

SCIENTIFIC REPORTS



OPEN

Molecular, physiological and phylogenetic traits of *Lactococcus* 936-type phages from distinct dairy environments

M. Chmielewska-Jeznach, J. K. Bardowski & A. K. Szczepankowska

Bacteriophage infection of *Lactococcus* species can cause serious disruption of dairy fermentation processes. The most common isolates from the dairy environment are *Siphoviridae* lytic 936-type phages. To gain specific knowledge about this group of phages in Polish dairies, we examined 90 isolates from 8 different locations. Based on restriction fragment length polymorphism analysis, coupled with physiological and molecular studies, the isolated phages were divided into 8 distinct groups. Whole-genome sequencing of single representatives from each phage group provided data about their biology and genetic composition. The phages present an overall conserved genome organization. High sequence homology to another Polish isolate, *Lactococcus* phage bBB29, indicates their close phylogenetic relatedness to this strain. Such similarity may be suggestive of a general genome conservation among phages persisting in Polish dairies. Comparative genome analyses with other 936-type phages revealed several discriminative traits, including the presence and position of HNH endonuclease genes, varying number of *orfs* in the early gene region, and a putative TpeX gene. Interestingly, host range of the sequenced phages was restricted to *L. lactis* subsp. *lactis* biovar. *diacetylactis* strains. The results provide new data regarding phages present in the Polish dairy environment and permit analysis of their biology, genome composition and relatedness to other *Lactococcus* 936-type phages.

The majority of starter cultures used in the dairy industry for the production of sour cream, buttermilk, kefir or cheese contain, among other microorganisms, a mixture of *Lactococcus lactis* (*L. lactis*) strains. The role of starter bacteria is to promote fermentation, in which lactose is metabolized into lactic acid, with a simultaneous reduction of pH. In cheese production, approximately 1–10% of fermentations are slow or fail to produce lactic acid and do not reach the pH required for proper product formation¹. It is estimated that 60–70% of technological disruptions encountered during industrial milk processing are caused by bacteriophage infections of sensitive lactococcal starter strains². Despite various precautions, phage infections pose a real threat in milk factories^{3–6}. The dairy environment serves as a constant source of phages, which may originate from raw milk, growth supplements, spontaneous prophage induction, factory equipment, or workers⁷. New phages can also arise in a factory through mutations and through recombination events between existing phages that infect the same sensitive host⁸. *Lactococcus*-infecting phages are among the most commonly isolated phages infecting any bacterial species⁹.

L. lactis phages belong to the order *Caudovirales* and have double-stranded DNA (dsDNA) genomes. According to the classification proposed by Deveau *et al.*¹⁰, lactococcal phages are grouped into ten distinct types based on nucleotide sequence homology. Eight of these phage types belong to the *Siphoviridae* family (with long, non-contractile tails and isometric or prolate heads), and the remaining two belong to *Podoviridae* (with short, non-contractile tails and isometric heads). Most lactococcal bacteriophages isolated from failed fermentations share morphological features with members of *Siphoviridae* and fall into one of the three most prevalent groups: 936, c2 and P335¹¹. Additionally, the 936 and c2 types consist of only virulent bacteriophages, whereas the P335 group includes both virulent and temperate phages. The remaining *L. lactis* phage groups (1358, Q54, P087, 949, 1706, P034, and KSY1) are isolated less frequently, more often from raw milk than from failed fermentations¹⁰.

Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Pawinskiego 5a, 02-106, Warsaw, Poland. Correspondence and requests for materials should be addressed to A.K.S. (email: agaszczep@ibb.waw.pl)

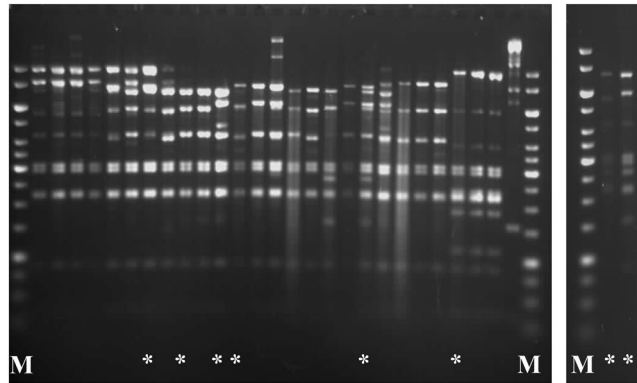


Figure 1. Restriction enzyme profiles of representative *Lactococcus* phages analysed in this study. DNA samples were cut overnight with *EcoRV* at 37°C and treated with 50% formamide to dissociate *cos* ends prior to electrophoresis. M: 1-kb DNA Ladder (Fermentas). Unique phage DNA restriction patterns are marked by an asterisk (*).

As shown by numerous studies on dairy phage populations, 936-type species predominate in dairy environments, with a few exceptions^{3,12–15}. This particularity is attributed to their high thermal resistance and ability to become air-borne and spread within the facility^{3,4,16}. Time-based studies on the evolution of phages in a single dairy plant provide proof of the prevalence of 936-type phages in this environment over long periods of time¹⁷. Worldwide industrial and financial consequences of fermentation failures motivated studies on the dynamics of this specific bacteriophage community (its abundance, development, diversity, evolution, etc.). Various methods, including plaque assays, morphological characterization, host range, burst size, restriction fragment length polymorphism (RFLP) analysis, DNA-DNA hybridization, PCR-based approaches and, finally, the recently developed high-throughput sequencing, have been employed to characterize industrially emerging lactococcal phage isolates^{8,10,12,18–24}. A better understanding of the bacteriophages that are prevalent in dairy plants is considered a crucial aspect of controlling phage infections in this environment. In an attempt to characterize 936-type phages collected during a 3-year period from 8 dairies located in various geographical regions in Poland, we initially screened 90 isolates for RFLPs. Based on distinct restriction patterns, phages were divided into 8 groups, and a representative from each group was subjected to whole-genome sequencing. The obtained data provide further information about the phages present in the milk environment, allowing the elucidation of their biology and genetic components and their comparison to already known 936-type phages that infect *Lactococcus* spp.

