AEM Accepts, published online ahead of print on 12 September 2014
Appl. Environ. Microbiol. doi:10.1128/AEM.02771-14

Exposing the secrets of two well known *Lactobacillus casei* phages: Genomic and structural analysis of J-1 and PL-1

Running title: Lactobacillus phages J-1 and PL-1

Maria Eugenia Dieterle[a,b], Charles Bowman[b], Carlos Batthyany[c] , Esteban Lanzarotti[a], Adrián Turjanski[a], Graham Hatfull[b] and Mariana Piuri[1a#]

[a]Departamento de Química Biológica, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, IQUIBICEN-CONICET, Buenos Aires, Argentina.
[b]Department of Biological Sciences and Pittsburgh Bacteriophage Institute, University of Pittsburgh, Pittsburgh, USA.
[c]Unidad de Bioquímica y Proteómica Analíticas, Institut Pasteur de Montevideo, Montevideo, Uruguay.

Address correspondence to Mariana Piuri, mpiuri@qb.fcen.uba.ar

Abstract

Bacteriophage J-1 was isolated in 1965 from an abnormal fermentation of Yakult using the strain *Lactobacillus casei* “Shirota”, and a related phage, PL-1, was subsequently recovered from a strain resistant to J-1. Complete genome sequencing shows that J-1 and PL-1 are almost identical, but PL-1 has a deletion of 1.9 kbp relative to J-1, resulting in loss of four predicted gene products involved in immunity regulation. The structural proteins were identified by mass spectrometry analysis. Similarly to phage A2, two capsid proteins are generated by a translational frameshift and undergo proteolytic processing. The structure of gp16, a putative tail protein, was modeled based on the crystal structure of baseplate distal tail proteins (Dit) that form the baseplate hub in other *Siphoviridae*. However, two regions of the C-terminus of gp16 could not be modeled using this template. The first region accounts for the differences between J-1 and PL-1 gp16 and showed sequence similarity to carbohydrate-binding modules (CBMs). J-1 and PL-1 g*fp-gp16* fusions bind specifically to *Lactobacillus casei/paracasei* cells, and addition of L-rhamnose inhibits binding. Phage adsorption inhibition assays revealed a higher affinity of J-1 gp16 for cell walls of *L. casei* subsp. *casei* ATCC 27139 in agreement with differential adsorption kinetics observed for both phages in this strain. The data presented here provide insights into how *Lactobacillus* phages interact with their hosts at the first steps of infection.

41 **Introduction**

42 Biotechnological processes employing bacterial cultures are subject to interference by
43 infection with bacteriophages (1). The deleterious impact of phages in fermentations is
44 common in the dairy industry where the organoleptic properties of cheese and other dairy
45 products rely on the use of starter cultures with combinations of specific strains. Phage
46 infection of these carefully selected bacteria can substantially delay the process, alter the
47 quality of the product, and in the worst-case scenario abort of the entire fermentation batch
48 (2).

49

50 Contamination in the dairy industry is common as phages can reside in the substrates for
51 fermentation and on surfaces of vessels and other equipment (3-6). Strategies to control
52 this contamination include milk pasteurization and sanitation of equipment (6-9), but
53 phages are rarely completely eliminated. Cultures of starter strains themselves can be
54 phage contaminated, likely due to induction of resident prophages and concomitant cell
55 lysis (10-13). In general, it is helpful to avoid lysogenic Lactic Acid Bacterium (LAB), but
56 the nature of starter strains is not always well defined, and in some strains multiple
57 prophages may be present (14). Rotation of starter strains generally helps to minimize the
58 negative impact of phages on fermentation (1). The increasing number of phage and
59 bacterial genomes sequenced has helped to elucidate phage-host interactions and the
60 engineering of resistant strains that are more robust during fermentation (15, 16).
61 *Lactobacillus casei* is widely employed in fermentation of vegetables, meat and dairy
62 products – including cheese and fermented milk. This bacterium is of interest because in
63 addition to its organoleptic properties in fermentation, some strains may have probiotic
64 properties (17, 18). *Lactobacillus casei* can tolerate the low pH of the stomach and
65 colonize the gastrointestinal tract (GI) (19), with potential beneficial outcomes (20, 21).

Forty-three phages that infect *L. casei* have been reported, sixteen of which morphologically belong to the *Myoviridae*, twenty-one are *Siphoviridae*, and the other six have not been morphologically classified (22). Among *L. casei* phages, A2 and phiAT3 are the best characterized and the complete genome sequences have been determined. A2 was isolated in Spain from the whey of a failed homemade blue cheese product using *L. casei* 393, and phiAT3 was recovered following induction from *L. casei* 393 using mitomycin C (23, 24). Both phages belong to the *Siphoviridae* family and are temperate, although they share little nucleotide sequence similarity (23, 24).

Among the other phages of *L. casei*, J-1 was isolated in 1965 in association with fermentation failures during the production of Yakult, a Japanese beverage fermented from skimmed milk and Chlorella extracts (25). Upon isolation, J-1 was shown to behave as a virulent phage in the *L. casei* "Shirota" strain used for manufacturing of Yakult. A *L. casei* strain resistant to J-1 infection was isolated, but after two years of use a second phage designated PL-1 was isolated that infects the resistant strain (26). J-1 and PL-1 were characterized extensively [reviewed in Sechaud *et al*. (27)] and shown to be serologically related (28-30). Conditions for transfection of protoplasts with phage DNA (31-34), phage inactivation (35-38), and characterization of DNA content (39, 40) have been reported. The lysis cassette of PL-1 encoding a holin and an endolysin – a *N*-acetylmuramoyl-L-alanine amidase – has been characterized (41), but the complete genome sequences of J-1 and PL-1 were only recently reported (42).

Here we describe the genomic and structural features of phages J-1 and PL-1. Interestingly PL-1 and J-1 genomes are almost identical and among sequenced phages

are most closely related to phage A2. Compared to J-1, PL-1 has a deletion of 1.9 kbp comprising 4 genes associated with the immunity cassette. PL-1 also differs from J-1 in a gene (*16*) that has a predicted structure similar to distal tail proteins that form the baseplate hub in other *Siphoviridae*. J-1 gp16 recognizes sugar motifs and specifically binds to *Lactobacillus casei/paracasei* cells blocking phage adsorption, suggesting that gp16 is involved in host recognition.

Material and Methods

Strains, Bacteriophages and Growth Conditions

Lactobacillus paracasei subsp. *paracasei* ATCC 27092 and *Lactobacillus casei* subsp. *casei* ATCC 27139 were grown in MRS medium (Difco, USA) at 37°C under static conditions. *E. coli* DH5α was used for cloning, and *E. coli* BL21(DE3) [pLysS] (Invitrogen, USA) was used for protein expression. *E. coli* strains were grown in LB-Broth or nutritive medium (Difco, USA) at 37°C under moderate shaking. When appropriate, antibiotics were added at the following concentrations: kanamycin (15 µg/ml) (Sigma, USA) and chloramphenicol (34 µg/ml) (Sigma, USA) for *E. coli*.

Lactobacillus phages J-1 and PL-1 used in this study were a kind gift of Dr. Quiberoni from INLAIN (Instituto de Lactología Industrial). Phage J-1 was propagated on *L. casei* subsp. *casei* ATCC 27139 and PL-1 on *L. paracasei* subsp. *paracasei* ATCC 27092 in MRS broth. Bacteriophage stocks were stored at 4°C in Phage Buffer (20 mM Tris-HCl, 100mM NaCl and 10 mM $MgSO_4$).

Electron Microscopy

FIve µl of *Lactobacillus* phage J-1 and *Lactobacillus* phage PL-1 lysates (10^{11} PFU/ml) were allowed to sit on freshly glow-discharged 400-mesh carbon-Formvar-coated copper grids for approximately 15 seconds. The grids were then rinsed with distilled water and stained with 1% uranyl acetate. Virus particles were imaged on an FEI Morgagni transmission electron microscope at 80 kV at a magnification of 56,000.

Identification of PL-1 and J-1 proteins

Approximately 50 µl (a total of 5 x 10^{10} PFU) of CsCl purified J-1 and PL-1 particles were collected by centrifugation at 20000g for 30 min. The pellet was resuspended in 37.5 µl of distilled water, frozen at -70°C and then mixed by vortexing. This cycle was repeated three times and the solution was then heated to 75°C for 4 minutes. DNAse (2U) (Fermentas, USA) was added and incubated for 30 minutes to reduce viscosity. Finally, 4X SDS sample buffer was added and boiled for 2.5 min. Approximately 25 µl were loaded onto a 12% SDS–polyacrylamide gel and electrophoresed at 90 V until the dye ran off the gel. The gel was stained with Coomassie blue. The visible bands were compared to a protein standard to determine the approximate molecular mass.

For protein identification by mass spectrometry (MS), the same protocol described above was applied but the gel was stained with colloidal Coomassie blue. The protein bands were excised and digested *in situ* with trypsin or Lys-C followed by peptide elution, chromatography, and tandem mass spectrometry (MS/MS) on an LTQ Velos Orbitrap mass spectrometer and MALDI TOF MS. Peptides were matched against predicted J-1 and PL-1 phage proteins.

Computational Methods

A *Lactobacillus* phages database was created and genome maps diagrams were drawn using Phamerator (43), with the threshold parameters of 32.5% identity with ClustalW and a BlastP E-value of 10^{-50}, as described previously.

Protein domains in gp16 and gp17 of J-1 and PL-1 were detected using Pfam server (44). In gp16 a Sipho_tail domain (PF05709) was detected achieving E-values of 1.1e-72 and 4.6e-72 for J-1 and PL-1 respectively. The template for the comparative modeling of the Sipho_tail domain structure was the PDB file 2x8k (chain A), which belongs to this family according to Pfam, and was retrieved by BLAST. To align gp16 with the template sequence we used the HMMER software and the HMM profile corresponding to Sipho_tail family. For both J-1 and PL-1, two regions (Dom1 and Dom2) were not covered with this template structure corresponding to an insertion of high probability in the Pfam HMM logo. Dom1 and Dom2 were submitted separately to the HHPRED server (45) in order to detect known structures with remote homology. Dom1 resulted in a significant hit against a carbohydrate binding module (PDB chain: 1xc1A) with 98.4 and 93.2 probability values for J-1 and PL-1 respectively. In the case of J-1, alignment presents an E-value of 4.1E-06, a P-value=1.3E-10, an identity of 15% and a coverage of 80%. For PL-1, Dom1, achieved an E-value=1.1, a P-value=3.5E-05, an identity of 14% and a coverage of 61%. Dom2 does not result in any significant hit for both J-1 and PL-1, so no further analysis was done. No hits were found when submitting gp17 of J-1 and PL-1 to the Pfam server. Both sequences were further submitted to the HHPred server achieving a high probability hit (97.8%) against chain A in the PDB 1wru (E-value=0.0055 and Pvalue=1.7E-07), a sequence that belongs to the Prophage_tail (PF06605) family. The corresponding HMM profile was used to align gp17 sequences with the template structure sequence. Sequence identities between template structure (1wruA) and both PL-1 and J-1 in the context of Pfam family (Prophage_tail) only cover about 40% of the sequences for the two of them

achieving 7% identity each. In this case, HHPred alignment was preferred to use for comparative modeling, which presents 8% identity and 35% coverage. MODELLER (46) was used to build models based on mentioned alignments over the aligned regions. For each target sequence, 10 different models were built, and their quality measures were assigned using the GA341 when globular domains were analyzed.

