AEM
Journals.ASM.org

Investigation of the Relationship between Lactococcal Host Cell Wall Polysaccharide Genotype and 936 Phage Receptor Binding Protein Phylogeny

Jennifer Mahony,[a] Witold Kot,[b] James Murphy,[a] Stuart Ainsworth,[a] Horst Neve,[c] Lars H. Hansen,[d] Knut J. Heller,[c] Søren J. Sørensen,[d] Karin Hammer,[e] Christian Cambillau,[f] Finn K. Vogensen,[b] Douwe van Sinderen[a,g]

Department of Microbiology, University College Cork, Cork, Ireland[a]; Department of Food Science, Faculty of Science, University of Copenhagen, Copenhagen, Denmark[b]; Max Rubner Institut, Kiel, Germany[c]; Department of Biology, Faculty of Science, University of Copenhagen (KU), Copenhagen, Denmark[d]; Department of Systems Biology, Technical University of Denmark, Lyngby, Denmark[e]; Architecture et Fonction des Macromolécules Biologiques, UMR 7257 CNRS and Aix-Marseille University, Marseille, France[f]; Alimentary Pharmabiotic Centre, Biosciences Institute, University College Cork, Cork, Ireland[g]

Comparative genomics of 11 lactococcal 936-type phages combined with host range analysis allowed subgrouping of these phage genomes, particularly with respect to their encoded receptor binding proteins. The so-called pellicle or cell wall polysaccharide of *Lactococcus lactis*, which has been implicated as a host receptor of (certain) 936-type phages, is specified by a large gene cluster, which, among different lactococcal strains, contains highly conserved regions as well as regions of diversity. The regions of diversity within this cluster on the genomes of lactococcal strains MG1363, SK11, IL1403, KF147, CV56, and UC509.9 were used for the development of a multiplex PCR system to identify the pellicle genotype of lactococcal strains used in this study. The resulting comparative analysis revealed an apparent correlation between the pellicle genotype of a given host strain and the host range of tested 936-type phages. Such a correlation would allow prediction of the intrinsic 936-type phage sensitivity of a particular lactococcal strain and substantiates the notion that the lactococcal pellicle polysaccharide represents the receptor for (certain) 936-type phages while also partially explaining the molecular reasons behind the observed narrow host range of such phages.

Lactococcal phages are classified into 10 groups based on morphology and DNA hybridization studies (1). Of these, three species are most frequently isolated from dairy environments, namely, the 936, c2, and P335 species (1). The 936 phages are strictly lytic and frequently cause problems for the dairy fermentation industry (2). For this reason, members of this phage species have attracted significant attention in recent years, and the genomes of several phages of this species are now available (3–8). The genome organization of these phages is well conserved and consists of three clusters: the early-, middle-, and late-expressed regions. Comparative genomic analysis of these phages has revealed that while most of the structural genes are highly conserved, there are particular regions of diversity within other regions of their genomes. Among these are the early-expressed genes, which are assumed to encode the replication functions of the phage, as well as the late-expressed region, particularly within the genes encoding the receptor binding protein (RBP) and the tail tape measure protein (TMP) (3, 4, 8). It has been suggested that the genomes of the 936-type phages can be subgrouped based on these variable regions, which appear to correspond to the subspecies grouping (i.e., *Lactococcus lactis* subsp. *lactis* or *L. lactis* subsp. *cremoris*) of the host(s) which they infect (3, 9). However, there are a number of exceptions to this in terms of phages that are capable of infecting both subspecies of *L. lactis*. Phages 645 and P475, for example, are capable of infecting certain members of both subspecies and possess a RBP which is different from those of the two major subgroups of the 936 phages (9). Recently, the RBP structures of phages p2 and bIL170 (or domains thereof) have been solved, which revealed that these RBPs are comprised of three regions, the N-terminal shoulder domain, the central neck domain, and the C-terminal head domain, the latter of which is believed to be required for receptor recognition (10–13). The specificity of the RBP is assumed to be associated with amino acids forming a groove between two monomers of the trimeric RBP head (10, 13).

Host genes that are involved in facilitating adsorption of 936-type phages have been identified for phages bIL170 and 645 (which infect *L. lactis* subsp. *lactis* IL1403 and *L. lactis* subsp. *cremoris* Wg2, respectively) and represent genetic elements within a gene cluster that is believed to be responsible for the biosynthesis of a cell wall-associated polysaccharide (CWPS) (14). More recently, so-called long-chain mutants of the prototype *L. lactis* subsp. *cremoris* strain MG1363 were isolated, which carried a mutation in a gene within the presumed CWPS or so-called pellicle biosynthesis cluster and which were also shown to be insensitive to phage sk1 (a 936-type phage) infection (15). The evidence for an interaction between lactococcal 936-type phages and the CWPS on the host cell envelope is compelling, although conclusive proof for such direct phage-host binding has not yet been obtained.

A study relating to the lactococcal host receptor material has included phage inactivation assays with the cell wall, cell mem-

Received 1 March 2013 **Accepted** 3 May 2013
Published ahead of print 10 May 2013
Address correspondence to Douwe van Sinderen, d.vansinderen@ucc.ie.
Jennifer Mahony and Witold Kot contributed equally to this work.
Supplemental material for this article may be found at http://dx.doi.org/10.1128/AEM.00653-13.

doi:10.1128/AEM.00653-13

branes, and polysaccharides (16). This study has implicated a saccharide component in adsorption, although very little is known about the requirement by these phages for specific polysaccharides or if it is a general receptor that permits initial and reversible binding prior to irreversible binding to a secondary receptor. Interestingly, the RBP head domain of phages bIL170 and p2 exhibits a high avidity for saccharide components, including glycerol and phosphoglycerol (10, 13), and for this reason, the current dogma would suggest a singular role for cell envelope-associated phosphorylated polysaccharide or lipoteichoic acid components in adsorption and infection by this group of phages. It should be noted, however, that the notion of lipoteichoic acid as the lactococcal phage receptor would be at odds with the observed specificity of 936-type phages of being able to infect just one or a very limited number of *L. lactis* strains. Therefore, the discovery of more complex saccharides as putative lactococcal phage receptors opened a new field of promising investigations.

In this study, we further expand the bank of publicly available genome sequences of 936-type phages and exploit this information to reveal a correlation between the receptor binding head domain of these phages and the observed diversity among CWPS-encoding gene clusters of their respective lactococcal host.

MATERIALS AND METHODS

Lactococcal strains and bacteriophages. Lactococcal strains (Table 1) were grown in M17 broth supplemented with 0.5% glucose at 30°C without agitation. Phages were propagated on the relevant strains at 30°C in M17 broth (Oxoid, Hampshire, United Kingdom) supplemented with 0.5% glucose and 5 mM $CaCl_2$ without agitation, as previously described (17). The phages used in this study and relevant details are listed in Table 2. Plaque assays were performed by using the double-agar method, as previously described (18). This method was also used to determine the host range of phages against a bank of lactococcal strains (Tables 1 and 2).