Results

RFLP profiling of phage isolates from whey samples. Ninety phages isolated from disturbed fermentation samples from 8 factories were subjected to RFLP analysis, using chosen restriction endonucleases. In certain cases, the digested phage DNA samples varied only by a single band, making their discrimination rather difficult. The most differential and apparent DNA patterns were obtained by *EcoRV* digestion (Fig. 1). Through this analysis, 8 distinct phage groups with unique DNA restriction profiles were distinguished. Single representatives from each restriction group were subjected to whole-genome sequencing.

Host range and infection dynamics. Phages selected for sequencing, representing distinct DNA restriction profiles, were examined for their host range against a set of 60 industrial, laboratory and environmental *L. lactis* isolates (Table 1). Of these, 11 strains were sensitive to all 8 phages, 4 strains exhibited selective sensitivity, and the remaining 75% of the strains were resistant to infection (Supplementary Table S1). Specifically, no detectable disturbance of bacterial cell growth was observed for any of the *L. lactis* subsp. *cremoris* or for any of the environmental *L. garvieae*, *L. plantarum*, *L. raffinolactis* or *L. laudensis* isolates. Moreover, none of the studied phages lysed *L. lactis* subsp. *lactis* strains other than biovar. *diacetylactis* (Fig. 2). In general, the examined phages exhibited highly similar host infection patterns. Most of them produced clear plaques on sensitive strains. The exceptions were phages 5g1 and 14s, which presented varying lysis efficiencies, repeatedly giving either turbid or clear plaques on selected strains. This property was the only distinguishing feature with respect to host range among the 8 sequenced phages. Notably, all sensitive strains except IL1403 and F7/2 (reference strains) were used by the dairy plants at the time of sample collection. In respect to those strains, the infection patterns of the phages were comparable to that of the bIBB29 bacteriophage, which was isolated previously from a Polish dairy plant, and differed from that determined for other 936-type phages tested (sk1, bIL170, 712), selected as references due to their noted distinct host ranges confined to *L. lactis* subsp. *lactis* (bIL170) or *L. lactis* subsp. *cremoris* (sk1 and 712)²⁵. To gain further insight into the dynamics of infection, phage virulence was determined by one-step growth assays. The latent period of the phages was in the range of 15 ± 1 to 26 ± 3 min, indicating a fast infection rate. All phages were shown to complete the lytic cycle in 25 ± 1 to 34 ± 2 min (burst time), producing 43 ± 5 to 148 ± 6 progeny phage particles (burst size) (Table 2). No correlation between the three parameters of the phage lytic cycle was observed.

	Relevant features	Source/Reference
Bacterial strains		
IL1403	<i>L. lactis</i> subsp. <i>lactis</i> biovar. <i>diacetylactis</i> wild-type strain	67
F7/2	<i>L. lactis</i> subsp. <i>lactis</i> biovar. <i>diacetylactis</i>	68
MG1363	<i>L. lactis</i> subsp. <i>cremoris</i> wild-type strain	69
LM0230	<i>L. lactis</i> subsp. <i>cremoris</i> plasmid-free derivative of C2.	70
IBB477	<i>L. lactis</i> subsp. <i>cremoris</i> wild-type strain	71
NZ9000	<i>L. lactis</i> subsp. <i>cremoris</i> MG1363 derivative <i>pepN::nisRnisK</i> strain	MoBiTec®
IBB743, 749, 750, 752–754, 757–759, 761, 763, 764, 1784, 1796	<i>L. lactis</i> subsp. <i>lactis</i> industrial starter strains	Rhodia Food Biolacta
IBB734–738, 742, 744–748 751, 755, 756, 760, 762, 1286	<i>L. lactis</i> subsp. <i>lactis</i> biovar. <i>diacetylactis</i> industrial starter strains	Rhodia Food Biolacta
IBB694–700, 1788–1792	<i>L. lactis</i> subsp. <i>cremoris</i> industrial starter strains	Rhodia Food Biolacta
IBB338–342, 704	<i>L. lactis</i> subsp. <i>cremoris</i> natural isolates from traditional fermented milk products	IBB PAS collection
10LC	<i>L. lactis</i> natural isolate	courtesy of A. Rozga
3LC	<i>L. garvieae</i> natural isolate	courtesy of A. Rozga
76B	<i>L. plantarum</i> natural isolate	courtesy of A. Rozga
115	<i>L. laudensis</i> natural isolate	courtesy of A. Rozga
87B	<i>L. raffinolactis</i> natural isolate	courtesy of A. Rozga
Bacteriophages	Accession number*	
sk1	NC_001835.1	72
712	NC_008370.1	73
bIL170	NC_001909.1	74
c2	NC_001706.1	75
bIBBF14	MF346933	This study
bIBB5g1	MG253647	“
bIBB14s	MG253648	“
bIBB24tp1	MG253649	“
bIBB77s	MG253650	“
bIBBEg1	MG253651	“
bIBBF12	MG253652	“
bIBBF13	MG253653	“

Table 1. Bacterial strains and phages used in study. *Complete genomic sequences of *Lactococcus* phages analyzed in this study were deposited in GenBank under the given accession numbers.