Phage adsorption assays

For kinetics of adsorption assays *L. casei* subsp. *casei* ATCC 27139 or *L. paracasei* subsp. *paracasei* ATCC 27092 were grown till an OD_{600} of 1 and 150 µl of cells were infected with 150 µl of J-1 or PL-1 at a multiplicity of infection of 0.005. One tube per time point was prepared and phages were allowed to adsorb from 5 minutes to one hour at 37°C in the presence of 10mM $CaCl_2$. Cells were removed by centrifugation and unadsorbed phages in the supernatant were measured using the double agar method. Cell wall purification was done essentially as previously described (47). Briefly, bacterial cells were centrifuged at 3200g and resuspended in 50mM Tris-HCl pH 7.5. Cells were disrupted by sonication in the presence of glass beads (10 X 20 s followed by intervals of 1 minute on ice). Whole extracts were incubated with DNAse (30µg/ml) (Thermo Scientific, USA) and RNAse (5µg/ml) (Thermo Scientific, USA) at 37°C for 1h and non-lysed cells and debris were removed by a short centrifugation step (1500g for 5 min). Cell walls were recovered from the supernatant by centrifugation at 20000g for 20 min and the pellet washed three times with distilled water and stored at -20 °C. One part of this sample was lyophilized and the weight (µg) was measured. Phage adsorption assays using cell walls were done as described above but using 100 µg of cell walls instead of whole cells.

Molecular cloning

189 DNA sequences were amplified by PCR using Go Taq DNA polymerase (Promega, USA)
190 following manufacturer's instructions. A N-terminal GFP fusion expression vector was
191 constructed amplifying *Egfp* from pMP14 (48) with the following primers: ED74
192 (tgactCATATGgtgagcaagggcgaggagc) and ED75 (taccaGGATCCcttgtacagctcgtccatgcc).
193 The PCR product was digested with NdeI and BamHI and cloned into pET28b (Novagen,
194 Merck Millipore, USA) to render pet28-GFP. Gene *16* and *17* were amplified using J-1 or
195 PL-1 DNA as templates. Gene *16* was amplified using ED76
196 (tagcaGAATTCgcaaatttaatatttggagg) and ED77 (tgaacGAGCTCtcatagccatgcctcct) and
197 gene *17* using ED78 (gctaaGAATTCaaggatttttattttgtgga) and ED79
198 (gatccAAGCTTctaaactgcgtatacctca). Amplicons were digested with the appropriate
199 restriction enzymes EcoRI/SacI or EcoRI/HindIII (Promega, USA) and cloned into pET28-
200 GFP. Plasmids were named pet28- J-1 GFP-gp16, pet28- PL-1 GFP-gp16, pet28- J-1
201 GFP-gp17 and pet28- PL-1 GFP-gp17.
202
203 **Protein purification**
204 Plasmids described above were transformed into *E. coli* BL21 (DE3) cells (Invitrogen,
205 USA) by electroporation for protein expression and further purification. Transformed cells
206 were grown at 37°C to an OD_{600} of 0.5 in nutrient broth and expression was induced by
207 addition of 0.5 mM IPTG. Cells were left overnight at 21°C before harvesting. Cell pellets
208 were resuspended in binding buffer (50 mM Tris-HCl pH 8, 300 mM NaCl, and 1 mM
209 PMSF) and disrupted by sonication (4 X 15 s). The lysate was centrifuged at 16000 g for
210 20 min, and the supernatants incubated with 300 µl of pre-equilibrated Ni-Agarose resin
211 (Novagen, USA) for 2 h at 4°C. The matrix was washed with 10 volumes of binding buffer,
212 10 volumes of the same buffer containing 25 mM imidazole and 5 volumes containing 60
213 mM imidazole. Elution was done with 4 volumes of buffer with 120 mM imidazole. Eluted

samples were dialyzed twice against protein buffer (50 mM Tris- HCl pH 8, 150 mM NaCl, 1mM DTT), a last time in the same buffer but containing 30% glycerol and stored at -20°C.

gp16 fluorescence binding assays

Cell binding assays using purified J-1 GFP-gp16 and PL-1 GFP-gp16 and GFP proteins were carried out as described before (49) with some modifications. Briefly, 0.3 ml of exponentially growing bacterial cells were centrifuged and resuspended in 100 µl of modified phage buffer (20 mM Tris-HCl, 100mM NaCl and 10 mM $MgSO_4$, 0,1% Tween 20, 10 mM $CaCl_2$) and incubated with 1 µg of protein fusions for 20 minutes at room temperature. Cells were washed twice with PBS buffer and binding to the bacterial cells was detected by fluorescence microscopy (Axiostar Plus; Carl Zeiss) with a 100X objective with oil immersion, and phase contrast. When binding in the presence of sugars was tested, J-1 GFP-gp16 and PL-1 GFP-gp16 were preincubated with 0.25 M L- rhamnose, D- glucose or N-acetyl glucosamine for 30 minutes at room temperature and the protocol was followed as described above.

Inhibition of adsorption assays

The inhibition of adsorption assay is an adaptation of the adsorption assay using cell walls. In brief, 50 µl of cell walls (100 µg) were incubated with, 50 µl of buffer (control), J-1 GFP-gp16, PL-1 GFP-gp16 or GFP at different concentrations at room temperature for 30 minutes. Then, 50 µl of phage was added ($1x10^6$ pfu/ml). The mixture was incubated at 37°C for 1 hour and cell walls were removed by centrifugation at 3200g for 10 minutes. The unadsorbed phage in the supernatant was measured using the double agar method. To test inhibition of adsorption by sugars the same protocol was followed using the indicated carbohydrates instead of proteins. All sugars were used at 0.25 M.

239 **Nucleotide sequence accession number**

240 GenBank accession number for J-1 is KC171646 and for PL-1 is KC171647.

241

242 **Results**

243

244 **Genome sequencing of bacteriophages J-1 and PL-1**

245 The J-1 and PL-1 genome sequences were determined by pyrosequencing. J-1 and PL-1
246 virion DNAs are 40,931 bp and 38,880 bp long respectively, and their G+C contents are
247 44.8 and 44.9 % respectively. Both phages have cohesive ends with10 base, single-strand
248 3'-extensions (left end; 3'- CGGTCGGCCT), 4 bases shorter than the
249 GAACGGTCGGCCTC sequence previously described by Nakashima *et al*. (50). Dotplot
250 analysis shows that J-1 and PL-1 are closely related (Fig. 1), with PL-1 having a 1.9 kbp
251 deletion corresponding to coordinates 23,513 to 25,418 in J-1.

252

253 **Organization of J-1 and PL-1 genomes**

254 Analysis of the J-1 genome revealed 63 potential ORFs and no tRNA genes. The most
255 commonly used start codon is ATG (73.5%) with lower usage of GTG (17%) and TTG
256 (9.5%). In J-1, 57 ORFs are transcribed rightwards, and 6 leftwards, and the deletion in
257 PL-1 results in loss or alteration of four leftwards transcribed genes, *27* to *30* (Fig. 2A).
258 The genome organizations can be divided into the following modules: DNA packaging,
259 virion structure, lysis, integration, immunity, and replication. The presence of integrase
260 genes (*24*) of the tyrosine recombinase family suggests a temperate origin of these
261 phages.

262

Nucleotide sequence comparison with other LAB phages shows that J-1 and PL-1 are most closely related to phages A2 (23, 51-54) and Lrm1, especially in the DNA packaging and virion structure modules (Fig. 2B). Because the J-1 and PL-1 genomes are similar, we will describe the functional assignments for J-1 and indicate where they differ for PL-1 (Table 1).

DNA packaging and virion structure

The DNA packaging module of J-1 contains genes *1* and *2,* which are predicted to encode the terminase small and large subunits respectively. J-1 gp1 is similar to the terminase small subunit of phage Lrm1 (ORF1) (12), but the organization differs to that of phage A2 where the terminase small subunit is encoded by gene *61*, and biochemical evidence supports this functional assignment (55) (23). In J-1 and PL-1 a single *orf* at the right end of the genomes (genes *63* and *59* respectively) corresponds to a fusion of A2 *orf60* and *orf61*, similar to that of orf54 in phage Lrm1 (12). It is unclear whether J-1 gp63 and PL-1 gp59 play a role in DNA packaging in addition to gp1 and gp2.

Examination by electron microscopy shows that J-1 and PL-1 have siphovirial morphologies. J-1 has an isometric head of 65 ± 4 nm in diameter and a tail length of 283 ± 8 nm. PL-1 also has an isometric head of 62 ± 4 nm in diameter and a tail length of 290 ± 7 nm, and a baseplate structure at the tip of the tail (Fig. 3); this is consistent with prior reports (26). These reflect the most frequently found morphologies among LAB phages (22).

Genes *4* and *5* code for putative portal and protease proteins, respectively, and are closely

related to the corresponding genes in phage A2. The major capsid gene is organized similarly to that in A2, where two types of capsid subunits are expressed; the product of gene *5* (gp5A) and a C-terminally 85 aa extended form (gp5B) resulting from a -1 programmed translational frameshift at the 3' end of gene *5*. Both proteins appear to be essential for phage viability (53). In J-1 and PL-1 a slippery sequence (CCCAAAA) is present at the end of gene *6* and a -1 frameshift can facilitate expression of a longer form (gp7), 86 amino acids longer than gp6 (Fig. 4A). Expression of gp6 and gp7 is supported by SDS-PAGE of virion proteins (Fig. 3) and mass spectrometry (ms) (Table 2), including peptides unique to gp7 (Fig. 4A). Quantification suggests that the ratio of gp6:gp7 is 2.5:1 (Fig. 3). J-1 gp6 and gp7 are also proteolytically processed involving cleavage of the N-terminal 123 residues, similarly to that reported for the A2 capsid subunit. This is supported by the observation that cleavage with Lys-C (which cleaves after lysine residues) prior to MS/MS generates a peptide AVPTDAS reflecting cleavage following an arginine residue, and correspondence to the sequence (N-AVPTADAS) of the mature form of the A2 capsid. The 123-residue peptide presumably acts as a scaffold for capsid assembly (56).

J-1 genes *8-11* are organized similarly to the putative head-tail connector protein genes of both A2 and have partial sequence similarity to *Staphylococcus aureus* phage PVL head-tail connectors (23). We presume that they provide similar functions in J-1 but we did not identify the products by ms/ms analysis of intact virions, presumably because they are absent or present in low abundance.

