe was labeled by nick-translation with [α-^{32}P]dATP. The membrane was hybridized with probe at 42 °C for 12 h, followed by washing under high stringency conditions: 30 min with 2× SSC–0.1% SDS (1× SSC is 150 mM NaCl plus 15 mM sodium citrate) followed by 30 min with 0.5× SSC–0.1% SDS, both at room temperature, and then 30 min at 55 °C with 0.1× SSC–0.1% SDS.

Acknowledgments

This work was supported by Grants NSC91-2313-B007-001 and NSC92-2313-B007-002 from the National Science Council, Taiwan, ROC. We also thank the instrument center of National Chung Hsing University for its technical assistance in the electron microscopy.

References

Ackermann, H.W., DuBow, M.S., 1987. Viruses of Prokaryotes. CRC Press, Boca Raton, FL, pp. 103–142.

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

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Alvarez, M.A., Herrero, M., Suárez, J.E., 1998. The site-specific recombination system of the *Lactobacillus* species bacteriophage A2 integrates in Gram-positive and Gram-negative bacteria. Virology 250, 185–193.

Bazinet, C., King, J., 1988. Initiation of P22 procapsid assembly in vivo. J. Mol. Biol. 202, 77–86.

Black, L., Showe, M., 1983. Morphogenesis of the T4 head. In: Mathews, E., Kutter, E., Mosig, G., Berget, P. (Eds.), Bacteriophage T4. ASM Publications, Washington, pp. 219–245.

Brennan, R.G., Brian, W.M., 1989. The helix–turn–helix DNA binding motif. J. Biol. Chem. 264, 1903–1906.

Brüssow, H., 2001. Phages of dairy bacteria. Annu. Rev. Microbiol. 55, 283–303.

Brüssow, H., Desiere, F., 2001. Comparative phage genomics and the evolution of Siphoviridae: insights from dairy phages. Mol. Microbiol. 39, 213–222.

Bruttin, A., Foley, S., Brüssow, H., 2002. DNA-binding activity of the *Streptococcus thermophilus* phage Sfi21 repressor. Virology 303, 100–109.

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

Casjens, S., 2003. Prophages and bacterial genomics: what have we learned so far? Mol. Microbiol. 49, 277–300 (Review).

Coleman, D.C., Sullivan, D.J., Russell, R.J., Arbuthnott, J.P., Carey, B.F., Pomeroy, H.M., 1989. *Staphylococcus aureus* bacteriophages mediating the simultaneous lysogenic conversion of beta-lysin, staphylokinase and enterotoxin A: molecular mechanism of triple conversion. J. Gen. Microbiol. 135, 1679–1697.

Desiere, F., Lucchini, S., 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.

Desiere, F., Pridmore, R.D., Brüssow, H., 1999. Comparative genomics of *Streptococcus thermophilus* phage species supports a modular evolution theory. J. Virol. 73, 8647–8656.

el-Mabrouk, N., Lisacek, F., 1996. Very fast identification of RNA motifs in genomic DNA: application to tRNA search in the yeast genome. J. Mol. Biol. 264, 46–55.

Engel, G., Altermann, E., Klein, J.R., Henrich, B., 1998. Structure of a genome region of the *Lactobacillus gasseri* temperate phage Φadh covering a repressor gene and cognate promoters. Gene 210, 61–70.

Forsman, P., Alatossavam, T., 1994. Repeated sequences and the sites of genome rearrangements in bacteriophages of *Lactobacillus delbrueckii* subsp. *lactis*. Arch. Virol. 137, 43–54.

Garcia, P., Alonso, J.C., Suárez, J.E., 1997. Molecular analysis of the cos region of the *Lactobacillus casei* bacteriophage A2. Gene product 3, gp3, specifically binds to its downstream cos region. Mol. Microbiol. 23, 505–514.

Garcia, P., Ladero, V., Suárez, J.E., 2003. Analysis of the morphogenetic cluster and genome of the temperate *Lactobacillus casei* bacteriophage A2. Arch. Virol. 148, 1051–1070.

Hendrix, R.W., Smith, M.C.M., Burns, R.N., Ford, M.E., Hatfull, G.F., 1999. Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. Proc. Natl. Acad. Sci. U.S.A. 96, 2192–2197.

Herrero, M., de los Reyes-Gavilan, C.G., Caso, J.L., Suárez, J.E., 1994. Characterization of Φ393-A2, a bacteriophage that infects *Lactobacillus casei*. Micobiology 140, 2585–2590.

Horinouchi, S., Weisblum, B., 1982. Nucleotide sequence and functional map of pE194, a plasmid that specifies inducible resistance to macrolide, lincosamide, and streptogramin type B antibodies. J. Bacteriol. 150, 804–814.

Josephsen, J., Neve, H., 1998. Bacteriophages and lactic acid bacteria. In: Salminen, S., von Wright, A. (Eds.), Lactic Acid Bacteria, Microbiology and Functional Aspects, 2nd ed. Marcel Dekker Inc., New York, pp. 385–436.

Kaneko, J., Kimura, T., Narita, S., Tomita, T., Kamio, Y., 1998. Complete nucleotide sequence and molecular characterization of the temperate staphylococcal bacteriophage ΦPVL carrying Panton–Valentine leukocidin genes. Gene 215, 57–67.