Phage purification and DNA preparation. Phage purification by cesium chloride gradient centrifugation was performed as previously described (19). The generated purified phage suspension (1 ml) was precipitated with 10% polyethylene glycol 8000 (PEG 8000) (Sigma-Aldrich, MO, USA) and 0.5 M sodium chloride at 4°C overnight. Subsequently, the suspension was centrifuged at 4°C at 17,700 × *g* for 15 min, and the supernatant was removed (phages 340, 645, JM2, JM3, and P113g). Alternatively, the phage suspension was dialyzed as described previously by Sambrook and Russell for phage λ (phages 936 [named 936^P throughout the text to identify this as the prototype phage of the species], fd13, P272, P680, P475, and 7) (19). The PEG-salt-induced precipitate was resuspended in 0.5 ml of Tris-EDTA (TE) buffer (pH 9.0) and treated with 20 μl of 20 mg ml^{-1} proteinase K (Sigma-Aldrich, MO, USA) for 20 min at 56°C, followed by treatment with SDS at a final concentration of 2% at 65°C for 20 min. This mixture was then phenol-chloroform (25:24:1 phenol-chloroform-isoamyl alcohol; Sigma-Aldrich, MO, USA) treated at least twice, and the aqueous phase was precipitated with 2.5 volumes of ice-cold 96% ethanol and 0.1 volume of sodium acetate (pH 4.8). After centrifugation at 20,000 × *g* at 4°C for 15 min, the pellet was washed in 70% ethanol and resuspended in 100 μl of TE buffer (pH 8.0).

Genome sequencing, assembly, and annotation. For DNA sequencing, 5 μg of DNA of phages 645, 340, *Viridus*JM2 (JM2), *Pastus*JM3 (JM3), and P113g was extracted and verified by nanodrop quantification, and confirmatory molecular identification tests were conducted on the DNA extract prior to shipment to the contract sequencing facility (Macrogen Inc., Seoul, South Korea). A 40- to 65-fold sequencing coverage was obtained by using pyrosequencing technology on a 454 FLX instrument. The files generated by the 454 FLX instrument were assembled with GSassembler (454, Branford, CT, USA) to generate a consensus sequence. Phages P475, fd13, 7, 936^P, and P272 were sequenced by using the 454 Roche Titanium platform available at the Department of Systems Biology, Technical University of Denmark, Lyngby, Denmark. These phages were sequenced as part of tagged pools of unrelated phages, built as MID-tagged Rapid libraries, and sequenced in one region (half a picotiter plate) using the GS FLX Titanium XLR70 sequencing kit. One phage, P680, was sequenced as 96-base reads by using the Illumina HighSeq2000 platform (Department of Systems Biology, Technical University of Denmark), again as part of a pool of unrelated phages, tagged with an index as part of one lane of the flow cell. Custom indexing primers were used to build libraries as described previously (20). Reads were assembled into contigs by using CLC Genomics Workbench 5.0.1 (CLC Bio, Aarhus, Denmark). Quality improvement of the genome sequence involved sequencing of 15 to 25 PCR products across the entire genomes to ensure correct assembly, double stranding, and the resolution of any remaining base conflicts occurring within homopolynucleotide tracts. Protein-encoding open reading frames (ORFs) were predicted by using Zcurve_V and Genmark.hmm (21, 22), followed by manual assessment and, where necessary, correction (18). Preliminary identification and functional annotation of ORFs were performed on the basis of BLASTP (23) analysis against the nonredundant protein database (nr) provided by the National Center for Biotechnology Information (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

CWPS typing of lactococcal strains by multiplex PCR. The relevant DNA regions encompassing the cell wall polysaccharide biosynthesis operon in the sequenced lactococcal strains IL1403 (GenBank accession no. AE005176.1), KF147 (accession no. NC_013656.1), MG1363 (acces-

TABLE 1 Features of the lactococcal strains used in this study

Strain	Reference or source[a]	Organism	CWPS type	Infected by phage(s)
IL1403	35	*L. lactis* subsp. *lactis*	IL/KF (B)	+
F7/2	36	*L. lactis* subsp. *lactis* biovar *diacetylactis*	IL/KF (B)	+
Bu2-60	37	*L. lactis* subsp. *lactis*	IL/KF (B)	+
455	Dupont-Danisco	*L. lactis* subsp. *cremoris*	IL/KF (B)	+
UC77	UCC	*L. lactis* subsp. *cremoris*	IL/KF (B)	+
229	UCC	*L. lactis* subsp. *lactis*	IL/KF (B)	−
WM1	UCC	*L. lactis* subsp. *lactis*	UC/CV (A)	−
ML8	1	*L. lactis* subsp. *cremoris*	UC/CV (A)	−
C10	UCC	*L. lactis* subsp. *cremoris*	UC/CV (A)	−
IE-16	1	*L. lactis* subsp. *cremoris*	UC/CV (A)	−
SMQ-450	1	*L. lactis* subsp. *cremoris*	UC/CV (A)	−
SMQ-562	1	*L. lactis* subsp. *cremoris*	UC/CV (A)	−
111	1	*L. lactis* subsp. *cremoris*	UC/CV (A)	−
UC063	UCC	*L. lactis* subsp. *cremoris*	UC/CV (A)	−
UC509.9	UCC	*L. lactis* subsp. *cremoris*	UC/CV (A)	−
UL8	UCC	*L. lactis* subsp. *lactis*	UC/CV (A)	−
275	UCC	*L. lactis* subsp. *lactis*	UC/CV (A)	−
KH	28	*L. lactis* subsp. *cremoris*	UC/CV (A)	−
C3	38	*L. lactis* subsp. *cremoris*	UC/CV (A)	−
W22	39	*L. lactis* subsp. *cremoris*	MG/SK (C)	+
W34	7	*L. lactis* subsp. *cremoris*	MG/SK (C)	+
3107	40	*L. lactis* subsp. *cremoris*	MG/SK (C)	+
WG2	36	*L. lactis* subsp. *cremoris*	MG/SK (C)	+
FD13	9	*L. lactis* subsp. *cremoris*	MG/SK (C)	+
H2	41	*L. lactis* subsp. *cremoris*	MG/SK (C)	+
158	34	*L. lactis* subsp. *cremoris*	MG/SK (C)	+
V32.2	42	*L. lactis* subsp. *cremoris*	MG/SK (C)	+
JM3	UCC	*L. lactis* subsp. *cremoris*	MG/SK (C)	+
JM2	UCC	*L. lactis* subsp. *cremoris*	MG/SK (C)	+
1196	UCC	*L. lactis* subsp. *cremoris*	MG/SK (C)	+
SMQ86	1	*L. lactis* subsp. *cremoris*	MG/SK (C)	−
US3	43	*L. lactis* subsp. *cremoris*	MG/SK (C)	−
E8	KU	*L. lactis* subsp. *cremoris*	MG/SK (C)	+
MG1614	44	*L. lactis* subsp. *cremoris*	MG/SK (C)	−
901-1	40	*L. lactis* subsp. *cremoris*	MG/SK (C)	−
184	UCC	*L. lactis* subsp. *lactis*	Unknown	−

[a] UCC, University College Cork.