Phages belonging to the *Siphoviridae* family have long flexible tails containing many copies of the major tail subunit. J-1 gp12 has similarity to the major tail protein of A2 (*orf10*

312 (98% identity), but J-1 and PL-1 lack the putative frameshifting sequence (CCCAAAA)
313 present in the major tail subunit genes of phages A2 and Lrm1 (12, 57). J-1 genes *13* and
314 *14* likely encode tail assembly chaperones expressed *via* a highly conserved (58)
315 programmed translational frameshift (5'- AAAAAAAT, Fig. 4B). These assembly
316 chaperones are not expected to be virion components but were detected by ms, perhaps
317 as contaminants in the phage lysates. Peptides identified are consistent with frameshifting
318 at the predicted sequence. J-1 gp15 is the putative tape measure protein based on its
319 position and size (4,902 bp), as well as similarity with the putative tape measure proteins
320 of Lrm1 and A2 phages. A large number of peptides derived from this protein were
321 identified by ms (Table 2) and a product of the expected size was observed by SDS-PAGE
322 (Fig. 3).
323
324 A notable difference between J-1 and PL-1 is found in genes *16* and *17*. gp16 of J-1 and
325 PL-1 were detected by ms analysis and unique peptides for each protein were identified
326 (Fig. 3 and Table 2). gp16 has similarity to A2 gp13 and several genes annotated as
327 putative phage tail protein genes in the genomes of several *Lactobacillus* strains. Multiple
328 sequence alignment revealed three highly conserved regions, and two variable regions
329 (Fig. 5A). The N-terminus of gp16 is identical in J-1 and PL-1 (residues 1 to 146) and
330 conserved in the other analyzed phages. After that, a region of 171 aa's in J-1 and 122
331 aa's in PL-1 with low identity between each other is found. Divergent sequences were also
332 identified in Lrm1, Lc-Nu, A2 and phiAT3. A short highly conserved region follows; then a
333 237 aa region similar between J-1 and PL-1 but with lower identity among the other
334 phages was found and finally at the C-terminal the gene products are highly similar (Fig.
335 5A). A Sipho_tail Pfam domain that encompasses the whole gene *16* of J-1 and PL-1 was
336 identified. However, the Sipho_tail HMM logo presents two long insertions with high

337 probability. Further analysis by HHPred revealed similarity between the N-terminus of
338 gp16 and the distal tail proteins (Dit) of *Bacillus subtilis* SPP1 (59) and *Lactococcus lactis*
339 TP901-1 (60, 61) J-1 gp16 may thus act like Dit in providing a hub for anchoring the tail-
340 tube, tail-spike and baseplate (62). Modeling of the J-1 gp16 N-terminal domain on SPP1
341 Dit (PDB code 2x8k_A) suggests that they fold into very similar structures (Fig. 5B). The
342 first of the variable regions in gp16 (Dom1; ~120 aa) is a predicted carbohydrate- binding
343 module (CBM) present in enzymes that depolymerize plant cell wall polysaccharides into
344 simple sugars (Fig. 5C) to increase the catalytic efficiency by targeting the enzymes to its
345 substrate (63). J-1 and PL-1 gp16 Dom1 was modeled on the endo-β -1,4-galactanase
346 from *Thermotoga maritima* (PDB code 2xon: chain L), revealing putative structural
347 differences between the J-1 and PL-1 proteins that could reflect recognition of distinctive
348 sugar motifs or a differential sugar binding affinity (Fig. 5C). The second variable region
349 comprises 237 aa's (Dom2) (Fig. 5B), and lacks a suitable template in the PDB for
350 modeling. The predicted secondary structure contained mostly B sheets resembling
351 galectin.
352
353 J-1 and PL-1 gp17 have similarity to proteins of phages Lrm1 and A2 that have been
354 annotated as host specificity proteins – as well as several host-encoded proteins
355 (presumably prophage-encoded in *Lactobacillus* and *Streptococcus*) annotated as phage-
356 related tail-host interaction proteins. The J-1 and PL-1 gp17 proteins share 98% identity
357 and their N-termini is shared with related proteins of other *Lactobacillus* phages (Fig. 6A)
358 and contain the Prophage_tail Pfam family. HHPred analysis revealed that residues 1-400
359 share similarity to gp44 of phage Mu and the structure was modeled using this as a
360 template (PDB code 1WRU). The predicted structure of the N- terminal region can be
361 superimposed on Mu gp44 (Fig. 6B) and is also similar to both T4 gp27 (PDB 1K28) and

362 *Lactococcus* phage p2 Orf16 (PDB 2WZP and 2X53). These proteins, assemble as
363 identical trimers, but can adopt different architectural arrangements (62). Although the
364 differences between J-1 and PL-1 gp17 (19 residues in total) are all in the N-terminal
365 region, the predicted structures are near-identical. The rest of gp17 resembles host
366 recognition proteins in *Streptoccocus termophilus* phages (64, 65) and contain five
367 collagen-like repeats (Gly-X-Y; Fig. 6A). The numbers of Gly-X-Y repeats differs in J-1, PL-
368 1 and their homologues, perhaps through recombination or replication errors (64, 66).
369 Taken together, the analyses of gp16 and gp17 suggest that are tail components involved
370 in host recognition, and the variation in the gp16 proteins and their homologues contribute
371 to host specificity.
372
373 **J-1 and PL-1 adsorption assays**
374 While propagating J-1 and PL-1 we observed that although J-1 forms large (~ 2 mm)
375 plaques on lawns of both *L. casei* subsp. *casei* ATCC 27139 (host to J-1) and *L. paracasei*
376 subsp. *paracasei* ATCC 27092, PL-1 forms similarly sized plaques (~1.5 mm) on lawns of
377 27092 strain but tiny pinprick plaques on 27139, so we decided to use these two strains for
378 further experiments. Both J-1 and PL-1 absorb with similar kinetics to *L. paracasei* subsp.
379 *paracasei* ATCC 27092, with more than 90% of particles adsorbed after five minutes of
380 incubation (Fig. 7A). However, the adsorption kinetics for both phages is different with *L.*
381 *casei* subsp. *casei* ATCC 27139. Only 60% of particles were adsorbed after 5 minutes,
382 and adsorption of all of the particles required 30 minutes incubation for J-1 and 60 minutes
383 for PL-1 (Fig. 7B). These results suggest that even though J-1 and PL-1 are very similar;
384 the small structural differences between each other account for the disparities observed in
385 this first step of infection.
386

Gp16 fluorescence binding assays

To test the roles of gp16 and gp17 in adsorption, we constructed and expressed gfp fusions of both proteins encoded by J-1 and PL-1. Both gp17-gp fusions were expressed but insoluble and were not studied further. Whole cells were incubated with the recombinant proteins and then visualized by fluorescence microscopy (Fig. 8). The fluorescent images showed that J-1 gfp-gp16 and PL-1 gfp-gp16 could bind to the cell surface of *L. casei* subsp. *casei* ATCC 27139 (Fig. 8a-b) as well as *L. paracasei* subsp. *paracasei* ATCC 27092 (Fig. 8c-d**)**. Fluorescence was dependent on gp16 binding and specific to the strains tested in that no signal above the background level was observed with *Lactobacillus acidophilus* (Fig. 8e-f) or with gfp alone (not shown). The uniform distribution of fluorescence suggests the ligands for gp16 binding are not localized in any particular part of the cell but regularly distributed on the cell surface. These observations are consistent with J-1 and PL-1 recognizing saccharide-containing receptors within the outer layer of the cell wall, including a role for L- rhamnose as noted previously (67-69).

Examination of the effect of addition of different monosaccharides to the gfp-gp16 binding assay is consistent with L-rhamnose being an important component of the receptor. Of all the sugars tested only L-rhamnose showed strong interference with gp16 binding (Fig. 8 g-j) and also reduced adsorption of whole phage particles by over 80% (Fig. 9). Similar results were observed with *L. paracasei* subsp. *paracasei* ATCC 27092 (data not shown). We also tested whether the gfp-gp16 fusions were able to specifically interfere with phage adsorption. We observed concentration-dependent inhibition of adsorption of both J-1 and PL-1 to *L. casei* subsp. *casei* ATCC 27139 although maximal adsorption inhibition of J-1 was 48% in the presence of J-1 gp16 but only 20% if PL-1 gp16 was used as competitor (Fig. 10A). When adsorption inhibition of PL-1 was tested, these values were 39% and

412 25% for J-1 gp16 and PL-1 gp16 respectively (Fig. 10B). These observations are
413 consistent with the sequence variations between J-1 and PL-1 gp16 being sufficient for a
414 differential binding affinity to the cell surface and playing a role in host specificity.
415
416 **Lysis**
417 The lysis cassette follows the virion structural genes, with gp22 being the putative holin
418 containing a predicted signal sequence, two putative transmembrane helices, and a highly
419 charged C-terminus. J-1 gp23 is the endolysin and N-acetylmuramoyl-L-alanine amidase
420 activity to hydrolyze the amide linkage in the peptidoglycan of *L. casei* as has been shown
421 previously (41).
422
423 **Integration and immunity**
424 J-1 gene *24* is identical in nucleotide sequence to a putative tyrosine-integrase present in
425 the genome of *L. casei* BL23 (LCABL_10790), and gp24 is also related at the amino acid
426 sequence level to the phage Lrm1 integrase. At the 5' side of the integrase gene there is a
427 region of approximately 221 bp with no coding potential and is the likely location of *attP*.
428 Comparison of this region using BLASTN revealed a short segment of sequence identity
429 (49/50) in an intergenic region between an endolysin gene and a putative uncharacterized
430 protein in *L. casei* BL23, BDII, LC2W and ATCC 334 strains. This sequence is partially
431 present in phages AT3, FSW and Lrm1, temperate phages induced from *L. casei* ATCC
432 393, *L. casei* subsp *casei* ATCC 27139 and *L. rhamnosus* strain M1 respectively (12, 24,
433 70). We were not able to assign any function to genes *25* and *26* and there are no
434 homologous phage proteins to the correspondent gene products in the database.
435

Genes *27* and *28* of J-1 are transcribed leftwards and are similar to *orfs 21* and *22* of phage A2. It is noteworthy that the deletion of 1.9 kbp in PL-1 genome spans positions 23,513 – 25,418 in J-1 and includes J-1 genes *27* to *30* (see below). The N-terminus of J-1 gp27 shares similarity to gp21 of phage A2 which is predicted to be an excise, although this has not been experimentally validated. J-1 gp28 is similar to ORF25 of Lrm1 and these are related to a negative regulator of the *tcd* operon (TcdC) found in several Gram-positive bacteria. This regulator has been extensively studied in *C. difficile* where it regulates expression of toxin genes *tcdA* and *tcdB* (71-73).

Using HHpred and CDD (Conserved protein domain database) of NCBI (74) it was found that the products of gene *29* of J-1, that is transcribed leftwards and of gene *31* (gene *27* in PL-1) that is transcribed rightwards, have maximal homology to CI and Cro proteins of lambda (75, 76). It is plausible that this organization corresponds to the synteny found in other lambda like phages where CI is coded from right to left and Cro in the opposite direction. Genes *30* and *32* of J-1 (*28* in PL-1) have no homology to any known sequence at the DNA level but a HTH (helix turn helix) domain could be recognized, indicating that probably act as transcriptional regulators.