Examination of phage morphology. Examination of the sequenced phages by electron microscopy confirmed their typical *Siphoviridae* morphology (Fig. 3). All 8 phages were shown to possess small isometric heads and long tails. Similar to other 936-type isolates²⁶, a large disc-like structure (baseplate) was observed at the tip of the phage tails. On several electrographs, potential spiral-like fibres/cross-striations are visible. Conversely, collar-like structures (or neck passage structures; NPS) were not observed. The results of the electron microscopy analyses are summarized in Table 3.

Phage classification by multiplex PCR. Phage typing using a PCR-based method allowed the classification of the sequenced phages as 936-type. A specific 179-bp DNA fragment was detected for the 8 phages under study, as well as for sk1 (positive control, 936-type phage), but not for *Lactococcus* phage c2 (negative control, c2-type phage) (Fig. 4). These findings corroborate those of the electron microscopy studies and confirm that the examined phages belong to the 936 type.

General features of *Lactococcus* phage genomes. Single representatives of *Lactococcus* 936-type phages representing the 8 unique DNA restriction profiles determined in this study were subjected to whole-genome sequencing. The sequenced genomes were found to consist of linear dsDNA, 29.3–30.2 kb in size, and possess a G-C content (34.6–34.9%) similar to that reported for *Lactococcus* spp. (Table 4). An overall conserved genetic organization, with three transcription regions comprising early, middle and late genes, was identified in all phages sequenced in this study and was consistent with the genome organization of other known 936-type phages^{27,28}. The identified open reading frames (*orfs*) within each region were determined to be encoded on the same strand and oriented in the same direction (Fig. 5). No tRNA or tmRNA encoding sequences were detected on either strand. In total, 53–57 *orfs* were identified. Approximately 40% of them were assigned putative functions based on homology of respective products to proteins currently present in databases. The genomes of the sequenced phages were determined to possess 11-nt cohesive ends (*cos*) at their extremities, between the middle- and late-expressed genes. The identified *cos* sites were found to be conserved and identical in sequence

Phage	Burst size [no. of particles]	Latent period [min]	Burst period [min]
5g1	43 ± 5	20 ± 0	29 ± 1
14s	54 ± 4	28 ± 3	34 ± 2
24tp1	54 ± 2	18 ± 0	29 ± 1
F12	68 ± 1	20 ± 0	31 ± 1
F14	81 ± 3	20 ± 0	31 ± 0
Eg1	95 ± 4	20 ± 0	29 ± 1
F13	135 ± 5	20 ± 0	29 ± 0
77s	148 ± 6	15 ± 1	25 ± 1

Table 2. Burst size, burst time and latent period based on one-step growth assay of the studied phages. Values are the means of three determinations ± standard deviations.

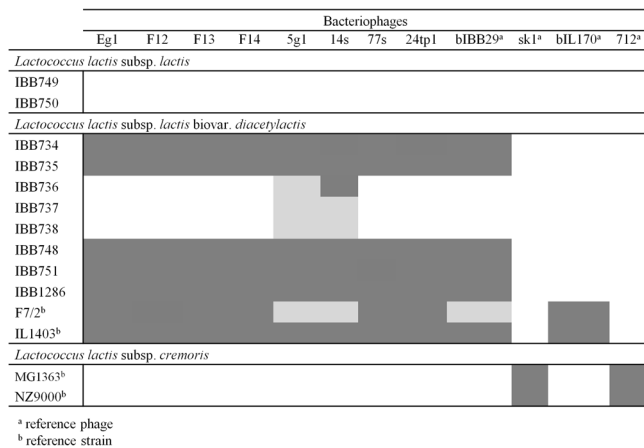


Figure 2. Host range of the studied phages tested on industrial *L. lactis* strains used at the time of whey sample collection. Grey-shaded boxes indicate clear plaques; light grey boxes indicate turbid plaques formed on the specified strains; white boxes signify lack of lysis. Phages bIBB29, sk1, bIL170 and 712 were used as references.

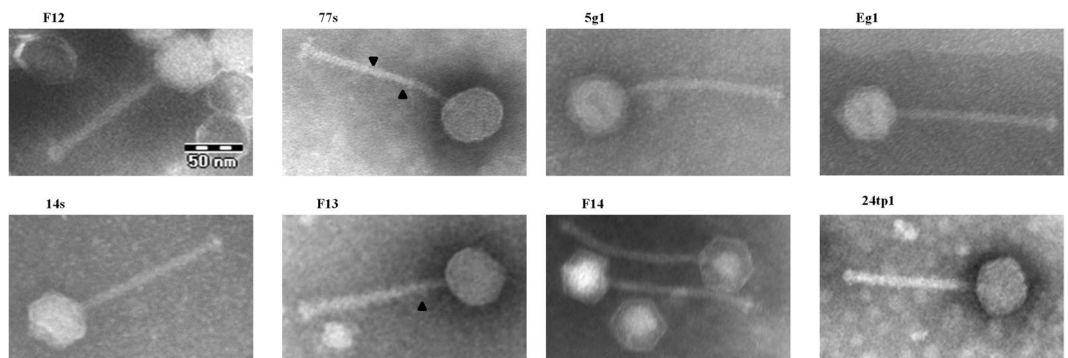


Figure 3. Electron micrographs of *Lactococcus* phages sequenced in this study. Arrowheads point to potential spiral-like structures around the phage tails.

to those of other 936-type phages (CACAAAGGACT)^{28,29}. The surrounding region was established to contain typical features described previously for 936-type *cos* sites: 10-bp inverted repeats on each side of the *cos* sites, 5–6 direct repeats (AATCT) and two complete sequences corresponding to two of the four previously identified putative terminase binding sites (R1 and R3)³⁰. Sequence alignment showed that one of the direct repeats (D6) present in 7 out of 8 studied phages (F14) was a truncated R2 terminase binding site (Fig. 6).