Katsura, I., 1990. Mechanism of length determination in bacteriophage lambda tails. Adv. Biophys. 26, 1–18 (Review).

Katsura, I., Hendrix, R.W., 1984. Length determination in bacteriophage lambda tails. Cell 39 (3 Pt. 2), 691–698.

Kleerebezem, M., Boekhorst, J., van Kranenburg, R., Molenaar, D., Kuipers, O.P., Leer, R., Tarchini, R., Peters, S.A., Sandbrink, H.M., Fiers, M.W., Stiekema, W., Lankhorst, R.M., Bron, P.A., Hoffer, S.M., Groot, M.N., Kerkhoven, R., De Vries, M., Ursing, B., de Vos, W.M., Siezen, R.J., 2003. Complete genome sequence of *Lactobacillus plantarum* WCFS1. Proc. Natl. Acad. Sci. U.S.A. 100, 1990–1995.

Ladero, V., Garcia, P., Bascaran, V., Herrero, M., Alvarez, M.A., Suárez, J.E., 1998. Identification of the repressor-encoding gene of the *Lactobacillus* bacteriophage A2. J. Bacteriol. 180, 3474–3476.

Ladero, V., Garcia, P., Alonso, J.C., Suarez, J.E., 1999. A2 Cro, the lysogenic cycle repressor, specifically binds to the genetic switch region of *Lactobacillus casei* bacteriophage A2. Virology 262, 222–229.

Learn, B.A., Um, S.J., Huang, L., McMacken, R., 1997. Cryptic single-stranded-DNA binding activities of the phage lambda P and *Escherichia coli* DnaC replication initiation proteins facilitate the transfer of *E. coli* DnaB helicase onto DNA. Proc. Natl. Acad. Sci. U.S.A. 94, 1154–1159.

Le Marrec, C., van Sinderen, D., Walsh, L., Stanley, E., Vlegels, E., Moineau, S., Heinze, P., Fitzgerald, G., Fayard, B., 1997. Two groups of bacteriophages infecting *Streptococcus thermophilus* can be distinguished on the basis of mode of packaging and genetic determinants for major structural proteins. Appl. Environ. Microbiol. 63, 3246–3253.

Loessner, M.J., Schneider, A., Scherer, S., 1995. A new procedure for efficient recovery of DNA, RNA, and proteins from *Listeria* cells by rapid lysis with a recombinant bacteriophage endolysin. Appl. Environ. Microbiol. 61, 1150–1152.

Lohman, T.M., Ferrari, M.E., 1994. *Escherichia coli* single-stranded DNA-binding protein: multiple DNA-binding modes and cooperativities. Annu. Rev. Biochem. 63, 527–570.

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

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

Matern, H.T., Henrich, J.R., Plapp, B., 1994. Determination and comparison of *Lactobacillus delbrueckii* ssp. *lactis* DSM 7290 promoter sequences. FEMS Microbiol. Lett. 122, 121–128.

Maurizi, M.R., Clark, W.P., Kim, S.H., Gottesman, S., 1990. C1pP represents a unique family of serine proteases. J. Biol. Chem. 265, 12546–12552.

Pouwels, P.H., Leunissen, J.A.M., 1994. Divergence in codon usage of *Lactobacillus* species. Nucleic Acids Res. 22, 929–936.

Proux, C., van Sinderen, D., Suarez, J., Garcia, P., Ladero, V., Fitzgerald, G.F., Desiere, F., Brüssow, H., 2002. The dilemma of phage taxonomy illustrated by comparative genomics of Sfi21-like Siphoviridae in lactic acid bacteria. J. Bacteriol. 184, 6026–6036.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Schnos, M., Zahn, K., Blattner, F.R., Inman, R.B., 1989. DNA looping induced by bacteriophage lambda O protein: implications for formation of higher order structures at the lambda origin of replication. Virology 168, 370–377.

Schouler, C., Ehrlich, S.D., Chopin, M.C., 1994. Sequence and organization of the lactococcal prolate-headed bIL67 phage genome. Microbiology 140, 3061–3069.

Silvestroni, A., Connes, C., Sesma, F., Savoy de Giori, G., Piard, J.-C., 2002. Characterization of the *melA* melA locus for α-galactosidase in *Lactobacillus plantarum Lactobacillus plantarum*. Appl. Environ. Microbiol. 68, 5464–5471.

Stephens, K.M., McMacken, R., 1997. Functional properties of replication fork assemblies established by the bacteriophage lambda O and P replication proteins. J. Biol. Chem. 272, 28800–28813.

Tremblay, D.M., Moineau, S., 1999. Complete genomic sequence of the

lytic bacteriophage DT1 of *Streptococcus thermophilus*. Virology 255, 63–76.

Tsui, L.C., Hendrix, R.W., 1983. Proteolytic processing of phage lambda tail protein gp *H*: timing of the cleavage. Virology 125, 257–264.

Ventura, M., Canchaya, C., Kleerebezem, M., de Vos, W.M., Siezen, R.J., Brüssow, H., 2003. The prophage sequences of *Lactobacillus plantarum* strain WCFS1. Virology 316, 245–255.

Zylicz, M., Gorska, I., Taylor, K., Georgopoulos, C., 1984. Bacteriophage lambda replication proteins: formation of a mixed oligomer and binding to the origin of lambda DNA. Mol. Gen. Genet. 196, 401–406.