Replication module

Genes *37* to *41* (*33* to *37* in PL-1) are predicted to be involved in DNA replication and recombination. J-1 gp37 (gp33 of PL-1) encodes an Erf–family protein involved in DNA single-strand annealing (SSAPs) (77), and gp39 (gp35 in PL-1) is a putative ssDNA binding protein (78). J-1 genes *40* to *44* in (*37* to *40* in PL-1) are similar to *orfs 34* to *38* of phage Lc-Nu and likely perform similar functions. Gp40 (Gp36 of PL-1) has a HTH binding domain at the N-terminus and is similar to several putative replication proteins. In gene *40*, a 19 bp AT rich region was observed flanked by several direct and inverted repeats,

features common to phage replication origins (79). Upstream of this sequence, three directed repeats comprised of another direct and inverted repeats were found. The AT rich region is followed by inverted repeats capable of forming a stem-loop structure similar to the one found in phage Lc-Nu (80). J-1 gp41 (gp37 in PL-1) is similar to DnaB helicases (81) and gp43 (gp39 in PL-1) is a putative DNA binding protein. J-1 gp44 (gp40 in PL-1) is a putative RusA-like Holliday Junction resolvase (82).

Other genes at the extreme right end of the genome code for small proteins mostly of unknown function, although J-1 *47* (*43* in PL-1) encodes a putative cytosine DNA methyltransferase. J-1 genes *54* and *55* (*50* and *51* in PL-1) both encode potential transcriptional regulators (83). J-1 gene *59* (*55* in PL-1) codes for a putative HNH endonuclease, and the possible function of the product of gene *63* (*59* in PL-1) has already been described.

Discussion

We described here the genomic and structural analysis of *Lactobacillus* phages J-1 and PL-1. Although both phages were isolated about 50 years ago and extensively studied (specially PL-1), their genome sequences have only recently become publicly available (42). J-1 was isolated from a failed fermentation during manufacture of Yakult (25) and PL-1 was isolated subsequently when using a derived strain resistant to J-1 (26, 27). Both phages differ by only four gene products in the immunity region and in a tail protein (Fig. 2).

The deletion in PL-1 genome removes genes corresponding to *27*, *28*, *29* and *30* of J-1. The presence of an integrase gene (*24*) strongly suggests that these phages are either temperate or derived from a temperate parent. Stetter *et al*. (84) described the isolation of

PL-1 lysogens in *L. casei* ATCC 334. A putative *attB* was found in the genome of *L. casei* BL23 and turbid plaques could be detected after infection with J-1 and PL-1. However, we were not able to isolate lysogens in this strain (data not shown). All the components of the module are currently being studied in order to propose a mechanism to regulate lysis-lysogeny in these phages.

J-1 and PL-1 are closely related to *L. casei* phage A2 among the packaging and structural genes (*1* to *17*). The predicted structural gene products could be detected in a SDS-PAGE and identified by MS analysis as components of the virion (Fig. 3 and Table 2). Similar to A2, two capsid proteins are present in the viral particle. These two gene products are the result of a translational frameshift, so gp6 and gp7 share the amino termini but gp7 is 86 amino acids longer that gp6. In both proteins the N-terminus is proteolytically processed (Fig. 4).

The virion structural gene module encoding the tail components following the canonical organization depicted by Veesler *et al*. (62) with the tail terminator, the MTP (major tail protein), the two chaperones (with the conserved translational frameshift), the tape measure protein (TMP), the baseplate hub (Dit), the gp27-like/Tal (tail associated lysozyme or tail fiber) and baseplate/tip peripheral proteins, seems to be conserved but with some interesting differences. Only two high molecular weight proteins, gp16 (75 and 69 kDa in J-1 and PL-1, respectively) and gp17 (113 kDa) appear to be part of the baseplate and host recognition apparatus. The predicted structure of gp16 is similar to Dit proteins that form the baseplate hub. The N-domain and the belt present in orf 19.1 of Spp1 (62) could be superimposed with the N-terminus of gp16. This conserved structure suggests that gp16 monomers, similarly to other phage Dit proteins, could connect to each

511 other through this region to form a circular shaped hexamer with a central wide channel
512 that allows DNA traffic during phage infection (59, 61, 62, 85, 86). Also in analogy with orf
513 19.1 of Spp1, could be expected that the C-domain protrude out of the cylinder core. The
514 C-terminus of gp16 is longer compared to other characterized Dit proteins and a variable
515 region between J-1, PL-1 and other analyzed *Lactobacillus* phages was recognized (Fig.
516 5A). The predicted structure of these variable regions (so called Dom1 of J-1 and PL-1
517 gp16) resulted similar to carbohydrate binding modules (CBMs). After modeling, the
518 structural differences between both domains became more evident, suggesting a
519 distinctive binding or affinity for sugars (Fig. 5B).
520
521 A *gfp-gp16* fusion yielded a product able to bind to *Lactobacillus casei/paracasei* cells (Fig.
522 8). Binding of this protein was inhibited in the presence of L-rhamnose. This sugar also
523 inhibited phage adsorption to purified cell walls (Fig. 9). The presence of L-rhamnose was
524 demonstrated in the cell walls of *L. casei* subsp. *casei* ATCC 27139 (87). Data shown here
525 strongly suggest that this carbohydrate is being used by the phage for host recognition and
526 gp16 is involved in this process. Adsorption inhibition assays indicate that affinity of
527 binding of J-1 gp16 to the cell walls of the 27139 strain is higher compared to PL-1 gp16
528 (Fig. 10). This is in agreement with results from whole cells phage adsorption assays (Fig.
529 7) and the observation that fluorescence intensity was consistently slightly dimmer with
530 PL-1 gp16 in the decorating assays in this strain (Fig. 8). We speculate that these
531 differences might be sufficient for a distinctive host range between both phages in a
532 suitable strain.
533
534 Based on the similarity to other phage proteins, gp17 has been annotated as the host
535 interaction protein. The N-terminus of gp17 has a similar structure to the gp27-like proteins

536 that are part of baseplates and assemble as trimers (62). An alignment of the host
537 interaction proteins of other *Lactobacillus* phages (Fig. 6A), showed a highly conserved N-
538 terminal moiety probably indicating its involvement in interaction with other phage proteins.
539 Goulet *et al*. (88) have shown that in phage Spp1, gp19.1 (Dit) and the N-terminal of gp21
540 (Tal) form a complex of one Dit hexamer with one gp21 N-terminal trimer where gp21 can
541 display a closed or an open conformation delineating a central channel to allow DNA
542 passage during infection. The conservation of Dit and N- terminal of gp27-like proteins
543 suggests that the Tal opening mechanism could be conserved in Siphophages infecting
544 Gram-positive bacteria (88). The C-terminus of gp17 (even similar between J-1 and PL-1)
545 is not conserved in the other analyzed *Lactobacillus* phages and the presence of a number
546 of collagen like repeats resembles the host interaction proteins found in *S. thermophilus*
547 phages. In these phages, it has been experimentally demonstrated that proteins carrying
548 these motifs are responsible of host recognition even probably not exclusively since
549 mutants with an expanded host range mapped also in other tail genes (64, 66). In contrast
550 with Lactococcal phages, no baseplate/tip peripheral proteins (as Receptor Binding
551 Proteins, RBPs) could be detected. The high molecular weight of gp16 and gp17 in
552 addition to the data presented here suggests the C-terminal of these proteins could be
553 playing that role in J-1 and PL-1 phages.
554 A peptidoglycan-digesting domain could not be identified as part of gp17 or located
555 somewhere else in the genome. These activities digest the peptidoglycan to form a hole
556 permitting the passage of the dsDNA at the beginning of infection. Tail associated
557 lysozymes have been recognized and characterized in Tuc2009 and TP901-1 Lactoccocal
558 phages but not in p2 (89, 90).
559
560 High resolution crystal structures of the baseplate proteins in addition to electron

micrograph reconstruction of phage particles would provide extremely useful data to decipher the particular strategies of host recognition and infection used by these phages.

Acknowledgements

This work was partially supported by ANPCyT PICT2008-0218 and UBACYT GEF 20020100200006 to MP. MED and EL are doctoral fellows of CONICET (Consejo Nacional de Investigaciones Científicas y Tecnológicas, Argentina). We thank Andrea Quiberoni (Instituto de Lactología Industrial, Argentina) for providing J-1 and PL-1 phages. We thank Pittsburgh Bacteriophage Genome Center and Madelon Portela from Institut Pasteur de Montevideo for her assistance with MALDI experiments. We specially thank Carmen Sanchez-Rivas for helpful discussions.

References

1. **Garneau JE, Moineau S.** 2011. Bacteriophages of lactic acid bacteria and their impact on milk fermentations. Microbial cell factories **10 Suppl 1:**S20.
2. **Brussow H.** 2001. Phages of dairy bacteria. Annual review of microbiology **55:**283-303.
3. **del Rio B, Binetti AG, Martin MC, Fernandez M, Magadan AH, Alvarez MA.** 2007. Multiplex PCR for the detection and identification of dairy bacteriophages in milk. Food microbiology **24:**75-81.
4. **Madera C, Monjardin C, Suarez JE.** 2004. Milk contamination and resistance to processing conditions determine the fate of *Lactococcus lactis* bacteriophages in dairies. Applied and environmental microbiology **70:**7365-7371.
5. **Suarez VB, Quiberoni A, Binetti AG, Reinheimer JA.** 2002. Thermophilic lactic acid bacteria phages isolated from Argentinian dairy industries. Journal of food protection **65:**1597-1604.
6. **Verreault D, Gendron L, Rousseau GM, Veillette M, Masse D, Lindsley WG, Moineau S, Duchaine C.** 2011. Detection of airborne lactococcal bacteriophages in cheese manufacturing plants. Applied and environmental microbiology **77:**491-497.
7. **Ebrecht AC, Guglielmotti DM, Tremmel G, Reinheimer JA, Suarez VB.** 2010. Temperate and virulent *Lactobacillus delbrueckii* bacteriophages: comparison of their thermal and chemical resistance. Food microbiology **27:**515-520.
8. **Briggiler Marco M, De Antoni GL, Reinheimer JA, Quiberoni A.** 2009. Thermal, chemical, and photocatalytic inactivation of *Lactobacillus plantarum* bacteriophages. Journal of food protection **72:**1012-1019.

9. **Suarez VB, Reinheimer JA.** 2002. Effectiveness of thermal treatments and biocides in the inactivation of Argentinian *Lactococcus lactis* phages. Journal of food protection **65:**1756-1759.

10. **Lunde M, Aastveit AH, Blatny JM, Nes IF.** 2005. Effects of diverse environmental conditions on {phi}LC3 prophage stability in *Lactococcus lactis*. Applied and environmental microbiology **71:**721-727.

11. **Madera C, Garcia P, Rodriguez A, Suarez JE, Martinez B.** 2009. Prophage induction in Lactococcus lactis by the bacteriocin Lactococcin 972. International journal of food microbiology **129:**99-102.

12. **Durmaz E, Miller MJ, Azcarate-Peril MA, Toon SP, Klaenhammer TR.** 2008. Genome sequence and characteristics of Lrm1, a prophage from industrial Lactobacillus rhamnosus strain M1. Applied and environmental microbiology **74:**4601-4609.

