



**HAL**  
open science

# Induction, structural characterization, and genome sequence of Lv1, a prophage from a human vaginal *Lactobacillus jensenii* strain

Rebeca Martín, Susana Escobedo, Juan E. Suárez

► **To cite this version:**

Rebeca Martín, Susana Escobedo, Juan E. Suárez. Induction, structural characterization, and genome sequence of Lv1, a prophage from a human vaginal *Lactobacillus jensenii* strain. *International Microbiology*, 2010, 13 (3), pp.113-121. 10.2436/20.1501.01.116 . hal-02667445

**HAL Id: hal-02667445**

**<https://hal.inrae.fr/hal-02667445>**

Submitted on 31 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

# Induction, structural characterization, and genome sequence of Lv1, a prophage from a human vaginal *Lactobacillus jensenii* strain

Rebeca Martín,<sup>1</sup> Susana Escobedo,<sup>1</sup> Juan E. Suárez<sup>1,2\*</sup>

<sup>1</sup>Microbiology Unit, University Institute of Biotechnology, University of Oviedo, Oviedo, Spain.

<sup>2</sup>Institute of Dairy Products of Asturias-CSIC, Villaviciosa, Spain

Received 24 May 2010 · Accepted 26 July 2010

**Summary.** The prophage Lv1, harbored by a vaginal *Lactobacillus jensenii* isolate, was induced by several different anti-cancer, antimicrobial, and antiseptic agents, suggesting that they contribute to the adverse vaginal effects associated with their therapeutic use. Of special interest with respect to its novelty was the inducing effect of nonoxynol-9, a non-ionic detergent commonly used as a spermicide. The Lv1 genome consists of a 38,934-bp dsDNA molecule with cohesive ends, in which 48 ORFs were recognized, and is organized into functional modules. Lv1 belongs to the family Siphoviridae and, more precisely, to the proposed Sfi21-like genus. The capsid-tail junction of the Lv1 virions is fragile such that most particles become disrupted, suggesting that the virus is defective and thus unable to generate fertile progeny. However, genome analysis did not provide evidence of the defective nature of the prophage, other than the finding that its genome is shorter than those of other, related, phages. Further analysis indicated that prophage Lv1 suffered deletions in its right half to the extent that it no longer fulfill the minimum packaging limits, thereby generating the observed unstable particles. [Int Microbiol 2010; 13(3):113-121]

**Keywords:** *Lactobacillus jensenii* · phage Lv1 genome · bacteriophages · vaginal lactobacilli · SOS response

## Introduction

Lactobacilli constitute over 70% of all bacteria sampled from the vaginas of healthy, fertile women [40,31]. These bacteria protect the mucosal surface from pathogen colonization by three complementary mechanisms: (i) direct competition for attachment to the epithelial surface, (ii) production of antimi-

crobials, such as lactic acid and hydrogen peroxide, and (iii) co-aggregation with pathogens, further neutralizing their adherence to the mucosa thus enhancing the antimicrobial effect of the lactobacilli [5,40].

Based on these properties, lactobacilli have been envisaged as probiotic agents in the control of vaginal infections as well as in relapse prevention [26]. A further advantage of lactobacilli as vaginal probiotics is that they are naturally resistant to most antibiotics used, either systemically or in situ, to combat vaginal disorders. Thus, they are not affected by metronidazole and clindamycin [24], first-line agents in the treatment of bacterial vaginosis and, in the case of metronidazole, for trichomoniasis as well. Furthermore, a significant proportion is resistant to nonoxynol-9 [24], a widely used spermicide commonly included in creams and in condom lubricants. Conversely, vaginal lactobacilli are extremely susceptible to  $\beta$ -lactams, macrolides, and tetracy-

---

\*Corresponding author: J.E. Suárez  
Área de Microbiología, Facultad de Medicina  
Universidad de Oviedo  
Julián Clavería, 6  
33006 Oviedo, Spain  
Tel. +34-985103559. Fax +34-985103148  
E-mail: evaristo@uniovi.es

cline. However, in the few isolates with moderate resistance towards the latter two antibiotics, genes encoding ribosomal protection or antibiotic-efflux proteins, usually linked to conjugative transmission of the resistances, have not been detected [24]. Lactobacilli tend also to be susceptible to anti-neoplastic drugs, many of which are active on gram-positive bacteria [32]. The sensitivity of lactobacilli to these drugs might explain the increase in vaginal disorders observed after treatment of systemic infections and tumors [9,36]. The chemotherapeutics probably permeate the vaginal surface and inhibit the development of indigenous lactobacilli, thus allowing the overgrowth of minor components of the vaginal microbiota, such as *Candida albicans* and *Gardnerella vaginalis*, or even colonization by enteric microorganisms.

In a previous report, we postulated that activation of the SOS response plays a complementary role in the development of vaginal side effects produced by systemic radio- and chemotherapy, through induction of prophages resident in the genomes of vaginal lactobacilli [23]. This conclusion was based on the following: firstly, most antineoplastic agents and several antimicrobials inhibit DNA synthesis, which is known to activate the SOS response [17,37]; furthermore, even certain  $\beta$ -lactams and a septum-blocking bacteriocin exhibit the same property [22,27]. Secondly, lysogeny is widespread among environmental lactobacilli [8] and it has been shown that a component of tobacco smoke induces a prophage harbored by a vaginal *Lactobacillus* [28], which might explain the high frequency of bacterial vaginosis observed in female smokers [16]. However, we could not isolate any *Lactobacillus* phages from vaginal exudates nor did we obtain viral progeny after treatment of 45 strains of vaginal lactobacilli with mitomycin C. Nevertheless, lysogeny appeared to be widespread since phage-specific sequences could be amplified from more than one-third of the strains. In some cases, these sequences were responsive to SOS activation, as judged by their concentration increase following challenge with mitomycin C. Finally, some of the induced culture supernatants were inhibitory when placed over lawns of other strains but never led to the formation of lytic plaques upon dilution. Taken together, these findings were considered indicative of the presence of defective prophages in a substantial proportion of *Lactobacillus* vaginal isolates [23]. It should be noted that most of these lactobacilli produce hydrogen peroxide ( $H_2O_2$ ), which has been shown to be a SOS-activating compound [24,25], suggesting that  $H_2O_2$  promotes the selection of bacterial cells harboring SOS-insensitive, defective prophages [23].

That study provided further proof of the involvement of prophage induction in the increase in vaginal disorders that follow radio- and chemotherapy. In fact, the culture super-

natant of a *L. jensenii* strain inhibitory to other lactobacilli contained virions with isometric capsids and long, non-contractile tails. This allowed classification of the virus, named Lv1, within the family Siphoviridae [23]. The current report addresses the responsiveness of this prophage towards treatment with potentially SOS-activating agents. Furthermore, the structural and genetic characteristics of this virus are discussed, as are the reasons why it might have become defective.

## Materials and methods

**Bacterial strain and culture conditions.** *Lactobacillus jensenii* Lv39 was isolated from the vaginal exudate of a healthy, reproductive-aged woman [24]. It was grown on MRS agar as previously described [23].