TABLE 2 Features of the phages used in this study

Phage	Source or reference	Source location	Yr of first report/isolation	Propagating host	Genome length (kb)	G+C content (%)	No. of ORFs	RBP group
fd13	KU	Denmark	2004	FD13	30.674	34.7	53	I
P113g	9	Germany	1986	IL1403	30.796	34.1	58	II
P272	9	Germany	1986	IL1403	30.778	34.1	61	II
936	34	New Zealand	1984	158	27.302	34.5	49	II
P475	Danisco Culto[a]	Europe	Unknown	455	30.961	34.3	57	III
P680	Z. Atamer and J. Hinrichs[b]	Germany	2009	IL1403	29.631	35.1	49	II
7	KU	Denmark	2004	V32.2	32.382	34.2	57	I
645	This study	Denmark	2004	IL1403	29.247	35.0	51	III
340	This study	Denmark	2010	IL1403	32.337	34.5	58	III
*Viridus*JM2	This study	Ireland	2010	JM2	31.090	34.3	59	I
*Pastus*JM3	This study	Ireland	2010	JM3	28.674	34.4	52	I

[a] Danisco Cultor Niebüll GmbH, Niebüll, Germany.
[b] University of Hohenheim, Germany.

sion no. NC_009004.1), SK11 (accession no. NC_008527.1), UC509.9 (accession no. CP003157.1), and CV56 (accession no. CP002365.1) were analyzed and compared by using BLASTP (23) analysis as described above. By using these data, conserved and unique regions within the operons of these strains were identified (Fig. 1). Primers were designed based on LLKF_205 of IL1403/KF147 (CWPS type B) (product size of 183 bp), llmg_0226 of MG1363/SK11 (CWPS type C) (product size of 686 bp), and UC509_0206 of UC509.9/CV56 (CWPS type A) (product size of 442 bp) CWPS types, as indicated in Fig. 1 (Table 3). A control was also included, in which primers based on the conserved *rmlB* gene were used to generate a product of 891 bp to verify that the PCR was working in all samples. The multiplex PCR included these four sets of primers and was applied to the strains assessed in the host range analysis (Table 1) under the following conditions: 95°C for 6 min followed by 31 cycles of 95°C for 15 s, 57°C for 30 s, and 72°C for 1 min, followed by a final extension step at 72°C for 7 min.

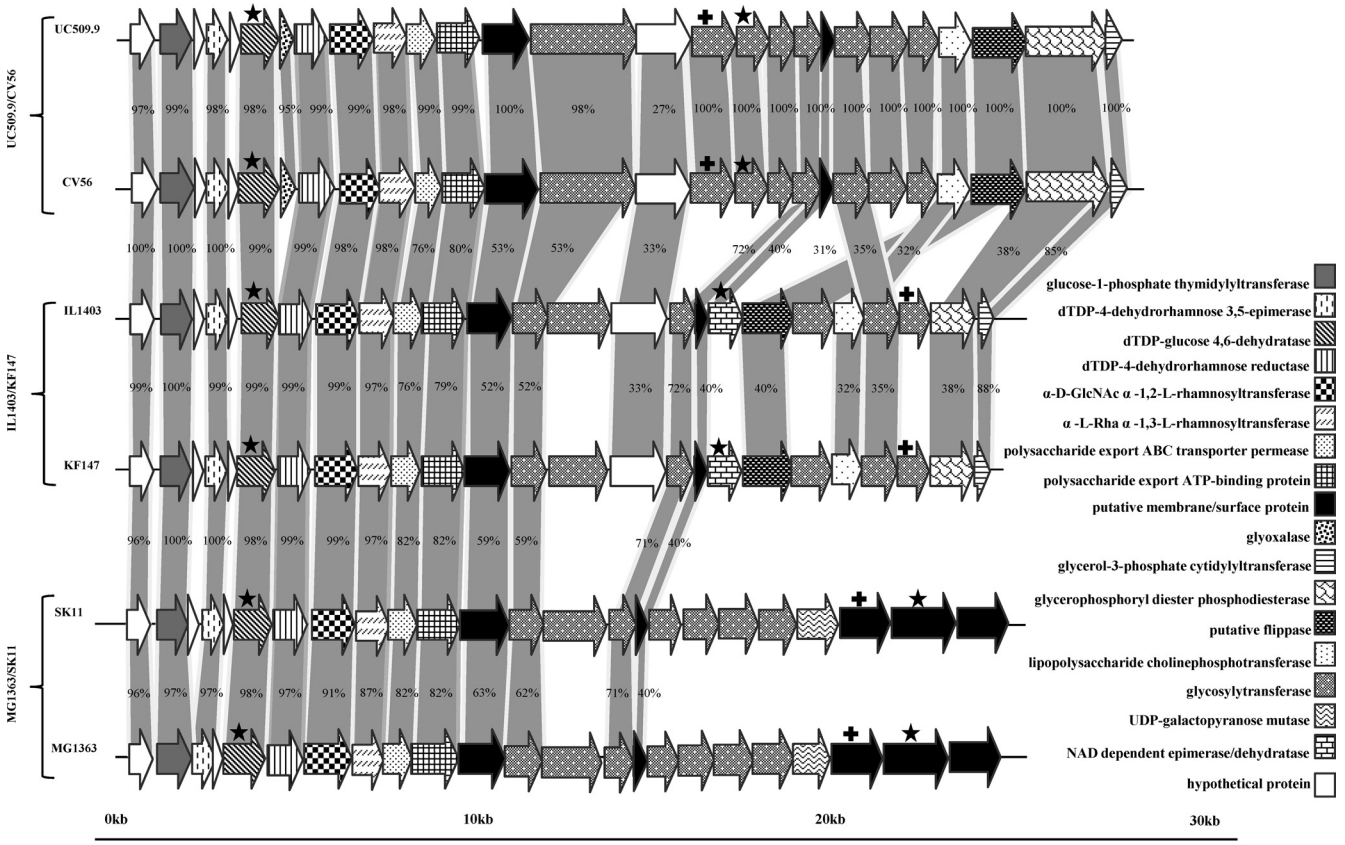

FIG 1 Comparison of the genomic regions encoding the CWPS biosynthesis clusters of six lactococcal strains (UC509.9, CV56, IL1403, KF147, SK11, and MG1363). Homologous protein-encoding genes are joined by gray blocks including the level of identity (percent amino acids) relative to UC509.9. The three subgroups of CWPS are the UC509.9/CV56 (UC/CV) (CWPS A), the IL1403/KF147 (IL/KF) (CWPS B), and the MG1363/SK11 (MG/SK) (CWPS C) subgroups, with unique regions in the operons of UC509.9, IL1403, and MG1363 in the schematic highlighted by stars to indicate the genes upon which the multiplex PCRs are based. The control for the multiplex PCR was based on the *rmlB* gene, which is conserved in each of the strains.