Based on the arrangement of the phage genomes and functional annotation of predicted *orfs*, 4 regions could be identified: (i) DNA replication/recombination in the early- and middle-expressed region and (ii) lysis, (iii) head and tail morphogenesis, and (iv) DNA packaging modules in the late-expressed region (Fig. 5). In general, Mauve progressive alignment revealed significant sequence homology and strong conservation of gene synteny between the genomes of all 8 phages examined in this study (data not shown). However, subsequent in-depth sequence analyses using bioinformatic tools allowed the mapping of genomic regions of distinct divergence.

Phage	Electron microscopy			Plaque diameter* [mm]
	Tail diam. [nm]*	Head diam. [nm]*	Additional structures	
5g1	127 × 8	50	nd	3.6
Eg1	143 × 12	58	nd	4.2
F12	132 × 10	55	nd	3.8
F13	135 × 11	55	spiral-like structures	3.7
F14	131 × 9	57	nd	3.6
14s	145 × 12	65	nd	4.0
24tp1	147 × 12	66	nd	2.9
77s	137 × 10	55	spiral-like structures	4.0

Table 3. Characteristic morphological features of the studied phages. The size of phage particles was calculated using measure IT 5.0 (Olympus Soft Imaging Solutions GmbH) software. *Mean from n = 5 phages. nd – none detected.

Phage	Dairy plant	Genome length (bp)	GC content (%)	No. of predicted genes	No. of ORFs with predicted functions
Eg1	1–5	29,659	34.7	54	21
5g1	2, 4	30,132	34.6	55	22
24tp1	3	29,736	34.8	55	21
F12	6	29,454	34.9	55	21
F13	6	29,374	34.8	53	20
F14	6	29,837	34.8	56	21
14s	7	30,249	34.6	57	22
77s	8	29,412	34.8	54	20

Table 4. Main characteristics of the studied phages based on sequencing data.

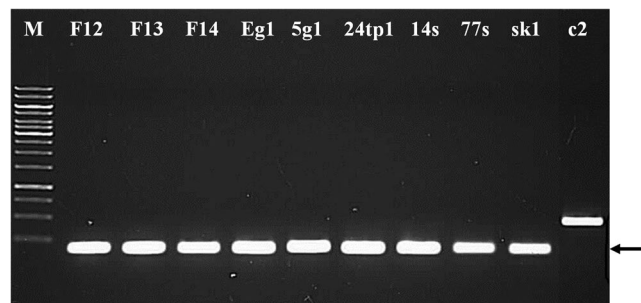


Figure 4. Phage typing using multiplex PCR. Unique fragments of phage genomes were amplified using a mix of primers complementary to the conserved genomic regions of 936- and c2-type phages. Arrows indicate the 936-type-specific band of 179 bp. Lane 1: 1-kb DNA Ladder (Fermentas), lanes 2–9: DNA fragments generated from respective phage lysates, lane 10: DNA fragment generated from phage sk1 lysate as a positive control for 936-type, lane 11: DNA fragment generated from phage c2 lysate as a negative control for 936-type.

DNA replication/recombination region. Among the 31–36 *orf*s identified in the DNA replication/recombination region of the sequenced phages, only 5–7 have been assigned putative functions based on significant similarities to entries found in public databases (Fig. 5). These included a putative single-strand DNA annealing protein (SSAP), single-strand DNA binding protein (SSB), DNA polymerase subunit, lysozyme-like superfamily protein, HNH homing endonuclease (HNHE) and Holliday junction resolvase (HJR). SSAP proteins play an essential role in recombination and proteins from the SSB family are known to play an accessory function in the process^{31,32}. The putative SSAPs and SSBs of the sequenced phages were all identical, except for phage F12, and exhibited homology (100% identity) respectively to GP42 and GP41 of *Lactococcus* phage bIBB29, which was isolated previously from a different Polish dairy plant²⁹. For phage F12, the putative SSAP (GP41_{F12}) contained short sequence deletions in the N-terminal and central region, and more closely resembled the SAK protein of *Lactococcus* phage Phi10.5 (97% identity vs. 58% identity to GP42_{bIBB29}). In turn, the hypothetical SSB protein of phage F12 (GP40_{F12}) contained single amino acid substitutions throughout the central and C-terminal part, compared to the remaining phages, and was most homologous (99% identity) to the putative SSB of *Lactococcus* phage Phi19.3. Interestingly, an *orf* identified just upstream of *gp41*_{F12} was determined to encode a putative HNHE protein (GP42_{F12}). The BLASTp search showed high sequence homology of GP42_{F12} to the NUMOD4 motif found in *Chlamydia trachomatis* (92% identity). The presence of a putative HNHE could indicate a possible horizontal acquisition of the