**Prophage excision assays.** Prophage excision assays were carried out essentially as described elsewhere [32]. Briefly, samples (25 ml) of *L. jensenii* cultures in the early exponential phase ( $OD_{600} = 0.1$ ) were treated with the following potentially SOS-inducing agents [usually at half of the minimal inhibitory concentration (MIC)]:  $H_2O_2$  (1 mM), UV radiation (1 J/m<sup>2</sup>), mitomycin C (0.45  $\mu$ g/ml), mitramycin (0.03  $\mu$ g/ml), ampicillin (0.5  $\mu$ g/ml), biliary salts (10 mg/ml), and nonoxynol-9 (0.4%). Incubation was continued under the same conditions and aliquots of 100  $\mu$ l were taken from the cultures every 30 min and immediately frozen. Real-time PCR samples were prepared in a final volume of 15  $\mu$ l, including 1.25  $\mu$ l of the cell culture, using the PCR Q SYBR Green Supermix (Bio-Rad) according to the manufacturer's instructions. The primers used matched phage sequences placed at each side of the *attP* attachment site: 5'-GTCACAGTAGTTACAA-GAGTTCTGCG-3' and 5'-CAGCGACGTGGCAAAC-3'. Amplifications were carried out on a iCycler iQ Multicolor Real-Time PCR detection system (Bio-Rad) with a thermocycle profile as follows: stage 1, 95°C (10 min); stage 2, 30 cycles of 95°C (1 min), 50°C (1.5 min), and 72°C (1 min). The data obtained were recorded as threshold cycle ( $C_T$ ) values; i.e., the cycle number during which the fluorescence signal crossed the threshold set by the manufacturer of the thermocycler.

**Purification and structural characterization of virions.** Virions were purified as described previously [23] through precipitation with polyethylene glycol from induced *L. jensenii* Lv39 cultures and banding into two successive CsCl-gradient centrifugations. The concentrated suspensions were stored at 4°C in the presence of CsCl. For electron microscopy, drops of dialyzed virion suspensions were placed onto carbon-coated grids and then negatively stained with 2% (w/v) uranyl acetate. Phage structural proteins were analyzed by heating dialyzed suspensions at 70°C for 15 min, followed by incubation for 1 h with 1  $\mu$ g DNase I/ml at 37°C. The samples were then subjected to SDS-PAGE (6% and 12% acrylamide in the stacking and separating gels, respectively) and stained with Coomassie brilliant blue following standard methods. The bands of interest were excised manually, digested with porcine trypsin (Promega), and the resulting peptides analyzed by MALDI-TOF-TOF mass spectrometry, essentially as previously described [15].

**Phage DNA purification, sequencing, and analysis.** Genomic DNA was extracted from purified virions using the phenol-chloroform method after disruption of the particles by incubation in a 50 mM EDTA solution at 70°C. Restriction enzymes were used as indicated by the supplier (Takara). For sequencing, approximately 10  $\mu$ g of phage DNA was sheared by sonication, size-selected (2–3 kbp), and cloned into pUC18. Individual clones were sequenced and assembled. Trace assembly was done

with the phredPhrap package [12]. Generally, the assembled phage genome sequences had at least five-fold coverage in both orientations. Primer walking was sometimes used to close gaps in the sequence.

The *cos* sites were localized through comparisons of the restriction patterns of ligated and unligated DNA samples. To obtain the sequence of the single-stranded *cos* extensions, the ends of the DNA were sequenced directly using primers that pointed outwards. The nucleotide sequences obtained were aligned with that of the circular genome and found to overlap by 12 bases that were considered to constitute the *cosN* site.

Open reading frames (ORFs) were predicted with Clone Manager 7 and ORF Finder [<http://ncbi.nlm.nih.gov/gorf/gorf.html>] and by visual inspection. The primary nucleotide sequence was scanned in all reading frames for the start codons AUG and GUG with a threshold of 50 codons. BLASTX and BLASTP [<http://www.ncbi.nlm.nih.gov/blast/Blast.cgi>] were used to search for homologous proteins. Structural predictions and motif searches were performed with InterProScan [<http://www.ebi.ac.uk/InterProScan/>], Pfam [<http://pfam.sanger.ac.uk/search?tab=searchSequenceBlock>], YASPIN [<http://zeus.cs.vu.nl/programs/yaspinwww/>] and TMHMM Server v. 2.0 [<http://www.cbs.dtu.dk/services/TMHMM-2.0/>]. Sequences of the  $\sigma^{70}$  promoter were identified using Bprom [<http://www.softberry.com>]. Putative terminator sequences were detected with the Terminator function of GCG (version 10.2). Putative tRNA encoding genes were searched for with tRNAscan-SE 1.21 [<http://selab.janelia.org/tRNAscan-SE/>]. Prediction of +1 and -1 frameshifting was carried out with FSFinder 2 [<http://wilab.inha.ac.kr/fsfinder2/>].

The accession number for the phage Lv1 genome is EU871039.

## Results and Discussion

**Inducibility of the Lv1 prophage.** Phages modulate the composition of bacterial communities through a “kill the winner” mechanism, thus contributing to the maintenance of balanced ecosystems. This mechanism appears to be restricted in the vagina due to the generation of H<sub>2</sub>O<sub>2</sub> by indigenous lactobacilli. H<sub>2</sub>O<sub>2</sub>-mediated activation of the SOS response, by inducing resident prophages, may select for bacteria harboring defective viruses unable to generate progeny [23]. The SOS response can also be activated by many of the drugs used to treat infections and cancer as well as by therapeutic radiation. Some of these agents were therefore examined regarding their ability to induce prophage Lv1, resident in the genome of the vaginal isolate *L. jensenii* Lv39. As representative anticancer drugs, mitomycin C and mitramycin were chosen (others, such as rebeccamycin, could not be used because the strain was not susceptible). To mimic the antibacterial effect of radiotherapy, ultraviolet light was chosen. Since vaginal lactobacilli are naturally resistant to metronidazole and quinolones [24], the two most conspicuous antibacterial DNA inhibitors, ampicillin was used as it has been shown to activate the SOS response in *Escherichia coli* [27]. Finally, both nonoxynol-9, the most widely used spermicide, and bile salts, due to the possibility of passage of lactobacilli from the enteric tract to the vagina, were tested [3].