TABLE 3 Primers for CWPS multiplex PCR

Primer	Sequence (5′–3′)	Product size (bp)
IL-KFfw	GATTCAGTTGCACGGCCG	
IL-KFrv	AGTAAGGGGGCGGATTGTG	183
MG-SKfw	AAAGCTCATCTTTCCCCTGTTGT	
MG-SKrv	GCACCATAGTCTGGAATAAGACC	686
UC-CVfw	GTGCCTATGCTCCGTTAGTC	
UC-CVrv	CGAGGGCCAATCTCTTTACC	442
CONfw (control)	GTACACTATGTTTATAACAATCATCCAG	
CONrv	GCAAACCAGATTCAAAGTCAGTATG	891

Electron microscopy (EM). Adsorption of CsCl-purified phages to freshly prepared carbon film floated from a freshly coated mica sheet and negative staining with 2% (wt/vol) uranyl acetate were done as described previously (24). The film was picked up with a 400-mesh copper grid (Agar Scientific, Essex, United Kingdom), and specimens were examined with a Tecnai 10 transmission electron microscope (FEI, Eindhoven, The Netherlands) operated at an acceleration voltage of 80 kV.

Nucleotide sequence accession numbers. The GenBank accession numbers for the phages sequenced in this study are as follows: KC182542 (phage 340), KC182543 (phage 645), KC182544 (phage 936), KC182545 (phage fd13), KC182546 (phage JM2), KC182547 (phage JM3), KC182548 (phage P113G), KC182549 (phage P272), KC182550 (phage P475), KC182551 (phage P680), and KC182552 (phage 7).

RESULTS AND DISCUSSION

Selection of phages for this study. Eleven phages were selected for this study in order to assess genome diversity among phages belonging to the 936 species. The phages represent a broad range of 936 phages isolated across Europe (and one New Zealand phage) during a time period that spanned from the 1980s until 2010 (Table 2). First, phages P113g, P272, P680, and 645 were selected, as they had been used in several phage-host interaction studies conducted over the past decades (9, 17, 25, 26). Second, phage 936^P was selected to serve as the prototype member of the 936 phage species for comparative purposes. Finally, considering the geographical origin and year of isolation of the phages, those that had been isolated over the past 30 years in Ireland (JM2 and JM3), Germany (P113g, P272, and P680), Denmark (fd13, 645, 340, and 7), other European regions (P475), and New Zealand (936^P) were selected (Table 2).

Host range analysis. Thirty-six lactococcal strains (Table 1) were assessed for their sensitivity to the 11 phages sequenced in this study. All phages assessed in this study have a relatively narrow host range and are limited to infecting at most 6 different strains from this panel of 36 possible hosts. It is also noteworthy that host range overlap was observed for certain members of the sequenced group of phages. For example, the host ranges of P475, 645, and 340 overlap those of P680, P272, and P113g (Table 4). Since these phages are derived from similar geographical locations, it is perhaps unsurprising that these apparent subgroups of phages possess related host ranges. Phages JM2, JM3, fd13, and 936^P display very narrow host specificities, infecting only a single strain among those tested.

Genome sequence analysis. The genome sequences of the 11 936-type phage isolates were determined, and each of the genomes was analyzed and compared to each other to define subgroups within this close-knit group of phages (see Table S1 in the supplemental material). It is evident that the overall architecture and functional organization of the genomes are maintained relative to previously sequenced 936-type phages (3–6, 8). The prototype of the species, phage 936^P, was used as the primary comparator (see Table S1 in the supplemental material), and its genome size represents the smallest of those sequenced in this study (Table 2). The genome sizes of the phages of this study are in line with previously sequenced 936-type phage genomes, all being 30 ± 3 kb (Table 2). The general characteristics of the phages are presented in Table 2.

TABLE 4 Host range of phages assessed in this study[a]

Strain	Subspecies	CWPS	Infected by phage: P680	P113g	P272	936	P475	645	340	7	fd13	JM2	JM3
IL1403	*lactis*	IL/KF (B)	+	+	+		+	+	+				
F7/2	*lactis*	IL/KF (B)	+					+	+				
Bu2-60	*lactis*	IL/KF (B)	+	+	+		+	+	+				
UC77	*cremoris*	IL/KF (B)						+	+				
455	*cremoris*	IL/KF (B)					+	+	+				
3107	*cremoris*	MG/SK (C)						+	+			+	
V32.2	*cremoris*	MG/SK (C)								+			
FD13	*cremoris*	MG/SK (C)									+		
W22	*cremoris*	MG/SK (C)					+			+			
W34	*cremoris*	MG/SK (C)								+			
158	*cremoris*	MG/SK (C)				+				+			
1196	*cremoris*	MG/SK (C)								+			
E8	*cremoris*	MG/SK (C)								+			
JM2	*cremoris*	MG/SK (C)										+	
JM3	*cremoris*	MG/SK (C)											+

[a] In gray-shaded blocks are the phages that group into RBP groups I (7, fd13, JM2, and JM3), II (P680, P113g, P272, and 936), and III (P475, 645, and 340).

The genomes are arranged in three divergent clusters as follows: the morphogenesis module is situated back to back with the presumed replication module, which is head to head with the so-called "middle-expressed" region, which as yet is of unknown function (5). The morphogenesis modules of these phages are highly conserved, with only a few regions of divergence (see Table S1 in the supplemental material). These divergent regions are represented by genes that encode the predicted neck passage (NPS) or collar protein, the tail tape measure protein (TMP), the receptor binding protein (RBP), and the holin and lysin. The replication module is characteristically diverse for the sequenced phages, although there are a number of conserved open reading frames observed in this region. The middle-expressed clusters are highly conserved between most of the phages.

While the overall genomic content highlights the intricacies of this phage species, the most interesting dissection of the species lies within the subtle subdivision of the morphogenesis module based on the regions of genetic divergence, as mentioned above. These points of divergence highlight the subgrouping of the phages into those with most similarity to the sk1-like, bIL170-like, or unique/intermediate genotypes and cluster similarly with groups I, II, and III based on the RBP identified in this study (see below). For example, the TMP and RBP of phages P113g and P272 are more related to those of bIL170 and P008 (group II), while the equivalent proteins encoded by 936^{P} are more related to their sk1 counterparts (group I). This gives further substance to the notion that while the primary indicator of host range is the RBP, other features of the morphogenesis module, including the TMP, may interact with host components and potentially play a role in the infection process. It is also noteworthy that the majority of the phages possess predicted NPS/collar structural proteins (of various sizes), a feature that previously was thought to be present in the bIL170-like and absent from the sk1-like phages (see Table S1 in the supplemental material) (3). Our results show that the NPS is a rather common feature of 936-type phages and that, apparently due to its nonessential nature, the corresponding gene is absent in certain isolates, including sk1 and 936^{P}.