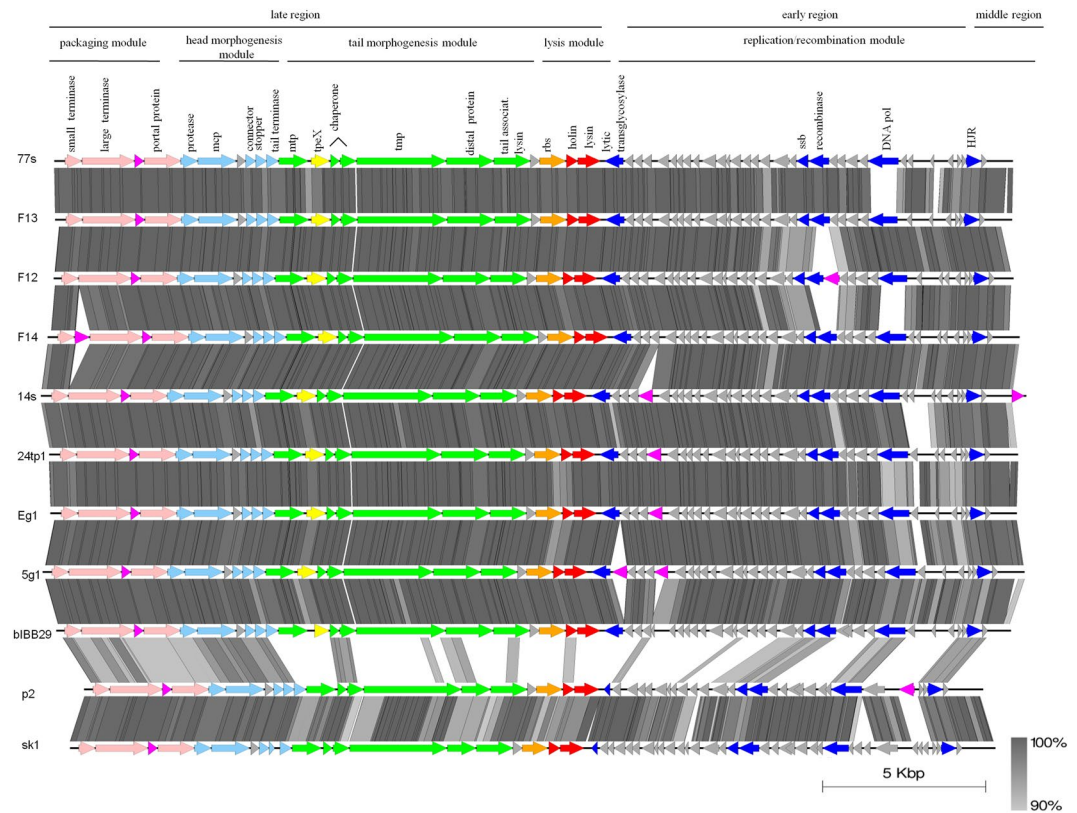


Figure 5. Schematic comparison of the *Lactococcus* phage genomes sequenced in this study. Arrows represent predicted genes and are coloured according to their putative function (indicated above each arrow). Genes encoding packaging proteins are in light pink; head morphogenesis, light blue; tail morphogenesis, green; TpeX-encoding genes, yellow, RBP-encoding genes, orange; host cell lysis, red; *orfs* of unknown function, grey; HNHE-encoding genes, pink; annotated functions of the early and middle genome region are in navy blue. The genomes of phages bIBB29, p2 and sk1 are shown as references. The grey shading between genomes represents blastx matches according to the BLAST identity scale (%).

adjacent DNA regions, including *gp40-gp41*. The third recognized function in the early-expressed gene region was DNA polymerase. In all 8 phages sequenced, this hypothetical protein best matched (97–100% identity) the putative DNA polymerase subunit of *Lactococcus* phage bIBB29 (GP47_{bIBB29}), with single amino acid differences (Supplementary Fig. S1). The exception was in phages F12 and F13, for which the predicted DNA pol exhibited high sequence divergence (45% identity) with respect to bIBB29, most closely resembling the putative DNA polymerase subunit of *Lactococcus* phages R3.4 and P680 (98% identity). A distinct feature of the early-expressed region among the 8 sequenced phages was the presence and location of putative HNH homing endonuclease genes (Fig. 5). Apart from F12, phages 14s, 24tp1, and Eg1 were also determined to possess one putative HNHE, and phage 5g1 was determined to possess two such endonucleases (GP23 and GP27). In contrast, phages 77s, F13 and F14 lacked a gene with such function in this particular genomic region. For phage 5g1, the putative protein products encoded by *gp23* and *gp27* exhibited the highest sequence homology to a predicted HNH endonuclease from *Staphylococcus fleurettii* (locus tag: B2G86_12885) (99% identity) and GP23 of *Lactococcus* phage 936 (96% identity), respectively. The three remaining phages (14s, 24tp1 and Eg1) encoded an HNH endonuclease identical to GP27_{5g1}. The HNH endonucleases predicted in the early-expressed gene region of the sequenced phages possessed a conserved zinc-binding loop domain also found in the HNHc superfamily. Proteins from this group are similar to the F-TfIII endodeoxyribonuclease of *E. coli* T5 phage, described to act on asymmetric DNA substrates (coding or template) to introduce single-strand breaks³³. HHPRED analysis allowed the determination of structural homology of the identified HNHEs to the catalytic motif of *Bacillus* phage SPO1 HNH, I-HmuI (E-values: 1.6e-30 for GP23_{5g1} and 4.8e-27 for the remaining products). The last *orf* of the early-expressed genes, directly neighbouring the late-expressed gene region, was found to encode a putative lysozyme-like superfamily protein containing a membrane-bound lytic murein transglycosylase MltE domain (COG0741). Lytic murein transglycosylases are involved in the cleavage of the β -1,4 glycosidic bond between the N-acetylglucosaminyl and N-acetylmuramoyl residues of the bacterial peptidoglycan layer³⁴. Based on studies in the phages T7, PRD1 and P1, lytic transglycosylases were suggested to play a role during the infection of host cells, possibly due to the localized destabilization of the bacterial cell wall peptidoglycan layer³⁵. The putative lytic transglycosylases of the phages sequenced in this study exhibited a high sequence similarity to the GP22 protein of phage bIBB29 (98–100% identity). At the structural level, the identified products resembled the *E. coli* EtgA muramidase (E-values: 2.9e-8 to 6.4e-10, depending on the phage). The remaining *orfs* in the early-expressed gene region could not be