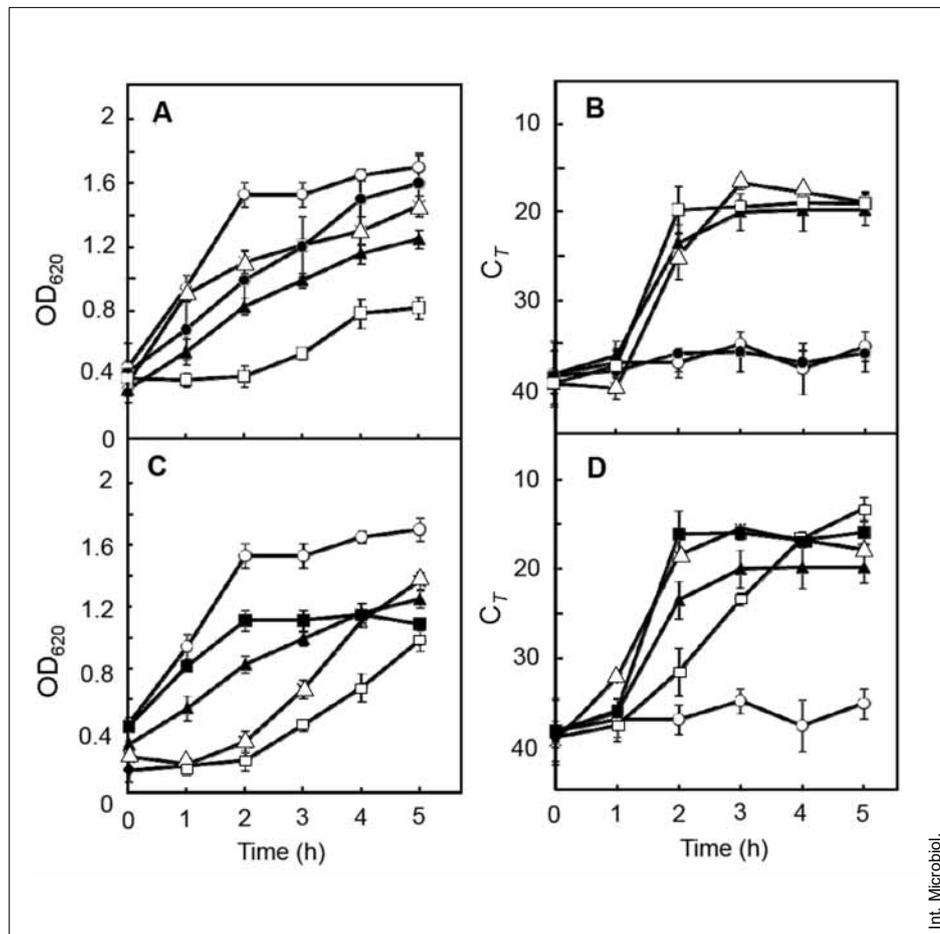
Since Lv1 did not produce lytic plaques on any of a collection of 45 strains of vaginal lactobacilli [23], converging

oligonucleotides, located at each side of the *attP* attachment site, were designed to quantify the induction of the phage's lytic cycle using real-time PCR [32].

All conditions tested, with the exception of bile salts, induced the lytic cycle of Lv1 at least as efficiently as achieved with H<sub>2</sub>O<sub>2</sub>, when used at concentrations that reduced but did not suppress the growth of *L. jensenii* Lv39 (Fig. 1A,D). In the case of the antitumor compounds tested and UV radiation, this result was expected because of their interactions with DNA: mitomycin C induces cross-linking of the two strands [34], mitramycin interacts with the minor groove and inhibits transcription [21], and UV induces the formation of pyrimidine dimers. Ampicillin most likely activated the DpiBA two-component signal transduction system, as it does in *E. coli*. DpiA binds A+T-rich sequences at the origin of replication, thus competing with DnaA and DnaB and transiently inhibiting bacterial cell division, which, in turn, activates the SOS response [27]. Note that homologues to *dpiA* and to *dpiB* were recognized in the sequenced genomes of two *Lactobacillus rhamnosus* strains (NC\_013198-1 and NC\_013199-1). The inducing effect of nonoxynol-9 deserves further study. This non-ionic detergent reduces the surface tension of the acrosomal membrane, thereby inhibiting the movement of spermatozoa. It may thus be assumed that the spermicide's effect on bacteria involves a disturbance of the phospholipid bilayer structure of membranes, a mechanism of action that has not previously been linked to SOS activation. However, only about half of the vaginal lactobacilli strains challenged with nonoxynol-9 were found to be susceptible [24]. This absence of a general effect and the capacity of the drug to induce the lytic cycle of Lv1 suggest that, in bacteria, the membrane might not be the main target of nonoxynol-9.

The increase in vaginal disorders that is commonly associated with cancer and antimicrobial therapies might be due, in part, to the induction of *Lactobacillus* resident prophages which, as is the case with Lv1, lyse their host cells, thus contributing to depletion of the beneficial *Lactobacillus*-dominant vaginal microbiota and subsequent colonization by undesirable microorganisms. However, this effect would be partial, as 50% of lysogenic vaginal lactobacilli harbor defective prophages unresponsive to SOS-activating conditions [23].

**Structural characteristics of Lv1 virions.** Electron microscopy of CsCl-purified Lv1 suspensions showed complete virions and an abundance of both empty capsids and isolated tails (Fig. 2A), although the samples were examined as soon as possible to avoid phage disruption due to storage. Since CsCl-gradients separate par-



**Fig. 1.** Influence of sublethal concentrations of several agents on the growth (A,C) and generation of autonomous Lv1 genomes (B,D) by a lysogenic culture of *Lactobacillus jensenii* Lv39, as measured by real-time PCR of the *attP* segment. Note that on the ordinate axis the threshold cycle ( $C_T$ ) values are depicted in descending order to provide a more intuitive representation (the lower the  $C_T$  value, the higher the initial concentration of the DNA segment being amplified). (A,B) Open circles, untreated cultures; closed triangles, 1 mM  $H_2O_2$  (positive control); closed circles, 10 mg biliary salts/ml; open triangles, 0.4% nonoxynol-9; open squares, 0.5  $\mu$ g ampicillin/ml. (C,D) Open circles, untreated cultures; closed triangles, 1 mM  $H_2O_2$ ; closed circles, 0.45  $\mu$ g mitomycin C/ml; open triangles, 0.03  $\mu$ g mitramycin/ml; open squares, 1 J UV radiation/m<sup>2</sup>. Error bars represent the standard deviations of three independent experiments.