CWPS typing of strains by multiplex PCR. While a considerable amount of information exists relating to phages belonging to the 936 species, little is known about the specific interactions between these phages and their hosts. The host receptor for members of the 936 phage species is believed to be (part of) the CWPS, whose biosynthetic machinery is encoded by a large operon, identified in three lactococcal strains (MG1363, IL1403, and Wg2) (14, 15). Furthermore, structural studies of the 936-type phage p2 have elucidated that glycerol was bound in the saccharide binding site of the baseplate, highlighting the carbohydrate affinity of these phages (13).While this presumption gives an insight into the initial interactions of these particular phages, it does not provide any explanation regarding the specificity of the host interaction of these phages.

BLASTP analysis of the proteins encoded by the CWPS-specifying clusters of the currently available lactococcal genomes identified three major CWPS subgroups based on conserved sequences, allowing classification into the UC509.9/CV56 (CWPS type A), KF147/IL1403 (CWPS type B), and SK11/MG1363 (CWPS type C) CWPS subgroups (Fig. 1). Each subgroup is defined by unique regions that were used to develop a multiplex PCR-based typing method (see Materials and Methods). This was applied to classify the CWPS type of each of the strains used in this study. Of the 36 strains assessed, 6 strains were in this way classified as CWPS type B (17%), and 16 belonged to CWPS type C (44%), while 13 belonged to CWPS type A (36%). Three strains (W22, W34, and 158) produced weaker bands relating to CWPS type C, which may be indicative of sequence variation in the DNA region corresponding to the primers used for this multiplex PCR. One strain (*L. lactis* subsp. *lactis* 184) did not generate an amplicon for any of the three CWPS types, although the conserved region, which is present in all three subtypes, was amplified (see Fig. S1 in the supplemental material), which may be indicative of an as-yet-unidentified CWPS type not represented by the three subtypes presented in this study. The dominance of CWPS types C and A is interesting, in addition to the finding that 12 out of 16 C-type strains and 5 of the 6 B-type strains were sensitive to phages in this study, whereas none of the strains belonging to CWPS type A was infected by phages in this study.

Correlation of CWPS type and host range. Of the 11 phages, 5 phages (fd13, 936^{P}, 7, JM2, and JM3) are largely limited to infecting strains of CWPS type C (Table 4). Conversely, the remaining six phages are almost completely limited to infecting hosts with CWPS type B. There are exceptions to this generalization, however, such as phages 645 and 340, which were shown to infect *L. lactis* subsp. *cremoris* 3107 (CWPS type C) as well as *L. lactis* subsp. *lactis* IL1403 (CWPS type B) and four other strains of this CWPS type. These phages can infect strains of both CWPS types with an equal relative efficiency, indicating that the receptor for these phages is conserved in both strain types. A similar observation was made for phage P475, which infects strains of CWPS type B but which also infects *L. lactis* W22 (CWPS type C), at a much lower efficiency of plaquing (EOP) (10^{-6} to 10^{-7}). Interestingly, none of the strains possessing the type A CWPS were infected by any of the 936-type phages assessed in this study, nor are they known to be infected by 936 phages from previous studies, with the exception of KH, which was previously reported to be infected by a number of 936 phages of the sk1-like genotype (14). Host range analysis of phages is routinely performed in studies related to emerging phage isolates, and such information may be correlated to host specificity and preference when strain typing information is known (3, 8). However, the range of strains used in such studies may be limited, and currently, little information is available relating to lactococcal strain typing. Therefore, we envisage that the system discussed here may be a useful starting point for such strain typing systems with respect to lactococcal strain collections to achieve a universal understanding and classification of such strain collections and their implications for phage susceptibility patterns.

Correlation between RBP group and host range. The RBPs of the sequenced phages were used to perform a comparative sequence analysis which also included sequences of previously sequenced 936 phage RBPs. Since the amino-terminal regions of these proteins are well conserved, the first 130 residues were removed from the RBP sequences, and a comparison of the much more variable sequences of the RBPs' carboxy termini was performed. Through this analysis, three subgroups of the RBP are identifiable (Fig. 2): group I corresponds to what was previously termed the sk1-like or *L. lactis* subsp. *cremoris*-infecting phages (25), group II contains the bIL170-like or *L. lactis* subsp. *lactis*-infecting phages (25), and group III represents a small yet distinct group of phages that infect primarily *L. lactis* subsp. *lactis* strains but are also capable of infecting strains of *L. lactis* subsp. *cremoris*