13. **Raya RRaH, E.M.** 2008. Isolation of phages via induction of lysogens, p. 23-32. *In* Clokie MR, Kropinski, A.M. (ed.), Bacteriohages: Methods and Protocols., vol. 1. Humana Press.

14. **Ventura M, Canchaya C, Bernini V, Altermann E, Barrangou R, McGrath S, Claesson MJ, Li Y, Leahy S, Walker CD, Zink R, Neviani E, Steele J, Broadbent J, Klaenhammer TR, Fitzgerald GF, O'Toole P W, van Sinderen D.** 2006. Comparative genomics and transcriptional analysis of prophages identified in the genomes of *Lactobacillus gasseri*, *Lactobacillus salivarius*, and *Lactobacillus casei*. Applied and environmental microbiology **72:**3130-3146.

15. **Barrangou R, Horvath P.** 2012. CRISPR: new horizons in phage resistance and strain identification. Annual review of food science and technology **3:**143-162.

16. **Sturino JM, Klaenhammer TR.** 2006. Engineered bacteriophage-defence systems in bioprocessing. Nature reviews. Microbiology **4:**395-404.

17. **Marranzino G, Villena J, Salva S, Alvarez S.** 2012. Stimulation of macrophages by immunobiotic *Lactobacillus* strains: influence beyond the intestinal tract. Microbiology and immunology **56:**771-781.

18. **Rochat T, Bermudez-Humaran L, Gratadoux JJ, Fourage C, Hoebler C, Corthier G, Langella P.** 2007. Anti-inflammatory effects of *Lactobacillus casei* BL23 producing or not a manganese-dependant catalase on DSS-induced colitis in mice. Microbial cell factories **6:**22.

19. **Greene JD, Klaenhammer TR.** 1994. Factors involved in adherence of lactobacilli to human Caco-2 cells. Applied and environmental microbiology **60:**4487-4494.

20. **Kawase M, He F, Kubota A, Harata G, Hiramatsu M.** 2010. Oral administration of lactobacilli from human intestinal tract protects mice against influenza virus infection. Letters in applied microbiology **51:**6-10.

21. **Galdeano CM, de Moreno de LeBlanc A, Vinderola G, Bonet ME, Perdigon G.** 2007. Proposed model: mechanisms of immunomodulation induced by probiotic bacteria. Clinical and vaccine immunology : CVI **14:**485-492.

22. **Villion M, Moineau S.** 2009. Bacteriophages of lactobacillus. Frontiers in bioscience : a journal and virtual library **14:**1661-1683.

23. **Garcia P, Ladero V, Suarez JE.** 2003. Analysis of the morphogenetic cluster and genome of the temperate *Lactobacillus casei* bacteriophage A2. Archives of virology **148:**1051-1070.

24. **Lo TC, Shih TC, Lin CF, Chen HW, Lin TH.** 2005. Complete genomic sequence of the temperate bacteriophage PhiAT3 isolated from *Lactobacillus casei* ATCC 393. Virology **339:**42-55.

25. **Hino Mal, N.** 1965. Lactic Acid Bacteria employed for beverage production. II. Isolation and some properties of a bacteriophage isolated during the fermentation of lactic acid beverage. Journal of Chemistry Society Japan **39:**472-476.

26. **Watanabe K, Takesue S, Jin-Nai K, Yoshikawa T.** 1970. Bacteriophage active against the lactic acid beverage-producing bacterium *Lactobacillus casei*. Applied microbiology **20:**409-415.

27. **Sechaud L, Cluzel PJ, Rousseau M, Baumgartner A, Accolas JP.** 1988. Bacteriophages of lactobacilli. Biochimie **70:**401-410.

28. **Watanabe K, Takesue S, Ishibashi K.** 1977. Reversibility of the adsorption of bacteriophage PL-1 to the cell walls isolated from *Lactobacillus casei*. The Journal of general virology **34:**189-194.

29. **Watanabe K, Ishibashi K, Nakashima Y, Sakurai T.** 1984. A phage-resistant mutant of *Lactobacillus casei* which permits phage adsorption but not genome injection. The Journal of general virology **65 (Pt 5):**981-986.

30. **Watanabe K, Hayashida M, Ishibashi K, Nakashima Y.** 1984. An N-acetylmuramidase induced by PL-1 phage infection of *Lactobacillus casei*. Journal of general microbiology **130:**275-277.

31. **Watanabe K, Kakita Y, Nakashima Y, Miake F.** 1992. Calcium requirement for protoplast transfection mediated by polyethylene glycol of *Lactobacillus casei* by PL-1 Phage DNA. Bioscience, biotechnology, and biochemistry **56:**1859-1862.

32. **Watanabe K, Kakita Y, Nakashima Y, Miake F.** 1995. Involvement of host cell energy in the transfection of *Lactobacillus casei* protoplasts with phage PL-1 DNA. Current microbiology **30:**39-43.

33. **Watanabe K, Kakita Y, Nakashima Y, Sasaki T.** 1990. Protoplast transfection of *Lactobacillus casei* by phage PL-1 DNA. Agric Biol Chem **54:**937-941.

34. **Kakita Y, Nakashima Y, Ono N, Miake F, Watanabe K.** 1996. Effects of some calcium-related agents on the protoplast transfection of *Lactobacillus casei* with phage PL-1 DNA. Current microbiology **33:**359-363.

35. **Kakita Y, Kashige N, Murata K, Kuroiwa A, Funatsu M, Watanabe K.** 1995. Inactivation of *Lactobacillus* bacteriophage PL-1 by microwave irradiation. Microbiology and immunology **39:**571-576.

36. **Kashige N, Kakita Y, Nakashima Y, Miake F, Watanabe K.** 2001. Mechanism of the photocatalytic inactivation of *Lactobacillus casei* phage PL-1 by titania thin film. Current microbiology **42:**184-189.

37. **Capra ML, Del LQA, Ackermann HW, Moineau S, Reinheimer JA.** 2006. Characterization of a new virulent phage (MLC-A) of *Lactobacillus paracasei*. Journal of dairy science **89:**2414-2423.

38. **Kakita Y, Kashige N, Miake F, Watanabe K.** 1997. Photocatalysis-dependent inactivation of Lactobacillus phage PL-1 by a ceramics preparation. Bioscience, biotechnology, and biochemistry **61:**1947-1948.

39. **Watanabe K, Takesue, S. and Ishibashi, K.** 1980. DNA of phage PL-1 active against *Lactobacillus casei* ATCC 27092. Agricultural and Biological Chemistry **44:**453-455.

40. **Khosaka T.** 1977. Physicochemical properties of a virulent *Lactobacillus* phage containing DNA with Cohesive Ends. Journal of General Virology **37:**209-214.

41. **Kashige N, Nakashima Y, Miake F, Watanabe K.** 2000. Cloning, sequence analysis, and expression of *Lactobacillus casei* phage PL-1 lysis genes. Archives of virology **145:**1521-1534.

42. **Dieterle ME, Jacobs-Sera D, Russell D, Hatfull G, Piuri M.** 2014. Complete Genome Sequences of *Lactobacillus* Phages J-1 and PL-1. Genome announcements **2**.

43. **Cresawn SG, Bogel M, Day N, Jacobs-Sera D, Hendrix RW, Hatfull GF.** 2011. Phamerator: a bioinformatic tool for comparative bacteriophage genomics. BMC bioinformatics **12:**395.

44. **Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, Sonnhammer EL, Tate J, Punta M.** 2014. Pfam: the protein families database. Nucleic acids research **42:**D222-230.

45. **Soding J, Biegert A, Lupas AN.** 2005. The HHpred interactive server for protein homology detection and structure prediction. Nucleic acids research **33:**W244-248.

46. **Webb B, Sali A.** 2014. Protein structure modeling with MODELLER. Methods in molecular biology **1137:**1-15.

47. **Palomino MM, Allievi MC, Grundling A, Sanchez-Rivas C, Ruzal SM.** 2013. Osmotic stress adaptation in *Lactobacillus casei* BL23 leads to structural changes in the cell wall polymer lipoteichoic acid. Microbiology **159:**2416-2426.

48. **Piuri M, Jacobs WR, Jr., Hatfull GF.** 2009. Fluoromycobacteriophages for rapid, specific, and sensitive antibiotic susceptibility testing of *Mycobacterium tuberculosis*. PloS one **4:**e4870.

49. **Habann M, Leiman PG, Vandersteegen K, Van den Bossche A, Lavigne R, Shneider MM, Bielmann R, Eugster MR, Loessner MJ, Klumpp J.** 2014. *Listeria* phage A511, a model for the contractile tail machineries of SPO1-related bacteriophages. Molecular microbiology **92:**84-99.

50. **Nakashima Y, Ikeda H, Kakita Y, Miake F, Watanabe K.** 1994. Restriction map of the genomic DNA of *Lactobacillus casei* bacteriophage PL-1 and nucleotide

718 sequence of its cohesive single-stranded ends. The Journal of general virology **75 (**
719 **Pt 9):**2537-2541.
720 51. **Ladero V, Garcia P, Bascaran V, Herrero M, Alvarez MA, Suarez JE.** 1998.
721 Identification of the repressor-encoding gene of the *Lactobacillus* bacteriophage A2.
722 Journal of bacteriology **180:**3474-3476.
723 52. **Garcia P, Ladero V, Alonso JC, Suarez JE.** 1999. Cooperative interaction of CI
724 protein regulates lysogeny of *Lactobacillus casei* by bacteriophage A2. Journal of
725 virology **73:**3920-3929.
726 53. **Garcia P, Rodriguez I, Suarez JE.** 2004. A -1 ribosomal frameshift in the transcript
727 that encodes the major head protein of bacteriophage A2 mediates biosynthesis of
728 a second essential component of the capsid. Journal of bacteriology **186:**1714-
729 1719.
730 54. **Moscoso M, Suarez JE.** 2000. Characterization of the DNA replication module of
731 bacteriophage A2 and use of its origin of replication as a defense against infection
732 during milk fermentation by *Lactobacillus casei*. Virology **273:**101-111.
733 55. **Garcia P, Alonso JC, Suarez JE.** 1997. Molecular analysis of the cos region of the
734 *Lactobacillus casei* bacteriophage A2. Gene product 3, gp3, specifically binds to its
735 downstream cos region. Molecular microbiology **23:**505-514.
736 56. **Benevides JM, Bondre P, Duda RL, Hendrix RW, Thomas GJ, Jr.** 2004. Domain
737 structures and roles in bacteriophage HK97 capsid assembly and maturation.
738 Biochemistry **43:**5428-5436.
739 57. **Rodriguez I, Garcia P, Suarez JE.** 2005. A second case of -1 ribosomal
740 frameshifting affecting a major virion protein of the *Lactobacillus* bacteriophage A2.
741 Journal of bacteriology **187:**8201-8204.

58. **Xu J, Hendrix RW, Duda RL.** 2004. Conserved translational frameshift in dsDNA bacteriophage tail assembly genes. Mol Cell **16:**11-21.