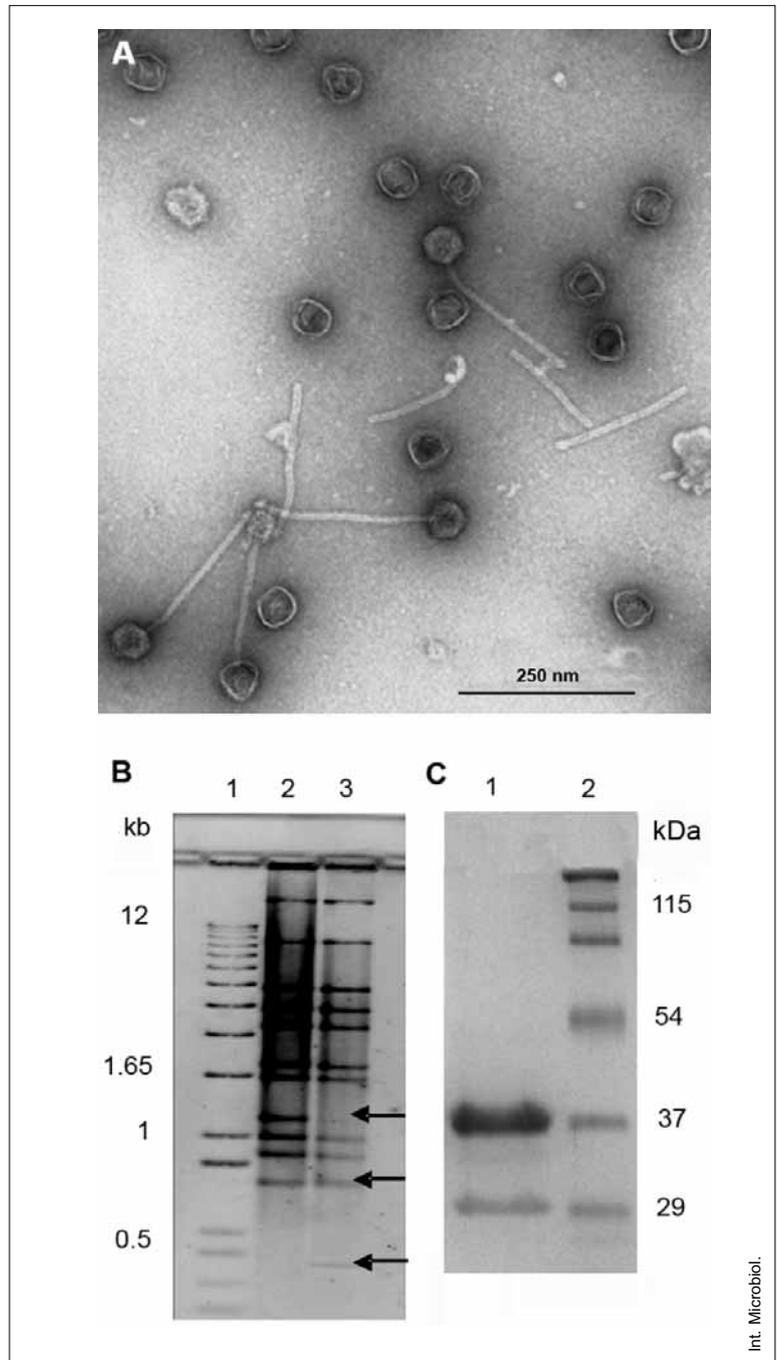
Int. Microbiol.

ticles as a function of their buoyant density, the band extracted had to be comprised exclusively of complete virions. In fact, viral DNA was easily obtained from these suspensions, which indicates that it was packaged into viral capsids. These findings suggest that the capsid-tail junction of Lv1 is unstable such that the virus is incapable of forming lytic plaques due not only to the lack of an appropriate host but also to the instability of its virions.

**Analysis of the phage Lv1 genome.** To determine the reason(s) for Lv1 virion fragility, its genome was sequenced and analyzed. The viral genome was located in a dsDNA molecule 38,934 bp long. Its G+C content was determined to be 37%, which is within the range of the values

recorded for *Lactobacillus* strains of species closely related to *L. jensenii* [2,4].

Treatment of the DNA with T4 DNA-ligase revealed the presence of cohesive ends (Fig. 2B), consisting of 13-bp single-stranded 3' extensions with the sequence 5'-CACACAACAGGGG-3'. This sequence is 70% identical to the *cosN* site of AT3, a *Lactobacillus casei* phage [20], and 60% identical to a sequence located in the 5' vicinity of the small terminase gene of a prophage harbored in the *Lactobacillus fermentum* genome (accession number NC\_010610). In the genome of Lv1, 48 putative ORFs using more than 50 codons were recognized (Fig. 3A, Table S1), all of which contained an AUG start codon. About 12.7% of the genome was non-coding; the start codon of four ORFs overlapped with the

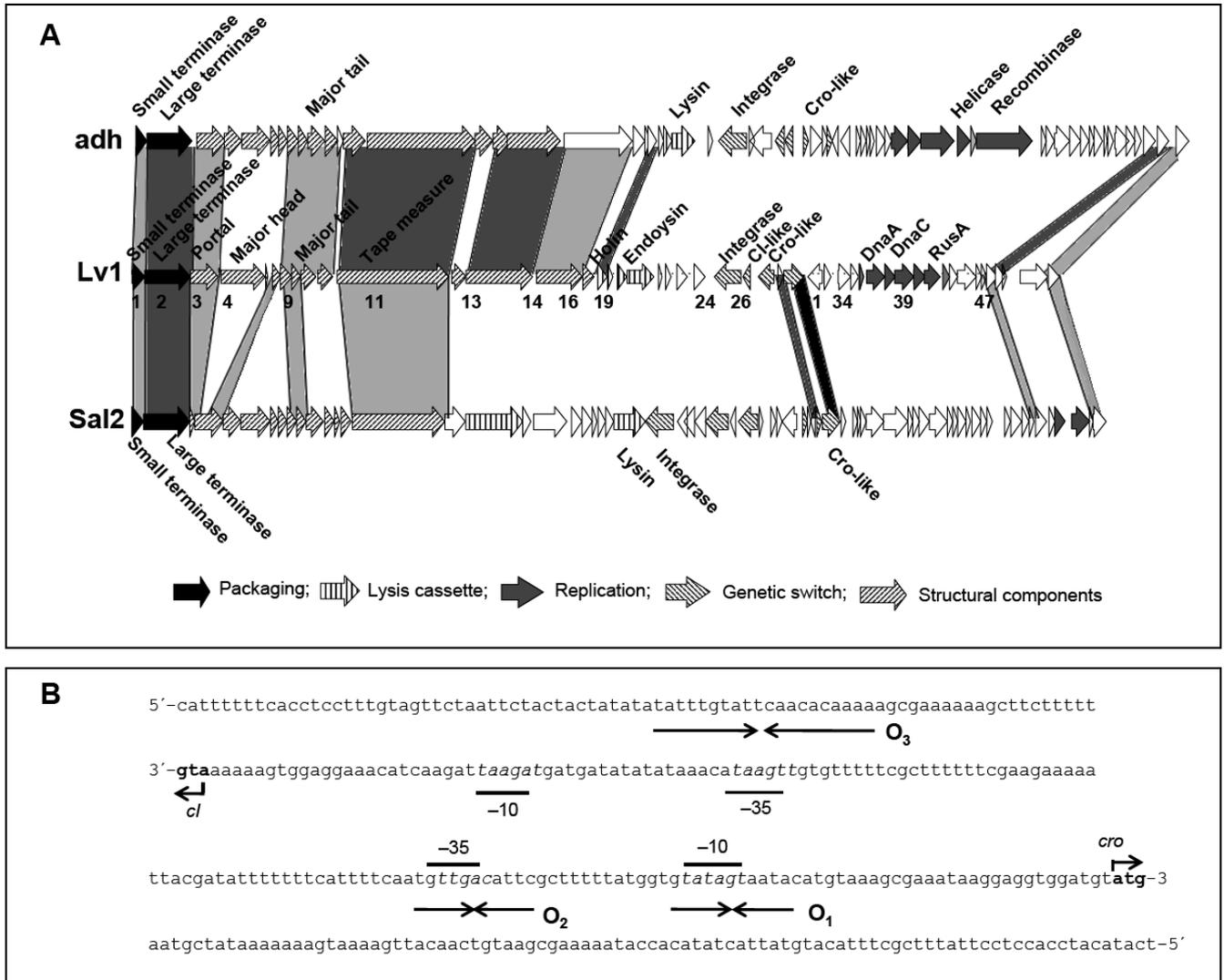


**Fig. 2.** (A) Electron micrograph of an Lv1 phage suspension obtained upon treatment of its host, *Lactobacillus jensenii* Lv39, with mitomycin C and purification of the virions through buoyant density CsCl gradient centrifugation. (B) Restriction pattern (*EcoRV*) of ligated (lane 2) and unligated (lane 3) DNA. The arrows point to the bands in the unligated DNA lane and to the result of their fusion after treatment with T4-ligase. (C) SDS-PAGE of the major structural proteins of Lv1.

stop codon of the previous one and nine had their ends inside the following ORF, indicative of expression through polycistronic mRNAs and translational coupling. Gene transcription was from right to left for the genes presumed to be involved in the establishment of the lysogenic cycle, i.e., from ORF26 to ORF24 and including those encoding the CI-like repressor and the integrase. All but one of the remaining ORFs (ORF29) read from left to right, starting at the *cro* homologue (ORF27), continuing through the cohesive ends,

and finishing at the *attP* site, where a  $\rho$ -independent terminator was detected.