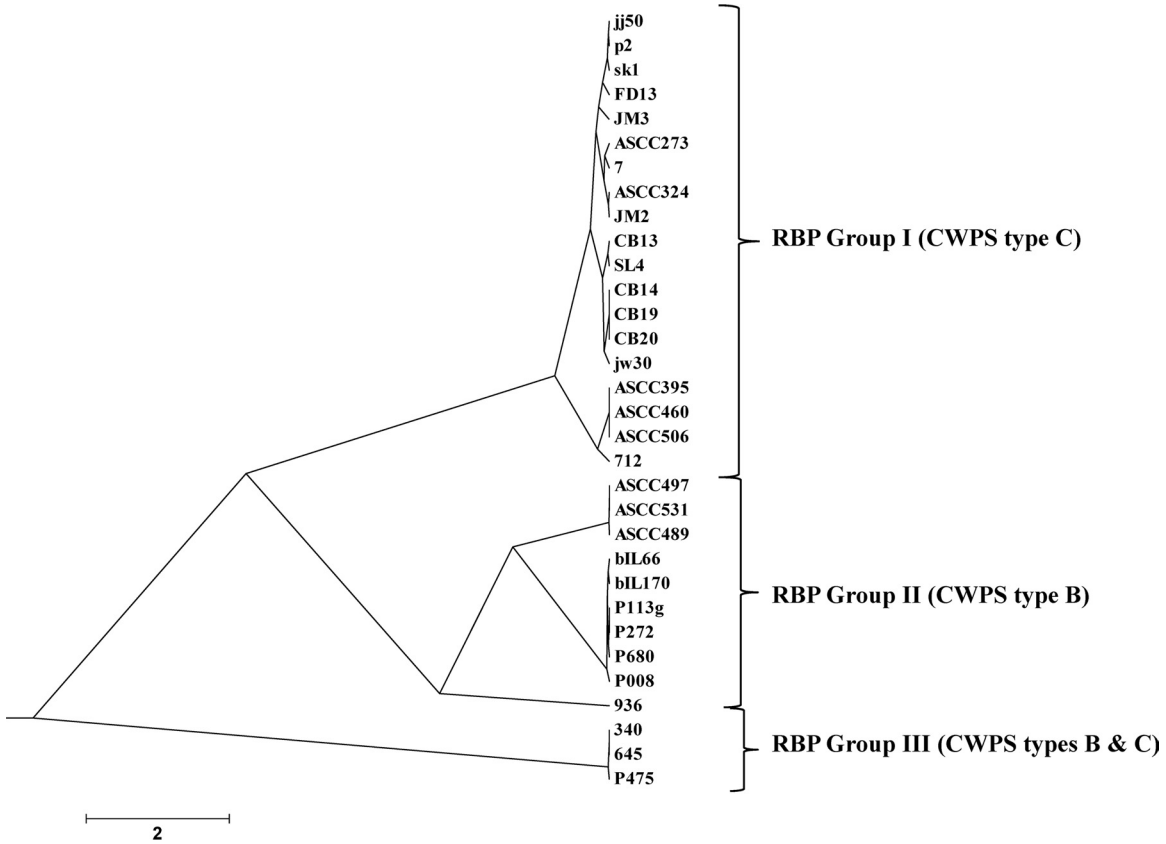

FIG 2 Representative phylogenetic tree of the RBP variable C termini of 32 936 phages, including the 11 sequenced phages of this study. Three major groups of the 936 RBPs are observed: group I includes the majority of phages in this study that infect strains with a MG/SK CWPS (except 936), group II includes those phages in this study that infect strains with an IL/KF CWPS, and group III represents a small group with a divergent RBP that predominantly infects strains with an IL/KF CWPS but also infects a strain with the MG/SK CWPS with equal efficiency.

(Table 4). While there are three main RBP groups, subtle subgroups within these groups can be distinguished (Fig. 2). The phages of RBP group I in this study infect exclusively strains that possess the type C CWPS. Similarly, the majority of phages of RBP group II infect strains that possess the type B CWPS, with the exception of 936^{P}, which infects *L. lactis* subsp. *cremoris* 158, while those belonging to group III appear to preferentially infect strains of CWPS type B, although some members of this group can also infect strains with a type C CWPS, as mentioned above (the cases of P475, 645, and 340). The exceptional case of 936^{P} may be a reflection of the potential variation in the CWPS sequence of *L. lactis* 158, which produced a weak amplicon for CWPS type C (see Fig. S1 in the supplemental material). This highlights the possibility of unique relationships, the specifics of which are as yet unknown.

Furthermore, since a link can be made between the RBP type of a phage and the CWPS-type strain that it will infect, a very controlled approach to defining dairy starter culture rotations could be implemented rapidly without the need for full genome sequencing of phages or starter strains in any given dairy facility. This reduces both the time taken to design such a starter rotational strategy and the cost of applying such systems. However, while the requirement for full genome sequencing may not be relevant to industry in all cases, there is a further wealth of information that dissects this group of phages into subclasses within the genome content. Such information includes the presence of particular features such as the neck passage structures (NPSs) of these phages, which have long been implicated in host range determination or extension (3, 27). Additionally, while the relationship between phage-encoded RBPs and the host CWPS type is a useful indicator of potential phage-host interactions, it must be mentioned that factors such as host-encoded phage resistance mechanisms may limit the host range of phages either within a collection of strains of a CWPS type or between different types, which may explain the narrow host range of some of these phages (26, 28–33). Furthermore, intricacies and additional unique genes within the CWPS biosynthesis cluster in as-yet-unsequenced strains may present an additional factor to explain why 936 phages of a certain RBP type cannot infect all strains of a particular CWPS type. It must also be considered that phages belonging to RBP group III may infect strains of the B and C CWPS types. This would indicate that these phages may recognize a component of the CWPS polymer that is common to all of these strains, for example, the saccharidic component, whose biosynthesis is directed by the predicted rhamnosyl transferases, which are common to all CWPS biosynthesis clusters (Fig. 1). In contrast, phages belonging to RBP groups I and II are largely limited to infecting strains of a single CWPS type, indicating a specific target rather than a generalized target. Furthermore, while this study relates to the interactions of 936 phages, there may be broader implications for other phage species with affinity for

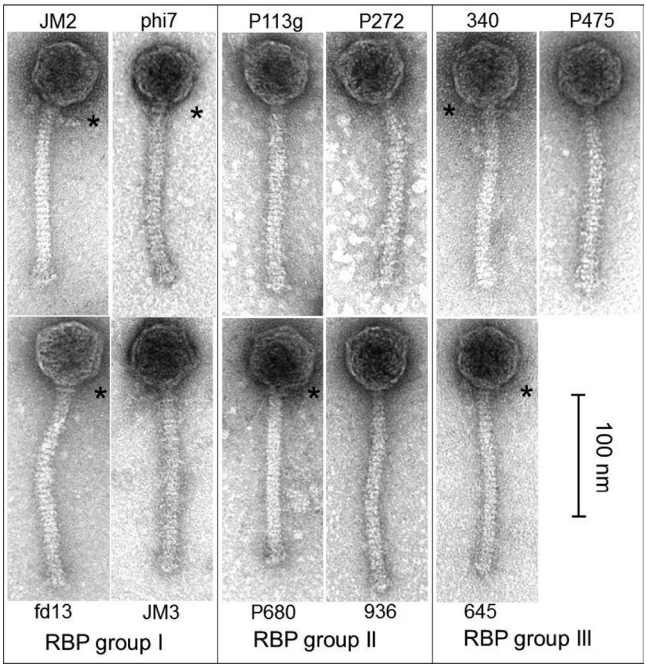

FIG 3 Representative electron micrographs of each of the 11 phages used in this study. The phages are grouped according to the RBP grouping assigned in this study (RBP groups I, II, and III). Where a collar-like structure is observed, it is highlighted with an asterisk in the image.

saccharidic receptors, including the P335 phages, although conclusive evidence of P335 receptors on lactococcal hosts remains to be identified. Structural analysis of the cell wall composition of lactococcal strain MG1363 has been performed, and the cell wall is composed of hexasaccharide repeating units containing two Glc, two GlcNAc, one Rha, and one Gal*f* residues, linked via a phosphodiester bond (15). This reported CWPS compositional analysis on a lactococcal strain provides an excellent basis for future studies relating to the identification of receptor material for phages of this and other species.