	IR	cos	IR
14s	ACCGCCCTCAATCGCTATGTTAAGG	<u>GAAATTTTCAGCACAAAGGACT</u>	CCCGGCTGCGAATCC
936	CCCCCATTAAATCGCTATGTTAAGG	<u>GAAATTTTCAGCACAAAGGACT</u>	CCCGGCTGCGAATCC
77s	C-CCCCATTAAATCGCTATGTTAAGG	<u>GAGATTTTCAGCACAAAGGACT</u>	CCCGGCTGCGAATCC
24tp1	C-CCCCATTAAATCGCTATGTTAAGG	<u>GAAATTTTCAGCACAAAGGACT</u>	CCCGGCTGCGAATCC
F13	C-CCCCATTAAATCGCTATGTTAAGG	<u>GAGATTTTCAGCACAAAGGACT</u>	CCCGGCTGCGAATCC
Eg1	CCCCCATTAAATCGCTATGTTAAGG	<u>GAAATTTTCAGCACAAAGGACT</u>	CCCGGCTGCGAATCC
5g1	C-CCCCATTAAATCGCTATGTTAAGG	<u>GAAATTTTCAGCACAAAGGACT</u>	CCCGGCTGCGAATCC
F12	CCCCCATTAAATCGCTATGTTAAGG	<u>GAAATTTTCAGCACAAAGGACT</u>	CCCGGCTGCGAATCC
F14	CCCCCATTAAATCGCTATGTTAAGG	<u>GAGATTTTCAGCACAAAGGACT</u>	CCCGGCTGCGAATCC
b1BB29	CCCCCATTAAATCGCTATGTTAAGG	<u>GAGATTTTCAGCACAAAGGACT</u>	CCCGGCTGCGAATCC
sk1	CCCCCATTAAATCGCTATGTTAAGG	<u>GAAATTTTCAGCACAAAGGACT</u>	CCCTTTGAAATCTCTCA

	R3	D1	D2	D3	D4
14s	AAAAATCAAAAAGAAAAGTTT	TAGCTGGTTAAATCTTGCTAAATCT	TGCTAAATCT	TGCTAAATCT	TGCTAAATCT
936	AAAAATCAAAAAGAAAAGTTT	TAGCTGATTAAATCTAAATTAATCT	TATTTAAATCT	TAGTAAATCT	TAGTAAATCT
77s	AAAAATCAAAAAGAAAAGTTT	TAGCTGGTAAATCTTGCTAAATCT	TGCTAAATCT	TGCTAAATCT	TGCTAAATCT
24tp1	AAAAATCAAAAAGAAAAGTTT	TAGCTGGTCAATCTTGCTAAATCT	TGCTAAATCT	TGCTAAATCT	TGCTAAATCT
F13	AAAAATCAAAAAGAAAAGTTT	TAGCTGGTTAAATCTTGCTAAATCT	TGCTAAATCT	TGCTAAATCT	TGCTAAATCT
Eg1	AAAAATCAAAAAGAAAAGTTT	TAGCTGGTTAAATCTTGCTAAATCT	TGCTAAATCT	TGCTAAATCT	TGCTAAATCT
5g1	AAAAATCAAAAAGAAAAGTTT	TAGCTGGTTAAATCTTGCTAAATCT	TGCTAAATCT	TGCTAAATCT	TGCTAAATCT
F12	AAAAATCAAAAAGAAAAGTTT	TAGCTGGTTAAATCTTGCTAAATCT	TGCTAAATCT	TGCTAAATCT	TGCTAAATCT
F14	AAAAATCAAAAAGAAAAGTTT	TAGCTGGTTAAATCTTGCTAAATCT	TGCTAAATCT	TGCTAAATCT	TGCTAAATCT
b1BB29	AAAAATCAAAAAGAAAAGTTT	TAGCTGGTTAAATCTTGCTAAATCT	TGCTAAATCT	TGCTAAATCT	TGCTAAATCT
sk1	AAAAATCAAAAAGAAAAGTTT	TAGCTGATTAAATCTAAATTAATCT	TATTTAAATCT	TAGTAAATCT	TAGTAAATCT

	D5	D6/R2
14s	GCTAAATCTTGCTAAA-----	TCTTAATTATAGGGCTATATGCTCATATTTGGCTTGGA
936	GCTAAATCTTAATTAATCTCAAAATCAAAGGAAATTA	ATGGCTATATGCTAATATTTGGCTCGTAT
77s	GCTAAATCTTGCTAAA-----	TCTTAATTATAGGGCTATATGCTCATATTTGGCTTGGA
24tp1	GCTAAATCTTGCTAAA-----	TCTTAATTATAGGGCTATATGCTCATATTTGGCTTGGA
F13	GCTAAATCTTGCTAAA-----	TCTTAATTATAGGGCTATATGCTCATATTTGGCTTGGA
Eg1	GCTAAATCTTGCTAAA-----	TCTTAATTATAGGGCTATATGCTCATATTTGGCTTGGA
5g1	GCTAAATCTTGCTAAA-----	TCTTAATTATAGGGCTATATGCTCATATTTGGCTTGGA
F12	GCTAAATC-----	TAATTATAGGGCTATATGCTCATATTTGGCTTGGA
F14	GCTAAATCTTGCTAAA-TCTTGCTAAATCTTAATTA	TAGGGCTATATGCTCATATTTGGCTTGGA
b1BB29	GCTAAATCTTGCTAAA-TCTTGCTAAATCTTAATTA	TAGGGCTATATGCTCATATTTGGCTTGGA
sk1	GCTAAATCTTAATTAATCTCAAAATCAAAGGAAATTA	TATGGCTATATGCTCATATTTGGCTCGTAT