The genome was annotated after comparison with current databases (Fig. 3A, Table S1). Significant matches were obtained for 31 ORFs and biological functions were assigned to most of them. No tRNA genes were found. As is the case for most phages that infect lactobacilli [8], the Lv1 genome is divided into several modules comprising groups of genes whose products are involved in a common function.



**Fig. 3.** (A) Physical and genetic map of the phage Lv1 and alignment with those of phages adh and Sal2. The grey spaces linking ORFs between genomes indicate degrees of amino acid identity: light gray, between 25 and 39%; dark gray, between 40 and 50%; black: >50%. The 48 ORFs of Lv1 are correlatively numbered, indicated by arrows proportional to their length and pointing towards their direction of transcription. The code of the functional modules detected is indicated at the bottom of (A). Above the scheme, the names of several putatively or experimentally identified genes are shown. (B) Nucleotide sequence of the region between the putative *cl* and *cro*. Consensus promoter sequences (-35 and -10 boxes), the starts of translation, and the operator regions are indicated.

The left extreme of the genome is made up of the DNA packaging cluster, including the genes encoding the small and large terminase subunits. The product of the latter gene has a motif typical of large terminases that render 3' protruding cohesive ends [10], which is in agreement with our experimental data.

The capsid morphogenesis module comprises ORF3 and ORF4. Gp3 shares motifs with phage portal proteins, while gp4 is a composite protein of 615 amino acids that in the databases matches proteases in its amino terminus (up to residue 174) and major head proteins in its carboxyl terminus, with several  $\alpha$ -helix motifs predicted in its middle part. MALDI-

TOF/TOF analysis of the most abundant Lv1 virion protein identified peptides that matched gp4 from its 365th residue onwards, suggesting that native gp4 is autoprocessed by its protease domain to give rise to the major capsid polypeptide. This may explain its apparent size of 37 kDa, as observed by SDS-PAGE (Fig. 2C). The  $\alpha$ -helices found in the middle of gp4 are characteristic of scaffold proteins, which together with their location, i.e., fused to the major capsid polypeptide, indicate that this is precisely the function of the central part of gp4. The products of ORF5 to ORF8 present regions of homology with *Lactobacillus* phage proteins putatively involved in the morphogenesis of viral neck structures.

The major tail protein is encoded by ORF9, as deduced from the homology of its product with others of similar function and, especially, by its correspondence to the peptides obtained by MALDI-TOF/TOF of the second most abundant virion protein (Fig. 2C). The apparent size of this protein, 25 kDa, coincides quite well with the predicted 22.3 kDa of gp9. Gp11 is postulated to be the tape-measure protein and, as expected, its secondary structure is almost exclusively  $\alpha$ -helix. Between the major tail and the tape-measure genes lies ORF10, which does not have the translational frameshift site frequently encountered in this position in other dsDNA phages [38]. ORF13 encodes a long polypeptide of 972 residues that presents a motif for putative phage antireceptors. This and its localization in the genome suggest that it is the tail-tip protein, usually involved in host recognition.

The lysis cassette appears to extend from ORF16 to ORF19. A promoter-like sequence is placed 5' of the cluster, suggesting that these genes are expressed as an operon. Gp16 might be considered to be a class I phage holin because it presents three transmembrane domains and a predicted N-out, C-in topology [39]. Moreover, it harbors two adjacent start codons, suggestive of a holin-antiholin lysis regulatory mechanism in Lv1, as is true for phage  $\lambda$  [39]. The phage endolysin is predicted to be encoded by ORF19, which presents a glycoside hydrolase motif that would confer lysozyme activity to gp19. The products of the intervening genes, ORF17 and ORF18, do not match any of the sequences or motifs stored in the databases.

The stretch from ORF20 to ORF23 exhibits several features typical of lysogenic conversion genes (morons), such as a disparate G+C content (31% vs. 37% for the phage as a whole), no identity with described sequences, and a location between the lysis cassette and the *attP* site, which is preferentially occupied by virulence determinants in *Staphylococcus aureus* and *Streptococcus pyogenes* prophages [6].

The genetic switch of the phage may be executed by the alternative expression of ORF26 and ORF27, which encode, respectively, CI- and Cro-like repressors that present helix-turn-helix motifs in their amino-terminal ends. These genes are separated by a 168-bp intergenic region that contains two divergently oriented promoters and three stretches of hyphenated symmetry that might act as operators (Fig. 3B) [14,29]. Downstream of ORF26, and separated from it by a small gene lies a determinant containing the motifs of Cre integrases [33]. It is followed by a 385-bp non-coding stretch that might constitute the *attP* sequence, based not only on its location but also on its high A+T content (73%) and the presence of direct and inverted repeats. Following the *cro* determinant, a putative antirepressor gene and several ORFs without relatives in the databases are located.

The replication cluster seems to expand from ORF34 to ORF39, with the functions deduced from their protein sequences as follows: DnaA (ORF34), DnaC (ORF36), single-stranded DNA-binding determinant (ORF37), and RusaA resolvase (ORF39).

In the last 4.5 kb of the Lv1 genome are nine ORFs, most of which do not have counterparts in the databases. One exception is ORF44, which encodes an ArpU-related protein. ArpU is thought to control the export of muramidase-2, a peptidoglycan hydrolase found in the peptidoglycan of *Enterococcus hirae* [19]. The other two exceptions are ORF47 and ORF48, whose products have partial homology with putative proteins of other lactic acid bacteria phages and which might encode a methylase and an endonuclease, respectively.

Overall comparison of the genome of Lv1 with those of other phages revealed a general similarity with respect to the functional modules present and to their order in the genome. However, significant sequence homologies were found only with the proteins encoded by *adh*, a phage that infects *L. gasseri* strains [1], and with Sal2, a prophage detected in the genome of a *L. salivarius* strain and which is inducible with mitomycin C, although it is not known whether infective progeny are produced [35]. This relatedness is mainly restricted to the left half of the genome, i.e., the terminases and the structural protein cluster (Fig. 3A).