Correlation of genotype features and morphology. Each of the phages was examined with respect to their morphology and genome sequence to assess the relationship between features such as the proposed neck passage structure and the tail length correlation to the size of the TMP, as these are previously highlighted features of these phages. In this study, by electron microscopic analysis, it was determined that the tail lengths of the phages are as follows: 147 ± 2 nm ($n = 8$) for JM2, 149 ± 4 nm ($n = 5$) for 7, 162 ± 4 nm ($n = 5$) for fd13, 149 ± 4 nm ($n = 9$) for JM3, 148 ± 6 nm ($n = 5$) for P113g, 149 ± 4 nm ($n = 4$) for P272, 142 ± 5 nm ($n = 5$) for P680, 166 ± 4 nm ($n = 5$) for 936^P, 147 ± 7 nm ($n = 4$) for 340, 145.6 ± 2.7 nm ($n = 11$) for P475, and 149 ± 2 nm ($n = 4$) for 645. The longer tail length observed for fd13 and 936^P than for all other 936 phages in this study is in agreement with their larger tape measure proteins. For example, the TMP of fd13 is a 996-amino-acid (aa)-residue protein, and that of phage 936^P is 959 aa residues in length, while the TMPs of the remaining phages are 916 aa residues in length (with the exception of JM2, which is 895 aa residues long), concurring with their shorter tail lengths, as observed by EM analysis (Fig. 3). Furthermore, analysis of the presence of so-called neck passage structures (collar-like structures) between the head and tail of the phages has revealed that six of the phages appear to possess collars (645, 340, fd13, 7, P680, and JM2), while the remaining five appear not to have this structure (936^P, P113g, P272, P475, and JM3). This is largely in agreement with the genomic data in which phages 340, 645, fd13, 7, P680, and JM2 have been shown to possess ORFs that encode proteins of at least 650 aa residues, while phages JM3, P113g, P272, and P475 specify a seemingly truncated derivative of these proteins (approximately 140 to 300 aa residues), with phage 936^P lacking an analogous ORF. Therefore, it is logical that these phages exhibit no obvious collar-like features in their ultrastructure, as has been observed previously for phage 712 (3). The overall morphological data therefore correlate with the genomic data. It is notable that two distinct types of neck passage structures are distinguishable: phage JM2 (Fig. 3) possesses a collar with very short whiskers and large terminal globuli, while phages fd13 and P680 revealed a collar with significantly longer whiskers terminating with small globuli, which may be a reflection of the larger collar protein encoded by these phages than by phages 7, 340, and 645, for which collar structures but no whiskers/globuli are visible (Fig. 3). The overall dimensions and morphology of these phages are similar to those documented for 936 species phages that have previously been isolated (27, 34).

Conclusions. This study has devised a classification system for lactococcal strains based on the genetic diversity found within the CWPS biosynthesis operon and to derive a host range link between the phage population and the starter strain collection. Such a classification system could be used as a rapid approach to the grouping of starter strain collections into phage-unrelated groups that could be used in a rotational strategy to circumvent the problem of phage infection of a defined or mixed starter culture.

It is imperative that the intricacies involved in the interactions of the 936 species phages and their hosts are unraveled in order to design effective strategies to overcome the persisting phage problem in the dairy fermentation industry. Therefore, while this research has identified clear relationship markers for the 936 phages and their hosts, it is obvious that further studies are required to uncover further details concerning the molecular nature of the host-encoded receptor(s) as well as the molecular trigger(s) for DNA injection of these phages following their adsorption.

ACKNOWLEDGMENTS

Douwe van Sinderen is the recipient of a Science Foundation Ireland (SFI) principal investigator award (reference no. 08/IN.1/B1909). Witold Kot is the recipient of a Ph.D. grant from the University of Copenhagen, Denmark. James Murphy was supported by an Enterprise Ireland grant under the Innovation Partnership Scheme (reference no. IP2009-0013). Christian Cambillau is supported by a grant from the Agence Nationale de la Recherche (ANR-11-BSV8-004-01 Lactophages).

We acknowledge Zeynep Atamer and Jörg Hinrichs (University of Hohenheim, Germany) for providing phage P680 (work supported by the FEI [Forschungskreis der Ernährungsindustrie e.V., Bonn, Germany], the AiF, and the Ministry of Economics and Technology [project no. AiF-FV 14339N]).

REFERENCES

1. **Deveau H, Labrie SJ, Chopin MC, Moineau S.** 2006. Biodiversity and classification of lactococcal phages. Appl. Environ. Microbiol. **72:**4338–4346.
2. **Mahony J, Murphy J, van Sinderen D.** 2012. Lactococcal 936-type phages and dairy fermentation problems: from detection to evolution and prevention. Front. Microbiol. **3:**335. doi:10.3389/fmicb.2012.00335