14s	GAGCAGGTTACCT--AAACAGGTA	AACTTATAAGCAAAGTTGTTGCTGTTCTTAGAAACGAGAAT
936	GAGCGTGTATATACATTACGA-TATA	ATTCAGTCTAATTAACCTAACGCTCACAGGAACGAAAAT
77s	GAGTAGAGTTACCT--AAACAGGTA	AACTTATAAGCAAAGTTGTTGCTGTTCTTAGAAACGAGAAT
24tp1	GAGCAGGTTACCT--AAACAGGTA	AACTTATAAGCAAAGTTGTTGCTGTTCTTAGAAACGAGAAT
F13	GAGCAGGTTACCT--AAACAGGTA	AACTTATAAGCAAAGTTGTTGCTGTTCTTAGAAACGAGAAT
Eg1	GAGCAGGTTACCT--AAACAGGTA	AACTTATAAGCAAAGTTGTTGCTGTTCTTAGAAACGAGAAT
5g1	GAGCAGGTTACCT--AAACAGGTA	AACTTATAAGCAAAGTTGTTGCTGTTCTTAGAAACGAGAAT
F12	GAGCAGGTTACCT--AAACAGGTA	AACTTATAAGCAAAGTTGTTGCTGTTCTTAGAAACGAGAAT
F14	GAGCAGGTTACCT--AAACAGGTA	AACTTATAAGCAAAGTTGTTGCTGTTCTTAGAAACGAGAAT
b1BB29	GAGCAGGTTACCT--AAACAGGTA	AACTTATAAGCAAAGTTGTTGCTGTTCTTAGAAACGAGAAT
sk1	GAGCGTGTATATACATTACGA-TATA	ATTCAGTCTAATTAACCTAACGCTCACAGGAACGAAAAT

	R1
14s	ATAGGCTTTAAACGATATAACCA
936	ATAGGCTTTAAACGATATAACCA
77s	ATAGGCTTTAAACGATATAACCA
24tp1	ATAGGCTTTAAACGATATAACCA
F13	ATAGGCTTTAAACGATATAACCA
Eg1	ATAGGCTTTAAACGATATAACCA
5g1	ATAGGCTTTAAACGATATAACCA
F12	ATAGGCTTTAAACGATATAACCA
F14	ATAGGCTTTAAACGATATAACCA
b1BB29	ATAGGCTTTAAACGATATAACCA
sk1	ATAGGCTTTAAACGATATAACCA

Figure 6. Comparative sequence analysis of aligned *cos* regions of *Lactococcus* phages sequenced in this study using the ClustalOmega tool. Sequences characteristic of phage 936-type *cos* regions are shaded in grey. *Cos* sites are underlined. IR: inverted repeats, R1-3: putative terminase binding sites, D1-6: direct repeats.

assigned functions using the BLASTp, HHPRED or HMMER tools but exhibited high homology (>90% identity) to proteins encoded in the same genomic positions by *Lactococcus* phage b1BB29, with some exceptions (Fig. 5). The middle-expressed gene region of 936-type phages is also postulated to encode proteins engaged in phage DNA replication/recombination events. For all phages sequenced in this study, this region was found to contain 4 small *orfs*, similar to the case in other 936-type phages²⁸. Among them, the largest (483 bp), mid-*orf* 3, was determined to encode a putative Holliday junction resolvase (HJR) with a RuvC domain, which exhibited significant homology (99% identity) to HJRs of several *Lactococcus* phages, including ASCC281, ASCC284, ASCC310, ASCC358 and ASCC365. Notably, in phage 14s a putative HNHE function was identified for the protein encoded by the last *orf* of the middle gene cluster. The hypothetical protein product exhibited high sequence identity (99%) to the HNH endonuclease of the *Lactococcus* phage P113G, which has a conserved HNHc superfamily zinc-binding loop domain. Similar to the case for the other HNHEs identified in the early gene regions, the best structural match for this protein was the HNH catalytic motif of *Bacillus* phage SPO1 I-HmuI processing endonuclease (E-value 1e-21). Rest of the products encoded in the middle-expressed region remain unannotated; no relevant sequence or structural homology to proteins existing in databases was noted.

In 936-type phages, the late-expressed region encodes products engaged in lysis, head and tail morphogenesis and packaging. Comparative genome analysis of this region among the studied phages showed an overall high

sequence conservation (>90% identity). The late regions have identical genetic organization in all 8 phages except phage F14 (Fig. 5). This gene arrangement is distinct from that of other 936-type phages²⁸, differing essentially in the presence or lack of a putative HNHE gene between the small and large terminase subunits. In contrast, the organization of the phage F14 late region resembles that reported earlier for phage Phi5.12²⁸.