Observation of Lv1 virion suspensions indicated that the junction capsid-tail is unstable and, accordingly, that it is a defective phage. If this were indeed so, then genetic differences between Lv1 (and possibly Sal2) and the proficient *adh* would be expected. These would presumably involve the structural region of the genome and, more precisely, the stretch limited by the major head and major tail protein determinants, where the genes that putatively encode the proteins of the neck region are located. However, in all three phages this region encodes four short polypeptides, such that it is essentially impossible to recognize differences that might account for the different phenotypes observed. Nevertheless, the genome of *adh* is longer (43.8 kb) than those of Sal2 (40.3 kb) and Lv1 (38.9 kb), suggesting that the latter two suffered deletions affecting their infectivity. According to the maps shown in Fig. 3A, the gene shortage is located mainly in the right half of the genomes. The replication cluster and a series of *orfs*, mostly with undefined functions, are located in this stretch. As Lv1 is clearly able to replicate its own DNA, the defect should be attributed to the surrounding genes. Although this possibility cannot be excluded, it is improbable because no gene involved in the architecture of the virion is located in the right half of the phage genome. Furthermore, most deletion mutants obtained upon treatment of phages as

dissimilar as  $\phi$ C31 of *Streptomyces* and A2 of *Lactobacillus casei* with chelating agents, have lost DNA in this region [11,18], indicating that this region does not contain essential genes. However, the reduced genome size might still account for the defectiveness of Lv1, due to the packaging constraints shown by phages with cohesive ends. In the case of lambda, effective packaging only proceeds when the DNA molecule to be introduced into the capsid is between 78% and 105% of the unit-genome length [13], while in  $\phi$ C31 the packaging limits are 91% and 103% [11]. The genome of Lv1 is 11.2% shorter than that of adh, while their capsids have similar sizes ( $58.8 \pm 07$  nm vs. 62 nm) [23,30]. Therefore, it appears to be the deletion itself, rather than the genes affected, that has converted Lv1 into a defective phage.

The classification of Lv-1 into the family Siphoviridae, together with the presence of cohesive ends in its genome and of one gene whose product is processed to yield the scaffolding and major capsid proteins, allows its inclusion in a widely represented group of viruses that infect gram-positive bacteria, as represented by Sfi21 of *Streptococcus thermophilus* [7].

**Acknowledgements.** This work was supported by the CICYT grants BFU2007-65781 and AGL2010-15097 from the Spanish Ministry of Science and Technology and the EU-FEDER Plan. S.E. and R.M. are holders respectively of a fellowship associated with the former grant and a scholarship from FICYT (Principado de Asturias). We thank Alfredo F. Braña for the generous gift of the antitumor drugs.

## References

1. Altermann E, Klein JR, Henrich B (1999) Primary structure and features of the genome of the *Lactobacillus gasseri* temperate phage jadH. *Gene* 236:333-346
2. Altermann E, Russell WM, Azcarate-Peril MA, et al. (2005) Complete genome sequence of the probiotic lactic acid bacterium *Lactobacillus acidophilus* NCFM. *Proc Natl Acad Sci USA* 102:3906-3912
3. Antonio MAD, Rabe LK, Hillier SL (2005) Colonization of the rectum by *Lactobacillus* species and decreased risk of bacterial vaginosis. *J Infect Dis* 192:394-398
4. Azcarate-Peril MA, Altermann E, Goh YJ, et al. (2008) Analysis of the genome sequence of *Lactobacillus gasseri* ATCC 33323 reveals the molecular basis of an autochthonous intestinal organism. *Appl Environ Microbiol* 74:4610-4625
5. Boris S, Barbés C (2000) Role played by lactobacilli in controlling the population of vaginal pathogens. *Microbes Infect* 2:543-546
6. Brüssow H (2006) Prophage genomics. In: Calendar R (ed) *The bacteriophages*, 2nd ed. Oxford University Press, New York, pp 17-25
7. Brüssow H, Desiere F (2001) Comparative phage genomics and the evolution of *Siphoviridae*: insights from dairy phages. *Mol Microbiol* 39:213-222
8. Brüssow H, Suárez JE (2006) *Lactobacillus* phages. In: Calendar R (ed) *The bacteriophages*, 2nd ed. Oxford University Press, New York, pp 653-666
9. Carta G, Di Stefano L, Porzio G, Caserta D, Mascaretti G (1995) Bacterial vaginosis during chemotherapy for gynecological neoplasms. A therapeutic proposal. *Minerva Ginecol* 47:5-8
10. Casjens SR, Gilcrease EB, Winn-Stapley DA, et al. (2005) The generalized transducing *Salmonella* bacteriophage ES18: complete genome sequence and DNA packaging strategy. *J Bacteriol* 187:1091-1104
11. Chater KF, Bruton CJ, Springer W, Suárez JE (1981) Dispensable sequences and packaging constraints of DNA from the *Streptomyces* temperate phage  $\phi$ C31. *Gene* 15:249-256
12. Ewing B, Green P (1998) Base-calling of automated sequencer traces using *phred*. II Error probabilities. *Genome Res* 8:186-194
13. Feiss M, Fisher RA, Crayton MA, Egner C (1977) Packaging of the bacteriophage lambda chromosome: effect of chromosome length. *Virology* 77:281-293
14. García P, Ladero V, Alonso JC, Suárez JE (1999) Cooperative interaction of CI protein regulates lysogeny of *Lactobacillus casei* by bacteriophage A2. *J Virol* 73:3920-3929
15. García P, Rodríguez I, Suárez JE (2004) A -1 ribosomal frameshift in the transcript that encodes the major head protein of bacteriophage A2 mediates biosynthesis of a second essential component of the capsid. *J Bacteriol* 186:1714-1719
16. Hillier SL, Nugent RP, Eschenbach DA, et al. (1995) Association between bacterial vaginosis and preterm delivery of a low-birth-weight infant. *N Engl J Med* 333:1737-1742
17. Janion C (2001) Some aspects of the SOS response system: a critical survey. *Acta Biochim Pol* 48:599-610
18. Ladero V, García P, Bascarán V, Herrero M, Alvarez M, Suárez JE (1998) Identification of the repressor-encoding gene of the *Lactobacillus* bacteriophage A2. *J Bacteriol* 180:3474-3476
19. Lleó MM, Fontana R, Solioz M (1995) Identification of a gene (*arpU*) controlling muramidase-2 export in *Enterococcus hirae*. *J Bacteriol* 177:5912-5917
20. Lo TC, Shih TC, Lin CF, Chen HW, Lin TH (2005) Complete genomic sequence of the temperate bacteriophage  $\phi$ AT3 isolated from *Lactobacillus casei* ATCC 393. *Virology* 339:42-55
21. Lombó F, Menéndez N, Salas JA, Méndez C (2006) The aureolic acid family of antitumor compounds: structure, mode of action, biosynthesis, and novel derivatives. *Appl Microbiol Biotechnol* 73:1-14
22. Madera C, García P, Rodríguez A, Suárez JE, Martínez B (2009) Prophage induction in *Lactococcus lactis* by the bacteriocin Lactococcin 972. *Int J Food Microbiol* 129:99-102
23. Martín R, Soberón N, Escobedo S, Suárez JE (2009) Bacteriophage induction versus vaginal homeostasis: role of H<sub>2</sub>O<sub>2</sub> in the selection of *Lactobacillus* defective prophages. *Int Microbiol* 12:131-136
24. Martín R, Soberón N, Vanechoutte M, Flórez AB, Vázquez F, Suárez JE (2008) Characterization of indigenous vaginal lactobacilli from healthy women as probiotic candidates. *Int Microbiol* 11:261-266
25. Martín R, Suárez JE (2010) Biosynthesis and degradation of H<sub>2</sub>O<sub>2</sub> by vaginal lactobacilli. *Appl Environ Microbiol* 76:400-405
26. McLean NW, Rosenstein IJ (2000) Characterization and selection of a *Lactobacillus* species to re-colonize the vagina of women with recurrent bacterial vaginosis. *J Med Microbiol* 49:543-552
27. Miller C, Thomsen LE, Gaggero C, Mosseri R, Ingmer H, Cohen SN (2004) SOS response induction by  $\beta$ -lactams and bacterial defense against antibiotic lethality. *Science* 305:1629-1631
28. Pavlova SI, Tao L (2000) Induction of vaginal *Lactobacillus* phages by the cigarette smoke chemical benzo[a]pyrene diol epoxide. *Mutat Res* 466:57-62
29. Ptashne M (1992). *A genetic switch*, 2nd ed. Cell Press and Blackwell Scientific Publications, Cambridge, MA, USA
30. Raya RR, Kleeman EG, Luchansky JB, Kleenhammer TR (1989) Characterization of the temperate bacteriophage jadH and plasmid trans-