59. **Veesler D, Robin G, Lichiere J, Auzat I, Tavares P, Bron P, Campanacci V, Cambillau C.** 2010. Crystal structure of bacteriophage SPP1 distal tail protein (gp19.1): a baseplate hub paradigm in gram-positive infecting phages. The Journal of biological chemistry **285:**36666-36673.

60. **Bebeacua C, Bron P, Lai L, Vegge CS, Brondsted L, Spinelli S, Campanacci V, Veesler D, van Heel M, Cambillau C.** 2010. Structure and molecular assignment of lactococcal phage TP901-1 baseplate. The Journal of biological chemistry **285:**39079-39086.

61. **Veesler D, Spinelli S, Mahony J, Lichiere J, Blangy S, Bricogne G, Legrand P, Ortiz-Lombardia M, Campanacci V, van Sinderen D, Cambillau C.** 2012. Structure of the phage TP901-1 1.8 MDa baseplate suggests an alternative host adhesion mechanism. Proceedings of the National Academy of Sciences of the United States of America **109:**8954-8958.

62. **Veesler D, Cambillau C.** 2011. A common evolutionary origin for tailed-bacteriophage functional modules and bacterial machineries. Microbiology and molecular biology reviews : MMBR **75:**423-433, first page of table of contents.

63. **Shoseyov O, Shani Z, Levy I.** 2006. Carbohydrate binding modules: biochemical properties and novel applications. Microbiology and molecular biology reviews : MMBR **70:**283-295.

64. **Duplessis M, Moineau S.** 2001. Identification of a genetic determinant responsible for host specificity in *Streptococcus thermophilus* bacteriophages. Molecular microbiology **41:**325-336.

65. **Dupont K, Vogensen FK, Neve H, Bresciani J, Josephsen J.** 2004. Identification of the receptor-binding protein in 936-species lactococcal bacteriophages. Applied and environmental microbiology **70:**5818-5824.

66. **Duplessis M, Levesque CM, Moineau S.** 2006. Characterization of *Streptococcus thermophilus* host range phage mutants. Applied and environmental microbiology **72:**3036-3041.

67. **Yokokura T.** 1971. Phage receptor material in *Lactobacillus casei* cell wall. I. Effect of L-rhamnose on phage adsorption to the cell wall. Japanese journal of microbiology **15:**457-463.

68. **Yokokura T.** 1977. Phage receptor material in *Lactobacillus casei*. Journal of general microbiology **100:**139-145.

69. **Ishibashi KT, S. Watanabe, K. Oishi, K.** 1982. Use of Lectins to Characterize the Receptor Sites for Bacteriophage PL-1 of *Lactobacillus casei*. Journal of general microbiology**:**2251-2259.

70. **Shimizu-Kadota M, Sakurai T.** 1982. Prophage Curing in *Lactobacillus casei* by Isolation of a Thermoinducible Mutant. Applied and environmental microbiology **43:**1284-1287.

71. **Dupuy B, Govind R, Antunes A, Matamouros S.** 2008. *Clostridium difficile* toxin synthesis is negatively regulated by TcdC. Journal of medical microbiology **57:**685-689.

72. **Carter GP, Douce GR, Govind R, Howarth PM, Mackin KE, Spencer J, Buckley AM, Antunes A, Kotsanas D, Jenkin GA, Dupuy B, Rood JI, Lyras D.** 2011. The anti-sigma factor TcdC modulates hypervirulence in an epidemic BI/NAP1/027 clinical isolate of *Clostridium difficile*. PLoS pathogens **7:**e1002317.

73. **Matamouros S, England P, Dupuy B.** 2007. *Clostridium difficile* toxin expression is inhibited by the novel regulator TcdC. Molecular microbiology **64:**1274-1288.

74. **Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Lu F, Marchler GH, Mullokandov M, Omelchenko MV, Robertson CL, Song JS, Thanki N, Yamashita RA, Zhang D, Zhang N, Zheng C, Bryant SH.** 2011. CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic acids research **39:**D225-229.

75. **Hochschild A, Lewis M.** 2009. The bacteriophage lambda CI protein finds an asymmetric solution. Current opinion in structural biology **19:**79-86.

76. **Hall BM, Roberts SA, Heroux A, Cordes MH.** 2008. Two structures of a lambda Cro variant highlight dimer flexibility but disfavor major dimer distortions upon specific binding of cognate DNA. Journal of molecular biology **375:**802-811.

77. **Iyer LM, Koonin EV, Aravind L.** 2002. Classification and evolutionary history of the single-strand annealing proteins, RecT, Redbeta, ERF and RAD52. BMC genomics **3:**8.

78. **Arcus V.** 2002. OB-fold domains: a snapshot of the evolution of sequence, structure and function. Current opinion in structural biology **12:**794-801.

79. **Schnos M, Zahn K, Blattner FR, Inman RB.** 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.

80. **Tuohimaa A, Riipinen KA, Brandt K, Alatossava T.** 2006. The genome of the virulent phage Lc-Nu of probiotic *Lactobacillus rhamnosus*, and comparative genomics with *Lactobacillus casei* phages. Archives of virology **151:**947-965.

81. **Stephens KM, McMacken R.** 1997. Functional properties of replication fork assemblies established by the bacteriophage lambda O and P replication proteins. The Journal of biological chemistry **272:**28800-28813.

82. **Mahdi AA, Sharples GJ, Mandal TN, Lloyd RG.** 1996. Holliday junction resolvases encoded by homologous rusA genes in *Escherichia coli* K-12 and phage 82. Journal of molecular biology **257:**561-573.

83. **Lleo MM, Fontana R, Solioz M.** 1995. Identification of a gene (arpU) controlling muramidase-2 export in *Enterococcus hirae*. Journal of bacteriology **177:**5912-5917.

84. **Stetter KO.** 1977. Evidence for frequent lysogeny in lactobacilli: temperate bacteriophages within the subgenus Streptobacterium. Journal of virology **24:**685-689.

85. **Sciara G, Bebeacua C, Bron P, Tremblay D, Ortiz-Lombardia M, Lichiere J, van Heel M, Campanacci V, Moineau S, Cambillau C.** 2010. Structure of lactococcal phage p2 baseplate and its mechanism of activation. Proceedings of the National Academy of Sciences of the United States of America **107:**6852-6857.

86. **Flayhan A, Vellieux FM, Lurz R, Maury O, Contreras-Martel C, Girard E, Boulanger P, Breyton C.** 2014. Crystal structure of pb9, the distal tail protein of bacteriophage T5: a conserved structural motif among all siphophages. Journal of virology **88:**820-828.

87. **Yasuda E, Tateno H, Hirabayashi J, Iino T, Sako T.** 2011. Lectin microarray reveals binding profiles of *Lactobacillus casei* strains in a comprehensive analysis of bacterial cell wall polysaccharides. Applied and environmental microbiology **77:**4539-4546.

88. **Goulet A, Lai-Kee-Him J, Veesler D, Auzat I, Robin G, Shepherd DA, Ashcroft AE, Richard E, Lichiere J, Tavares P, Cambillau C, Bron P.** 2011. The opening of the SPP1 bacteriophage tail, a prevalent mechanism in Gram-positive-infecting siphophages. The Journal of biological chemistry **286:**25397-25405.

89. **Stockdale SR, Mahony J, Courtin P, Chapot-Chartier MP, van Pijkeren JP, Britton RA, Neve H, Heller KJ, Aideh B, Vogensen FK, van Sinderen D.** 2013. The lactococcal phages Tuc2009 and TP901-1 incorporate two alternate forms of their tail fiber into their virions for infection specialization. The Journal of biological chemistry **288:**5581-5590.

90. **Spinelli S, Veesler D, Bebeacua C, Cambillau C.** 2014. Structures and host-adhesion mechanisms of lactococcal siphophages. Frontiers in microbiology **5:**3.

91. **Krumsiek J, Arnold R, Rattei T.** 2007. Gepard: a rapid and sensitive tool for creating dotplots on genome scale. Bioinformatics **23:**1026-1028.

852 **Figure legends**

853 **Figure 1. Dot plot comparison of *Lactobacillus* phages J-1 and PL-1**. A sequence file
854 containing J-1 was compared against a file containing PL-1 using Gepard (91).
855
856 **Figure 2. A. Annotated genome maps of bacteriophages J-1 and PL-1**. The viral
857 genomes of J-1 and PL-1 are represented in four tiers with markers spaced at 1-kbp and
858 100-bp intervals. The predicted genes are shown as boxes either above or below the
859 genome, depending on whether they are rightwards or leftwards transcribed, respectively.
860 Gene numbers are shown within each box. Putative genes can be divided in the following
861 six modules: packaging (light blue), virion structure (yellow), lysis (purple), integration and
862 immunity (red) and replication (orange). The putative proteins found in the extreme right
863 region are colored in green while the ORFs lacking function are white colored. **B**. **Global**
864 **comparison of *Lactobacillus* phages**. Two pairwise nucleotide alignment of phages J-1,
865 PL-1 and related *Lactobacillus* phages (Lrm1, A2, phiAT3 and Lc-Nu) using Phamerator
866 (43). The genomes are represented by horizontal lines with putative genes shown as
867 boxes above (transcribed rightwards) or below (transcribed leftwards) each genome; the
868 number of each gene is shown within each box.
869
870 **Figure 3**. **Analysis and functional assignment of J-1 and PL-1 structural proteins**.
871 Illustrated are electron micrographs of J-1 (left) and PL-1 (right) phage particles and SDS
872 gel electrophoresis of virion proteins showing the predicted gene products. Molecular
873 mass markers (Mk) are from top to bottom, 170, 130, 100, 70, 55, 40, 35 and 25 kDa.
874
875 **Figure 4**. **Translational frameshift of capsid and chaperone mRNAs**. A translational -1