3. **Mahony J, Deveau H, McGrath S, Ventura M, Canchaya C, Moineau S, Fitzgerald GF, van Sinderen D.** 2006. Sequence and comparative genomic analysis of lactococcal bacteriophages jj50, 712 and P008: evolutionary insights into the 936 phage species. FEMS Microbiol. Lett. **261:** 253–261.
4. **Castro-Nallar E, Chen H, Gladman S, Moore SC, Seemann T, Powell IB, Hillier A, Crandall KA, Chandry PS.** 2012. Population genomics and phylogeography of an Australian dairy factory derived lytic bacteriophage. Genome Biol. Evol. **4:**382–393.
5. **Chandry PS, Moore SC, Boyce JD, Davidson BE, Hillier AJ.** 1997. Analysis of the DNA sequence, gene expression, origin of replication and modular structure of the *Lactococcus lactis* lytic bacteriophage sk1. Mol. Microbiol. **26:**49–64.
6. **Crutz-Le Coq AM, Cesselin B, Commissaire J, Anba J.** 2002. Sequence analysis of the lactococcal bacteriophage bIL170: insights into structural proteins and HNH endonucleases in dairy phages. Microbiology **148**(Part 4):985–1001.
7. **Josephsen J, Andersen N, Behrndt H, Brandsborg E, Christiansen G, Hansen MB, Hansen S, Nielsen EW, Vogensen FK.** 1994. An ecological study of lytic bacteriophages of *Lactococcus lactis* subsp. *cremoris* isolated in a cheese plant over a five year period. Int. Dairy J. **4**(2):123–140.
8. **Rousseau GM, Moineau S.** 2009. Evolution of *Lactococcus lactis* phages within a cheese factory. Appl. Environ. Microbiol. **75:**5336–5344.
9. **Dupont K, Vogensen FK, Josephsen J.** 2005. Detection of lactococcal 936-species bacteriophages in whey by magnetic capture hybridization PCR targeting a variable region of receptor-binding protein genes. J. Appl. Microbiol. **98:**1001–1009.
10. **Ricagno S, Campanacci V, Blangy S, Spinelli S, Tremblay D, Moineau S, Tegoni M, Cambillau C.** 2006. Crystal structure of the receptor-binding protein head domain from *Lactococcus lactis* phage bIL170. J. Virol. **80:**9331–9335.
11. **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. Proc. Natl. Acad. Sci. U. S. A. **107:**6852–6857.
12. **Spinelli S, Desmyter A, Verrips CT, de Haard HJ, Moineau S, Cambillau C.** 2006. Lactococcal bacteriophage p2 receptor-binding protein structure suggests a common ancestor gene with bacterial and mammalian viruses. Nat. Struct. Mol. Biol. **13:**85–89.
13. **Tremblay DM, Tegoni M, Spinelli S, Campanacci V, Blangy S, Huyghe C, Desmyter A, Labrie S, Moineau S, Cambillau C.** 2006. Receptor-binding protein of *Lactococcus lactis* phages: identification and characterization of the saccharide receptor-binding site. J. Bacteriol. **188:**2400–2410.
14. **Dupont K, Janzen T, Vogensen FK, Josephsen J, Stuer-Lauridsen B.** 2004. Identification of *Lactococcus lactis* genes required for bacteriophage adsorption. Appl. Environ. Microbiol. **70:**5825–5832.
15. **Chapot-Chartier MP, Vinogradov E, Sadovskaya I, Andre G, Mistou MY, Trieu-Cuot P, Furlan S, Bidnenko E, Courtin P, Pechoux C, Hols P, Dufrene YF, Kulakauskas S.** 2010. Cell surface of *Lactococcus lactis* is covered by a protective polysaccharide pellicle. J. Biol. Chem. **285:**10464–10471.
16. **Geller BL, Ngo HT, Mooney DT, Su P, Dunn N.** 2005. Lactococcal 936-species phage attachment to surface of *Lactococcus lactis.* J. Dairy Sci. **88:**900–907.
17. **Mahony J, McGrath S, Fitzgerald GF, van Sinderen D.** 2008. Identification and characterization of lactococcal-prophage-carried superinfection exclusion genes. Appl. Environ. Microbiol. **74:**6206–6215.
18. **Lillehaug D.** 1997. An improved plaque assay for poor plaque-producing temperate lactococcal bacteriophages. J. Appl. Microbiol. **83:**85–90.
19. **Sambrook J, Russell DW.** 2001. Molecular cloning: a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
20. **Kampmann ML, Fordyce SL, Avila-Arcos MC, Rasmussen M, Willerslev E, Nielsen LP, Gilbert MTP.** 2011. A simple method for the parallel deep sequencing of full influenza A genomes. J. Virol. Methods **178**(1–2): 243–248.
21. **Lukashin AV, Borodovsky M.** 1998. GeneMark.hmm: new solutions for gene finding. Nucleic Acids Res. **26:**1107–1115.
22. **Guo FB, Zhang CT.** 2006. ZCURVE_V: a new self-training system for recognizing protein-coding genes in viral and phage genomes. BMC Bioinformatics **7:**9. doi:10.1186/1471-2105-7-9.
23. **Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402.
24. **Deasy T, Mahony J, Neve H, Heller KJ, van Sinderen D.** 2011. Isolation of a virulent *Lactobacillus brevis* phage and its application in the control of beer spoilage. J. Food Prot. **74:**2157–2161.
25. **Atamer Z, Hinrichs J.** 2010. Thermal inactivation of the heat-resistant *Lactococcus lactis* bacteriophage P680 in modern cheese processing. Int. Dairy J. **20**(3):163–168.
26. **McGrath S, Fitzgerald GF, van Sinderen D.** 2002. Identification and characterization of phage-resistance genes in temperate lactococcal bacteriophages. Mol. Microbiol. **43:**509–520.
27. **Crutz-Le Coq AM, Cantele F, Lanzavecchia S, Marco S.** 2006. Insights into structural proteins of 936-type virulent lactococcal bacteriophages. Arch. Virol. **151:**1039–1053.
28. **Sanders ME, Klaenhammer TR.** 1981. Evidence for plasmid linkage of restriction and modification in *Streptococcus cremoris* KH. Appl. Environ. Microbiol. **42:**944–950.
29. **Haaber J, Moineau S, Hammer K.** 2009. Activation and transfer of the chromosomal phage resistance mechanism AbiV in *Lactococcus lactis.* Appl. Environ. Microbiol. **75:**3358–3361.
30. **Tukel C, Sanlibaba P, Ozden B, Akcelik M.** 2006. Identification of adsorption inhibition, restriction/modification and abortive infection type phage resistance systems in *Lactococcus lactis* strains. Acta Biol. Hung. **57:**377–385.
31. **Boucher I, Emond E, Parrot M, Moineau S.** 2001. DNA sequence analysis of three *Lactococcus lactis* plasmids encoding phage resistance mechanisms. J. Dairy Sci. **84:**1610–1620.
32. **McGrath S, Seegers JF, Fitzgerald GF, van Sinderen D.** 1999. Molecular characterization of a phage-encoded resistance system in *Lactococcus lactis.* Appl. Environ. Microbiol. **65:**1891–1899.
33. **Forde A, Daly C, Fitzgerald GF.** 1999. Identification of four phage resistance plasmids from *Lactococcus lactis* subsp. *cremoris* HO2. Appl. Environ. Microbiol. **65:**1540–1547.
34. **Jarvis AW.** 1984. Differentiation of lactic streptococcal phages into phage species by DNA-DNA homology. Appl. Environ. Microbiol. **47:**343–349.
35. **Bolotin A, Wincker P, Mauger S, Jaillon O, Malarme K, Weissenbach J, Ehrlich SD, Sorokin A.** 2001. The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. *lactis* IL1403. Genome Res. **11:**731–753.
36. **Schafer A, Geis A, Neve H, Teuber M.** 1991. Bacteriophage receptors of *Lactococcus lactis* subsp. '*diacetylactis* ' F7/2 and *Lactococcus lactis* subsp. *cremoris* Wg2-1. FEMS Microbiol. Lett. **62:**69–73.
37. **Neve H, Geis A, Teuber M.** 1984. Conjugal transfer and characterization of bacteriocin plasmids in group N (lactic acid) streptococci. J. Bacteriol. **157:**833–838.
38. **Jarvis AW, Parker VR, Bianchin MB.** 1992. Isolation and characterization of 2 temperate phages from *Lactococcus lactis* ssp. *cremoris* C3. Can. J. Microbiol. **38:**398–404.
39. **Petersen A, Josephsen J, Johnsen MG.** 1999. TPW22, a lactococcal temperate phage with a site-specific integrase closely related to *Streptococcus thermophilus* phage integrases. J. Bacteriol. **181:**7034–7042.
40. **Braun V, Hertwig S, Neve H, Geis A, Teuber M.** 1989. Taxonomic differentiation of bacteriophages of *Lactococcus lactis* by electron microscopy, DNA-DNA hybridization, and protein profiles. J. Gen. Microbiol. **135:**2551–2560.
41. **Boyce JD, Davidson BE, Hillier AJ.** 1995. Spontaneous deletion mutants of the *Lactococcus lactis* temperate bacteriophage Bk5-T and localization of the Bk5-T attP site. Appl. Environ. Microbiol. **61:**4105–4109.
42. **Dupont K, Vogensen FK, Neve H, Bresciani J, Josephsen J.** 2004. Identification of the receptor-binding protein in 936-species lactococcal bacteriophages. Appl. Environ. Microbiol. **70:**5818–5824.
43. **Wiederholt KM, Steele JL.** 1993. Prophage curing and partial characterization of temperate bacteriophages from thermolytic strains of *Lactococcus lactis* ssp *cremoris.* J. Dairy Sci. **76:**921–930.
44. **Gasson MJ.** 1983. Plasmid complements of *Streptococcus lactis* NCDO 712 and other lactic streptococci after protoplast-induced curing. J. Bacteriol. **154:**1–9.