- duction in *Lactobacillus acidophilus* ADH. *Appl Environ Microbiol* 55: 2206-2213
31. Redondo-López V, Cook RL, Sobel JD (1990) Emerging role of lactobacilli in the control and maintenance of the vaginal bacterial microflora. *Rev Infect Dis* 12:856-872
  32. Soberón N, Martín R, Suárez JE (2007) New method for evaluation of genotoxicity, based on the use of real-time PCR and lysogenic Gram-positive and Gram-negative bacteria. *Appl Environ Microbiol* 73: 2815-2819
  33. Swalla BM, Gumpert RI, Gardner JF (2003) Conservation of structure and function among tyrosine recombinases: homology-based modeling of the lambda integrase core-binding domain. *Nucleic Acids Res* 31: 805-818
  34. Tomasz M (1995) Mitomycin C: small, fast and deadly (but very selective). *Chem Biol* 2:575-579
  35. Ventura M, Canchaya C, Bernini V, et al. (2006) Comparative genomics and transcriptional analysis of prophages identified in the genomes of *Lactobacillus gasseri*, *Lactobacillus salivarius*, and *Lactobacillus casei*. *Appl Environ Microbiol* 72:3130-3146
  36. Von Gruenigen VE, Coleman RL, Li AJ, Heard MC, Miller DS, Hemsell DL (2000) Bacteriology and treatment of malodorous lower reproductive tract in gynecologic cancer patients. *Obstet Gynecol* 96:23-27
  37. Walker GC (1987) The SOS response of *Escherichia coli*. In: Ingraham JL, et al. (eds) *Escherichia coli and Salmonella typhimurium: cellular and molecular biology*. ASM Press, Washington, DC, pp 1346-1357
  38. Xu J, Hendrix RW, Duda RL (2004) Conserved translational frameshift in dsDNA bacteriophage tail assembly genes. *Mol Cell* 16:11-21
  39. Young R, Wang IN (2006) Phage lysis. In: Calendar R (ed) *The bacteriophages*, 2nd ed. Oxford University Press, New York, pp 104-125
  40. Zhou X, Bent S, Schneider M, Davis C, Islam M, Forney L (2004) Characterization of vaginal microbial communities in adult healthy women using cultivation-independent methods. *Microbiology* 150: 2565-2573

**Table S1.** Characteristics of bacteriophage Lv1 *orf*s and their deduced gene products

ORF	From	To	Length	aa	kDa (pI)	Predictive function	Closest hit (E value)	% aa identity (% similarity)	Accession number	Predicted domain (E value)
<b>1</b>	98	631	534	177	19749.47 (8.48)	Terminase small subunit	<i>Lactobacillus</i>	52% (77%)	YP_001843285	Terminase_4 (PF05119)(2.8e-25); Phage terminase, small subunit (TIGR1558)( 1e-18 )
							<i>fermentum</i> IFO 3956 (9e-41)			
<b>2</b>	631	2535	1905	634	74178.16 (5.86)	Terminase large subunit	<i>Lactobacillus</i>	67% (82%)	YP_001843286.1	Terminase_1 (PF03354) (2.2e-115); TMHMM transmembrane regions
							<i>fermentum</i> IFO 3956 (0.0)			
<b>3</b>	2545	3705	1161	386	42207.74 (4.85)	Phage portal protein	<i>Lactobacillus</i>	48% (72%)	YP_001843288.1	Phage portal protein, HK97 (TIGR01537) (4.4e-48); Phage portal protein (PF04860) (2.3e-56 )
							<i>fermentum</i> IFO 3956 (5e-97)			
<b>4</b>	3811	5658	1848	615	68706.56 (4.86)	Phage capsid protein	<i>Lactobacillus</i>	46% (63%)	YP_001843289.1	Peptidase U35, phage prohead HK97 (PF04586) (3e-48);
							<i>fermentum</i>			

										IFO 3956 (2e-98)	Proheadase_HK97: phage prohead protease, (TIGR01543) (6.9e-26); Phage major capsid protein, HK97 (PF05065) (5.8e-31); Phage major capsid protein (TIGR01554) (5.9e-41)
<b>5</b>	5680	5779	300	99	11100.59 (4.59)	Unknown					
<b>6</b>	5962	6333	372	123	13765.60 (8.76)	Head-tail joining protein	<i>Lactobacillus</i>	phage $\Phi$ AT3	30% (57%)	YP_025033.1	Phage head-tail joining (PF05521) (3.8e-07); Phage head-tail adaptor, putative (TIGR01563) (0.00025)
<b>7</b>	6284	6784	501	166	18279.54 (5.35)	Head-tail joining protein	<i>Lactobacillus</i>	phage $\Phi$ AT3	28% (45%)	YP_025034.1	Phage head/tail component (PF04883) (2e-05)
<b>8</b>	6777	7157	384	127	14689.42 (4.55)	Head-tail joining protein	<i>Lactobacillus</i>	phage $\Phi$ AT3	35% (54%)	YP_025035.1	
<b>9</b>	7160	7789	630	209	22264.79 (5.44)	Major tail protein	<i>Lactobacillus</i>	phage $\Phi$ adh	37% (55%)	NP_050156.1	Phage major tail protein (PF04630) (1.6e-22); Signalp signal-peptide