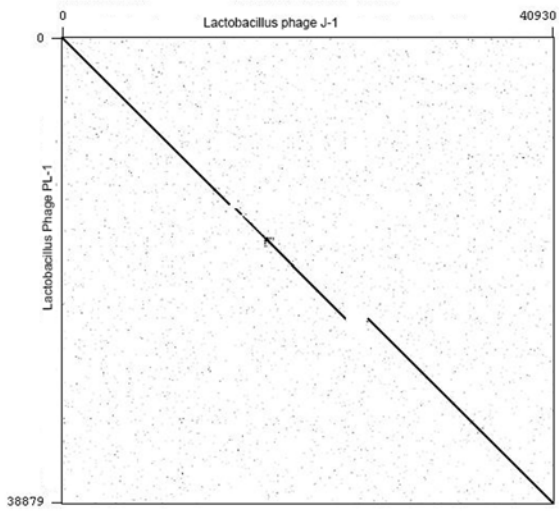
876 frameshifting near the end of transcribed genes *6* and *13* results in synthesis of two
877 different length products: gp6- gp7 (**A**) and gp13-14 (**B**), respectively. The proposed
878 slippery sequence is shaded in gray. Aminoacid sequences depicted in bold letters
879 correspond to peptides detected by MALDI-MS for the short and long forms of the protein.
880 The underlined sequence in gp6-7 (**A)** corresponds to the N-terminal peptide detected by
881 MALDI-MS and confirms the predicted proteolytical processing.
882
883 **Figure 5. Gp16 alignment and structure prediction. A.** Aminoacid sequence alignment
884 of gp16 with similar proteins found in *Lactobacillus* phages, orf 19.1 of Spp1 and orf 46 of
885 TP901-1. Upper case letters correspond to aligned Pfam domain (Sypho_tail). **B**.
886 Predicted structure of gp16 of J-1 and PL-1 based on Spp1 orf 19.1 crystal structure (PDB
887 code 2x8k_A). Dom 1 and 2 correspond to the regions that could not be modeled with this
888 PDB. Template is shown in yellow and the colors in the modeled structure correspond to
889 the domains shown in A. **C**. Predicted structure of Dom1 J-1 (orange) and Dom1 PL-1
890 (brown) of gp16 based on the CBM (carbohydrate-binding module) crystal structure of the
891 endo-β -1,4-galactanase from *Thermotoga maritima* (PDB code 2xon L) (gray).
892
893 **Figure 6. Gp17 alignment and structure prediction. A.** Amino acid sequence alignment
894 of gp17 with similar proteins found in *Lactobacillus* phages and gp44 of phage Mu. Upper
895 case letters correspond to aligned Pfam domain (Prophage_tail). **B**. Predicted structure of
896 the first 400 aminoacids of gp17 of J-1 and PL-1 (green) based on Mu gp44 crystal
897 structure (PDB code 1WRU) (yellow).
898
899 **Figure 7. Kinetics of adsorption of J-1 and PL-1.** *L. paracasei* subsp. *paracasei* ATCC
900 27092 (**A**) or *L. casei* subsp *casei* ATCC 27139 (**B**) cells were incubated with J-1 (circles)

or PL-1 (squares). At the indicated time points, PFU/ml in the supernatant were measured and percentage of adsorption was calculated.

Figure 8. Binding of gfp-gp16 to *Lactobacillus* cells. Recombinant proteins, J-1 gfp-gp16 or PL-1 gfp-gp16 were incubated with: *Lactobacillus casei paracasei* ATCC 27139 (**a-b**), *Lactobacillus acidophilus* (**c-d**) and *L. casei paracasei* ATCC 27139 in the presence of L-rhamnose (**e-f**) or glucose (**g-h**). Cells were visualized by phase contrast (left image) and fluorescence microscopy (right image). Magnification 1000X.

Figure 9. **Adsorption of J-1 and PL-1 in the presence of sugars.** J-1 (gray bars) and PL-1 (black bars) were preincubated with 0.25M of the indicated sugars and further incubated with cell walls of *L. casei* subsp *casei* ATCC 27139. PFU/ml in the supernatant were measured and percentage of adsorption compared to the control was calculated. The error bars represent the standard deviations of experiments done in triplicate.

Figure 10. Adsorption inhibition assays. Adsorption inhibition was determined when *L. casei* subsp *casei* ATCC 27139 cell walls were incubated with increasing amounts of J-1 gp16 (gray bars) or PL-1 gp16 (black bars), followed by adsorption assays using phage J-1 (**A**) or PL-1 (**B**). The error bars represent the standard deviations of experiments done in triplicate.



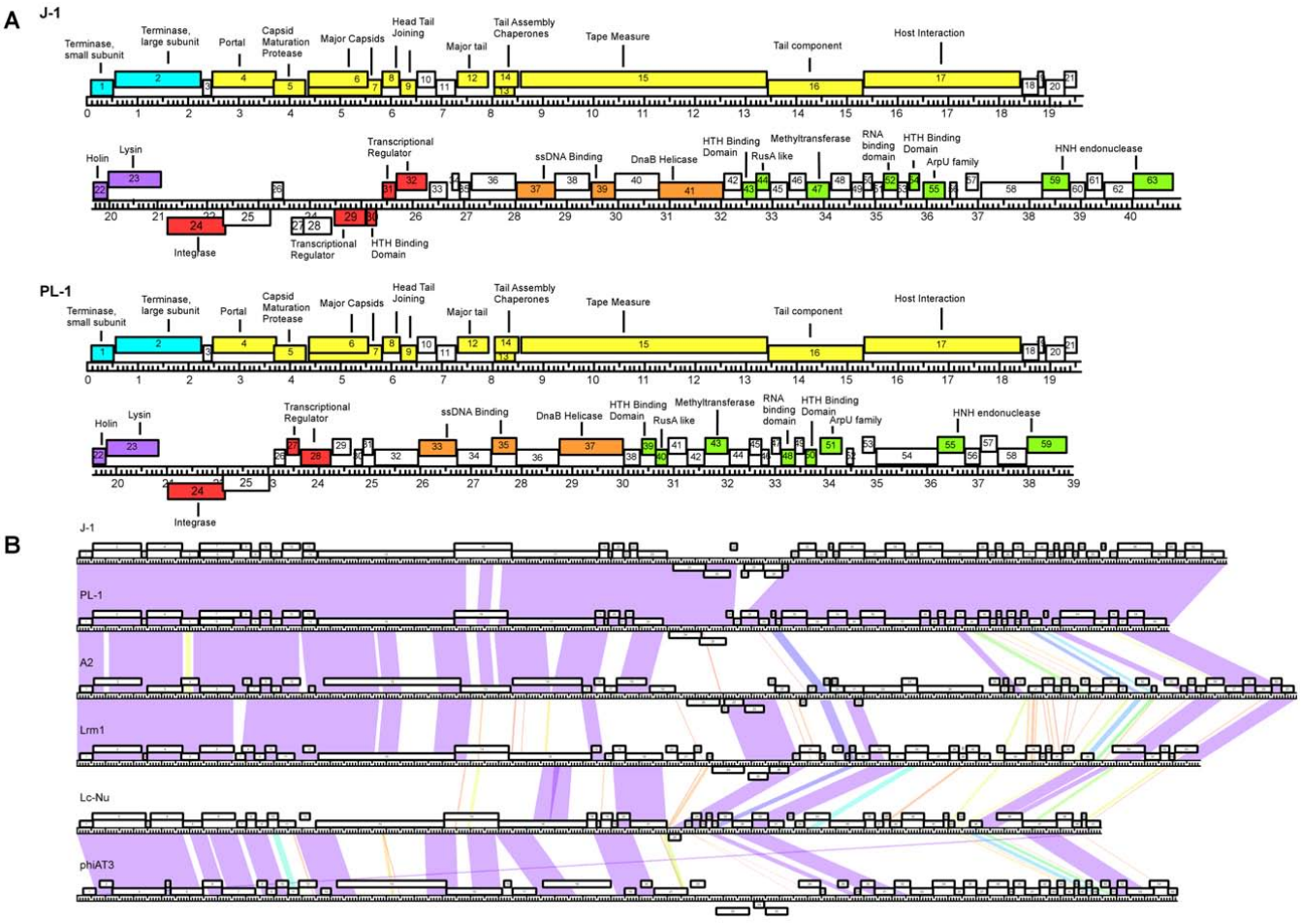
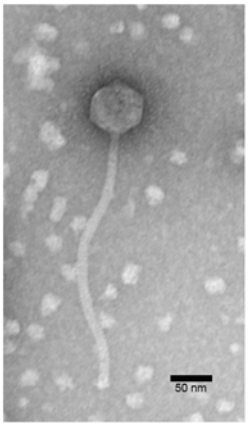
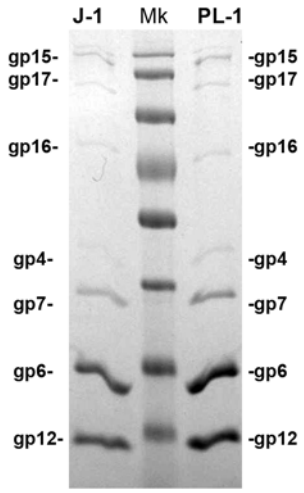


Table 1. *Lactobacillus* phages J-1 and PL-1 predicted genes and gene products.

Gene Strand	Start - Stop (Length-aa)	Gene Strand	Start – Stop (Length- aa)	Best database match (organism, *gene*)	% Ident.	Predicted Function
	J-1		PL-1			
1F	90 - 545 (152)	1F	90 - 545 (152)	*Lactobacillus phage Lrm1, 1*	97	Terminase, small subunit
2F	567 - 2279 (571)	2F	567 - 2279 (571)	*Lactobacillus phage A2, 2*	96	Terminase, large subunit
3F	2291 - 2482 (64)	3F	2291 - 2482 (64)	*Lactobacillus phage Lrm1,3*	96	
4F	2488 - 3741 (418)	4F	2488 - 3741 (418)	*Lactobacillus phage A2, 3*	96	Portal
5F	3695 - 4324 (210)	5F	3695 - 4324 (210)	*Lactobacillus phage A2, 4*	85	Capsid Maturation protease
6F	4366 - 5568(401)	6F	4366 - 5568(401)	*Lactobacillus phage A2, 5a*	99	Major Capsid
7F	4366 - 5825 (487)	7F	4366 - 5825 (487)	*Lactobacillus phage A2, 5b*	99	Major Capsid
8F	5836 - 6195 (120)	8F	5836 - 6195 (120)	*Lactobacillus phage A2, 6*	75	head-tail joining
9F	6185 - 6514 (110)	9F	6185 - 6514 (110)	*Lactobacillus phage A2, 7*	92	head-tail joining
10F	6514 - 6900 (129)	10F	6514 - 6900 (129)	*Lactobacillus phage Lrm1, 9*	98	
11F	6900 - 7286 (129)	11F	6900 - 7286 (129)	*Lactobacillus phage Lrm1,10*	98	
12F	7320 - 7937 (206)	12F	7320 - 7937 (206)	*Lactobacillus phage A2, 10*	98	Major Tail
13F	8036 - 8449 (138)	13F	8036 - 8449 (138)	*Lactobacillus phage Lrm1, 12*	99	Tail assembly Chaperones
14F	8036 - 8550 (172)	14F	8036 - 8550 (172)	*Lactobacillus phage Lrm1, 12*	90	Tail Assembly Chaperones
15F	8572 - 13434 (1621)	15F	8572 - 13434 (1621)	*Lactobacillus phage Lrm1, 13*	97	Tape Measure
16F	13435 - 15474 (680)	16F	13435 - 15327 (631)	*Lactobacillus phage A2, 13*	45	Tail Component
17F	15471 - 18590 (1040)	17F	15324 - 18443 (1040)	*Lactobacillus phage Lrm1, 15*	85	Host Interaction
18F	18600 - 18923 (108)	18F	18453 - 18782 (110)	*Lactobacillus phage A2, 15*	100	
19F	18916 - 19047 (44)	19F	18769 - 18900 (44)	*Lactobacillus rhamnosus Lc-Nu, 17*	84	
20F	19072 - 19464 (131)	20F	18925 - 19317 (131)	*Lactobacillus phage Lc-Nu, 18*	97	
21F	19445 - 19687 (81)	21F	19298 - 19540 (81)	*Lactobacillus phage Lc-Nu, 19*	79	
22F	19677 - 19949 (91)	22F	19530 - 19802 (91)	*Lactobacillus phage PL1*	99	Holin
23F	19951 - 21003 (351)	23F	19804 - 20856 (351)	*Lactobacillus phage PL1*	100	Lysin
24R	22375 - 21224 (384)	24R	22228 - 21077 (384)	*Lactobacillus casei BL23*	100	Integrase
25R	23246 - 22311 (312)	25R	23099 - 22164 (312)	*Lactobacillus casei BL23*	96	
26F	23265 - 23495 (77)	26F	23118 - 23348 (77)	*Lactobacillus casei BL23*	100	
27R	23632 - 23564 (24)			*Lactobacillus casei ATCC334*	95	
28R	23895 - 23656 (74)			*Lactobacillus Phage A2, 21*	66	
29R	25117 - 24488 (210)			*Latobacillus phage phiAT3, 25*	88	Repressor
30R	25311 - 25105 (69)			*Lactobacillus rhamnosus LMS2-1*	100	HTH binding domain
31F	25430 - 25684 (85)	27F	23379 – 23633 (85)	*Latobacillus rhamnosus LMS2-1*	100	Repressor
32F	25687 - 26298 (204)	28F	23636 – 24247 (204)	*L.paracasei ATCC 25302*	89	Anti-Repressor
33F	26326 - 26685 (120)	29F	24275 – 24634 (120)	*Lactobacillus phage phiAT3, 25*	97	
34F	26767 - 26919 (51)	30F	24716 – 24868 (51)	*Latobacillus phage Lrm1, 31*	98	
35F	26924 - 27127 (68)	31F	24873 – 25076 (68)	*Lactobacillus phage phiAT3, 26*	91	
36F	27145 - 28032 (296)	32F	25094 – 25981 (296)	*Enterococcus phage phief11*	29	
37F	28032 - 28787 (252)	33F	25981 – 26736 (252)	*Lactobacillus phage LBR48*	59	ssDNA binding protein
38F	28784 - 29458 (225)	34F	25094 – 25981 (225)	*Lactobacillus rhamnosus HN001*	94	
39F	29473 - 29958 (162)	35F	27422 – 27907 (162)	*L.paracasei subsp. paracasei 8700:2*	88	ssDNA binding protein
40F	29936 - 30808 (290)	36F	27888 – 28757 (290)	*Lactobacillus phage Lc-Nu, 34*	78	HTH binding domain
41F	30805 - 32067 (421)	37F	28754 – 30016 (421)	*Lactobacillus phage Lc-Nu, 35*	98	DnaB Helicase
42F	32069 - 32413 (115)	38F	30018 – 30362 (115)	*Lactobacillus phage Lc-Nu, 36*	90	
43F	32426 - 32713 (96)	39F	30375 – 30662 (96)	*Lactobacillus phage Lc-Nu, 37*	91	HTH binding domain
44F	32700 - 32954 (85)	40F	30649 – 30903 (85)	*Lactobacillus phage Lc-Nu, 38*	93	RusA like
45F	32951 - 33316 (122)	41F	30900 – 31265 (122)	*Lactobacillus phage Lc-Nu, 39*	99	
46F	33328 - 33666 (113)	42F	31277 – 31615 (113)	*Lactobacillus phage A2, 43*	96	
47F	33678 - 34133 (152)	43F	31627 – 32082 (152)	*Lactobacillus phage Lc-Nu, 40*	96	Methyltranferase
48F	34144 - 34551 (136)	44F	32093 – 32500 (136)	*Lactobacillus phage Lrm1, 48*	57	
49F	34544 - 34789 (82)	45F	32493 – 32738 (82)	*L. paracasei ATCC 25302*	94	
50F	34782 - 34952 (57)	46F	32731 – 32901 (57)	*L. paracasei ATCC 25302*	56	
51F	34977 - 35174 (66)	47F	32926 – 33123 (66)	*L.paracasei ATCC 25302*	88	
52F	35164 - 35457 (98)	48F	33113 – 33406 (98)	*Lactobacillus phage Lb338-1*	81	RNA binding domain
53F	35450 - 35641 (64)	49F	33399 – 33590 (64)	*Haemophihlus parasuis SH0165*	38	
54F	35662 - 35880 (73)	50F	33611 – 33829 (73)	*Lactobacillus phage Lc-Nu,44*	90	HTH binding domain
55F	35945 - 36382 (146)	51F	33894 – 34331 (146)	*Lactobacillus phage Lrm1, 49*	96	ArpU family
56F	36471 - 36617 (50)	52F	34417 – 34566 (50)	*Lactobacillus phage A2, 54*	98	
57F	36779 - 37033 (85)	53F	34728 – 34982 (85)	*Lactobacillus phage A2, 57*	89	
58F	37058 - 38275 (406)	54F	35007 – 36224 (406)	*Lactobacillus phage Lc-Nu, 47*	96	
59F	38262 - 38792 (177)	55F	36211 – 36741 (177)	*Lactobacillus phage A2,60*	97	HNH endonuclease
60F	38796 - 39116 (107)	56F	36745 – 37065 (107)	*Lactobacillus phage phi AT3, 52*	78	
61F	39119 - 39442 (108)	57F	37068 – 37391 (108)	*Lactobacillus phage A2, 61*	89	
62F	39455 - 40036 (194)	58F	37404 – 37985 (194)	*Lactobacillus phage A2, 62*	95	
63F	40026 - 40820 (265)	59F	37975 – 38769 (265)	*Lactobacillus phage Lrm1, 54*	98	HNH endonuclease / P-loop NTPase domain





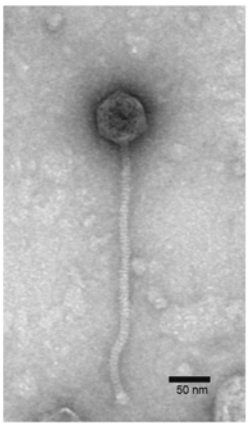


Table 2. Identification of virion-associated proteins

gp	MW [kDa]	Coverage[a]	# PSMs[b]	MW [kDa]	Coverage[a]	# PSMs[b]
		J-1			PL-1	
gp4 (portal)	46,3	51,91	45	46,3	52,39	51
gp6 (major capsid)	29,90	55,44	837	29,9	82,54	561
gp7 (major capsid)	38,3	69,61	644	38,3	74,33	548
gp12 (major tail)	22,1	74,27	220	22,1	80,58	337
gp15 (tmp)	173,2	60,27	882	173,2	59,04	476
gp16 (tail component)	75,1	51,91	91	69,2	25,99	29
gp17 (host interaction)	112,7	43,17	45	113,0	39,90	88

[a] Percentage of predicted protein sequence identified in peptides.
[b] PSMs, peptide spectrum matches.

A

```
1    AUGACUUUAGAUGAAAAAUUAGCUGCUGUUAAAAAGCAACUUGAU
1     M  T  L  D  E  K  L  A  A  V  K  K  Q  L  D

46   GAAAAGCGUUCAGCGUUACCAGCUAUGAAGACAGAACUUCGUUCU
16    E  K  R  S  A  L  P  A  M  K  T  E  L  R  S

91   UUACUUGAAGGUGAAGAUUCCGAGGAAAACCUGAAGAAGGCAGAA
31    L  L  E  G  E  D  S  E  E  N  L  K  K  A  E

136  GGCGUUCGUGCCAAGUAUGAUAAAGCUGGCAAAGAGAUCAAAGAU
46    G  V  R  A  K  Y  D  K  A  G  K  E  I  K  D

181  CUUGAAGAAAAACGUGACUUAUACGAGGCUGCGUUGAAAGGCAAU
61    L  E  E  K  R  D  L  Y  E  A  A  L  K  G  N

226  GAACAGCCGAGUGGGAAGAAGCCCGAUCAUCCGGAAGAGCAUAGC
76    E  Q  P  S  G  K  K  P  D  H  P  E  E  H  S

271  UAUCGCGAUGCACUGAAUGCUUAUUUGCAUACUCGUGGCCGUGAU
91    Y  R  D  A  L  N  A  Y  L  H  T  R  G  R  D

316  ACAGAAGGCGTCAAUUUUGAAAAGACUGAUGUUGGCACAUUUGCA
106   T  E  G  V  N  F  E  K  T  D  V  G  T  F  A

361  GUUUUACGAGCUGUUCCUACUGAUGCCAGUGAUGCGGUAAAUGCC
121   V  L  R  A  V  P  T  D  A  S  D  A  V  N  A

406  GGUGUCAAGGCUGCAGACGCGGCCUCUACCAUUCCAGAAACUAUU
136   G  V  K  A  A  D  A  A  S  T  I  P  E  T  I

451  AGCAAUACACCACAGCGUGAAUUGCAGACUGUUGUUGAUCUGAAA
151   S  N  T  P  Q  R  E  L  Q  T  V  V  D  L  K
-------------------------------------------------
5536 GGGUACUUCCUCACGUAUACCCCAAAAGCGUAA
391   G  Y  F  L  T  Y  T  P  K  A  *

5562                              AGCGUAACGCCUGACGGA
400                                S  V  T  P  D  G

5580 GUGACUUUGAGCCAGAAAACGUUCACGGGUGGUGUCGGUGCCACA
406   V  T  L  S  Q  K  T  F  T  G  G  V  G  A  T

5625 AAAGAUAUCACGGUGACAGUCACUCCUGAUGGCGCUCCUCAAGCA
421   K  D  I  T  V  T  V  T  P  D  G  A  P  Q  A

5670 GUCGAAGCUGUGUCGAGCAAUGAAAGCGUCGCUACGGUUGUUAAG
436   V  E  A  V  S  S  N  E  S  V  A  T  V  V  K

5715 AAGUCCGAUGGUGUUUACACCAUUACCAAUCUGGCAGCGGGUGCA
451   K  S  D  G  V  Y  T  I  T  N  L  A  A  G  A

5760 GCGACAAUCACAUUUAGCACUAAUGGCAUCAGCUCAACGCUUGCC
466   A  T  I  T  F  S  T  N  G  I  S  S  T  L  A

5805 GUUACUGUUAACGCUGGGUAG
481   V  T  V  N  A  G  *
```

B

```
8351 AAGGACACAGCAAAAAAAAUCACCGAAGCGGACGUCAAAGAAGCCA
106   K  D  T  A  K  K  S  P  K  R  T  S  K  K  P
111                K  I  T  E  A  D  V  K  E  A

8396 UUAGCAACCUUGACGACUUCUACAAAGCAAGGCUCUCUGAAGGCU
121   L  A  T  L  T  T  S  T  K  Q  G  S  L  K  A
121   I  S  N  L  D  D  F  Y  K  A  R  L  S  E  G

8441 ACCGAUUAGCUGACGUUGAUGCUAUGACGCUCCGCGAUAUUGAAA
136   T  D  *
136   Y  R  L  A  D  V  D  A  M  T  L  R  D  I  E

8486 AGCUUAACCAGAUUUACGAGGAACGGGAGACCACGAUCGACAAGG

151  K  L  N  Q  I  Y  E  E  R  E  T  T  I  D  K

8531 CCUUUCCGUUCCUUUUCUAG

166  A  F  P  F  L  F  *
```