<b>10</b>	7875	8489	645	204	23216.39 (5.33)	Tail protein	<i>Lactobacillus</i> phage $\Phi$ adh (3e-07)	29% (44%)	NP_050157.1	
<b>11</b>	8663	13345	4683	1560	168134.97 (9.78)	Tape measure protein	<i>Lactobacillus</i> <i>rhamnosus</i> Lc Nu-like prophage (0.0)	30% (48%)	AAX07957.1	Lytic transglycosylase-like, catalytic (PF01464) (0.00088); Tape measure domain (TIGR02675) (4.2e-05)
<b>12</b>	13470	14072	603	200	23359.45 (5.25)	Tail protein	<i>Enterococcus</i> <i>faecalis</i> V583 (2e-14)	30% (49%)	NP_815672.1	
<b>13</b>	14069	16852	2784	927	100922.12 (4.95)	Fiber tail protein	<i>Lactobacillus</i> phage $\Phi$ adh (8e-53)	47% (68%),	NP_050162.1	
<b>14</b>	17024	18934	1914	637	70997.65 (7.52)	Minor capsid protein	<i>Lactobacillus</i> phage $\Phi$ g1e (1e-10)	31% (52%)	NP_695155.1	
<b>15</b>	18961	19425	465	154	16820.78 (4.57)	Unknown				
<b>16</b>	19576	19968	393	130	14608.35 (9.21)	Holin	<i>Lactobacillus</i> phage Lc-Nu	30% (53%)	YP_358776.1	Signalp signal-peptide; Transmembrane regions

								(2e-10)		
<b>17</b>	19987	20220	234	77	9214.87 (10.16)	Unknown				Signalp signal-peptide
<b>18</b>	20400	20792	393	130	14269.59 (6.59)	Unknown				Signalp signal-peptide; Transmembrane regions
<b>19</b>	20808	21929	1122	373	40957.01 (5.54)	Lysin (lysozyme)	<i>Lactococcus lactis</i> (3e-29)	32% (50%)	CAA10710.1	Glycoside hydrolase, family 25 (PD004620) (7e-19); Glyco_hydro_25 2. (PF01183) (3e- 15); Glycoside hydrolase, catalytic core (G3DSA:3.20.20.80) (2.6e-27)
<b>20</b>	22421	22645	225	74	8108.60 (9.87)	Unknown				
<b>21</b>	22891	23322	432	143	16990.89 (5.36)	Unknown				
<b>22</b>	23334	23525	192	63	7720.57 (9.22)	Unknown				
<b>23</b>	23607	24071	465	154	17782.38 (4.05)	Unknown				
<b>24</b>	25562	24456	1107	368	41823.25	Phage	<i>Lactobacillus salivarius</i>	35% (55%)	YP_535622.1	N-6 adenine-specific DNA methylase, conserved site

					(9.82)	integrase	UCC118 (9e-58)			(PS00092); Integrase, catalytic core, phage (PF00589) (5.5e-24); integrase-like, catalytic core, phage (G3DSA:1.10.443.10) (7e-31)
<b>25</b>	25979	25653	327	108	11934.51 (5.56)	Unknown				
<b>26</b>	26956	26297	660	219	24066.32 (5.13)	CI-like repressor	<i>Lactobacillus reuteri</i> 100-23(1e-41)	46% (61%)	ZP_01274778.1	Helix-turn-helix type 3 (PS50943) (9.844); Peptidase S24, S26A and S26B, C-terminal (G3DSA:2.10.109.10) (1.1e-20), (PF00717) (1.2e-14)
<b>27</b>	27124	27354	231	76	8671.15 (9.79)	Cro-like repressor	<i>Lactobacillus</i> phage Sal2 (1e-04)	41% (62%)	YP_535148.1	Helix-turn-helix type 3 (PF01381) (2.8e-13), (SM00530) (1.2e-10), (PS50943) (13.603)
<b>28</b>	27374	28201	828	275	30370.81 (9.19)	Antirepressor	<i>Lactobacillus</i> prophage Lj928 (2e-100)	75% (84%)	NP_958516.1	BRO, N-terminal (PF02498) (1e-37); Phage antirepressor protein (PF03374) (2.8e-72)
<b>29</b>	28979	28362	621	206	23298.98 (7.01)	Unknown				Transmembrane regions

<b>30</b>	29059	29328	273	90	10894.25 (5.78)	Unknown				
<b>31</b>	29641	30165	525	174	20066.09 (4.68)	Unknown				
<b>32</b>	30168	30500	330	110	13308.29 (9.42)	Unknown				
<b>33</b>	30517	30726	210	69	8068.69 (9.45)	Unknown				Signalp signal-peptide; Transmembrane regions
<b>34</b>	30840	31589	750	249	28983.03 (9.66)	DnaA	<i>Lactobacillus</i> phage LL-H (1e-38)	36% (54%)	ref YP_001285918.1	Phage conserved hypothetical protein, C-terminal (TIGR02220) (9e-39)
<b>35</b>	31592	32026	435	144	16839.02 (6.73)	Unknown				
<b>36</b>	32010	32819	813	270	30925.33 (8.55)	DnaC	<i>Lactobacillus</i> phage LL-H (3e-45)	37% (61%)	ref YP_001285919.1	AAA+ ATPase, core (SM00382) (1.8e-05)
<b>37</b>	32801	33292	492	163	18309.37 (4.94)	Single stranded DNA binding	<i>Lactobacillus</i> <i>sakei</i> subsp. <i>sakei</i> 23K (4e- 40)	48% (67%)	YP_394621.1	Primosome PriB/single-strand DNA- binding (PF00436) (3.3e-34), (PS50935); Single-strand DNA- binding (PTHR10302) (9.7e-28) (TIGR00621) (2.5e-43); Nucleic

						protein				acid-binding, OB-fold (G3DSA:2.40.50.140) (1.1e-36)
<b>38</b>	33246	33908	663	220	25493.10 (9.11)	Unknown				
<b>39</b>	34024	34317	294	97	11079.84 (7.71)	RusA resolvase	<i>Lactobacillus reuteri</i> 100-23 (2e-25)	42% (68%)	ZP_01273467.1	Endodeoxyribonuclease, RusA-like (PF05866) (1.1e-11)
<b>40</b>	34314	34625	312	103	11863.78 (7.64)	Unknown				
<b>41</b>	34633	35415	783	260	30623.38 (7.70)	Unknown				Transmembrane regions
<b>42</b>	35405	35647	243	80	9243.22 (9.72)	Unknown				Transmembrane regions
<b>43</b>	35631	35840	210	69	8290.39 (4.98)	Unknown				
<b>44</b>	35844	36317	474	157	18007.66 (8.57)	Arp-U-like protein	<i>Lactobacillus reuteri</i> 100-23 (1e-14)	30% (55%)	ZP_01274484.1	Phage transcriptional activator ArpU (TIGR01637) (3e-21)
<b>45</b>	36223	36627	405	134	15488.60 (4.51)	Unknown				

<b>46</b>	36509	36697	189	62	7196.34 (10.42)	Unknown				
<b>47</b>	37292	38419	1128	375	43678.20 (8.83)	DNA methylase	<i>Lactobacillus</i> phage Lc-Nu (7e-48)	43% (62%)	YP_358799.1	
<b>48</b>	38428	38892	465	154	18256.08 (9.59)	HNH nuclease	<i>Lactobacillus</i> phage Lc-Nu (1e-13)	35% (50%),	YP_358809.1	HNH endonuclease (PF01844) (0.00036); HNH nuclease (SM00507) (0.033)

---