Lysis module. The lysis module of the phage late region directly neighbours the early-expressed gene region and encodes two putative proteins, holin and lysin, that play essential roles in the liberation of phage progeny from host cells. BLASTp analysis classified the endolysin encoded by the studied phages to N-acetylmuramoyl-L-alanine amidases, which cleave the amide bond between two moieties and thus liberate progeny phage particles from host cells³⁶. The identified hypothetical holin proteins were shown to contain 3 transmembrane domains, classifying them as class I holins³⁷ (Supplementary Fig. S2).

Head and tail morphogenesis module. The morphogenesis region, located upstream of the lysis module, carries genes encoding structural proteins that are engaged in head and tail formation and host infection. Based on analogies of position and *orf* size, coupled with high sequence homology ($\geq 94\%$ identity) to the corresponding gene products of *Lactococcus* phages bIBB29 and p2, hypothetical functions were assigned to the putative *orfs* within the morphogenesis modules of phages sequenced in this study (refer to Fig. 5 for details). Among the important proteins encoded in this region is the RBP, which is essential in phage–host interactions as it recognizes receptors (carbohydrate residues) on the bacterial cell surface and plays a crucial role in the attachment of the virion to the host cell³⁸. Divergences in RBP sequence (mainly in the C terminus) have been linked with variations in host range^{39,40}. Comparative analysis of the putative RBPs of the sequenced phages revealed their 100% sequence identity, except for phages 77s, 24tp1 and 5g1, which carried a single amino acid substitution. Multiple sequence alignment with representative proteins of the currently identified RBP groups in 936-type phages⁴¹ showed high conservation in the N-terminal part and a variable C terminus, which was consistent with previous reports (Supplementary Fig. S3)^{39,40,42}. The highest similarity, spanning the whole sequence, was found for the putative RBP of phage bIBB29 (GP19) (99–100% identity) and, to a lesser extent, *Lactococcus* phages P475 (97% identity) and 645 (94% identity). Phylogenetic analysis showed clustering of the putative products to group III RBPs, confirming their significant homology (Supplementary Fig. S4). Previous studies of 936-type phage genomes reported that the greatest sequence variations within the late-expressed gene region occur in the surrounding of genes encoding the major capsid and tail proteins (MCP and MTP), the tape measure protein (TMP) and the receptor binding protein (RBP)^{28,43}. For phages sequenced in this study, the greatest divergence in the region was detected within the *orf* localized just downstream of the putative *mtp*. A search for homologues using BLASTp identified the 151-aa putative tail structural protein of *Lactococcus* phage bIBB29 (GP12_{bIBB29}) as the best match (99–100% identity), but with only 76–78% sequence coverage. However, manual correction of the start position of bIBB29 *gp12* to match that automatically detected for the corresponding *orf* of phage 77s increased the coverage value to 96–100% and identity $\geq 98\%$. Further scrupulous inspection of the size and position of this uncharacterized *orf* in respect to the morphogenesis module of other 936-type phages, indicated the presence of a potential tail protein extension (*tpeX*) gene. Multiple sequence alignment with putative TpeX sequences identified previously in other 936-type phages (i.e., Phi93, Phi145) showed homology (65% identity) in the C-terminal part (Supplementary Fig. S5). In earlier studies, TpeX was described to contribute to the formation of an alternative elongated product, major tail protein extension, as the result of a +1 translational frameshift at the end of the *mtp* gene sequence. To determine whether this is also the case for phages sequenced in this study, the intergenic region between the *mtp* and the downstream putative TpeX-encoding *orf* was examined for signals of a possible translational shift. The analysis revealed a CCC.UAG sequence at the end of the *mtp* gene, which was previously described as the site of a ‘shifty stop’ for several other 936-type phages (e.g., PhiE1127, PhiB1127, Phi5.2, Phi4.2, PhiCO139)²⁸. Bioinformatic simulation of the +1 translational frameshift in this DNA region for phages sequenced in this study (CCC.UAG → CCC.UAG) permitted the formation of largely identical (with single amino acid differences) putative protein products (520 aa), which best matched the major TpeX of *Lactococcus* phage Phi5.12 (68–69% identity, 99% coverage) and Phi 4.2 (65% identity, 100% coverage). Close clustering of these putative protein products was confirmed by maximum-likelihood phylogenetic analysis (Supplementary Fig. S6). Notably, HHPRED analysis indicated that the C-terminal part exhibited high structural homology (E-values: 1.1e-34 to 6.7e-35) to the minor structural protein 5 (5E7T_B) of *Lactococcus* phage TUC2009 (P335 species), described to be a component of the baseplate structure that is involved in binding to host receptors^{44,45}. It is plausible that these proteins could play a role in the phage–host interactions of the newly sequenced phages. The presence of the MTP extension protein was found to manifest as spiral-like structures along the tail that were visible under an electron microscope. Similar structures were also observed in several of the phages under study (e.g., 77s, F13) (Fig. 3). However, both the +1 frameshift and the formation of a larger protein, as well as the localization of the protein (in spiral-like forms or in baseplate structures), require further experimental investigation.

Packaging module. The packaging module is located upstream of the morphogenesis region and contains genes encoding the portal protein and the small and large terminase units. Together, they form the DNA packaging complex that recognizes and cuts phage DNA concatemers into properly sized genome fragments, which subsequently pass through a hole formed by the portal protein into the empty phage capsid⁴⁶. The recognized putative small and large terminase units and the portal protein are identical among the studied phages and exhibit the highest sequence homology (100% identity) to corresponding proteins encoded by *Lactococcus* phage bIBB29. One exception is the predicted large terminase subunit of phage F14, which best matches proteins from other lactococcal phages, including bIL170 (99% identity). The packaging module of 936-type phages may carry additional genes, such as those encoding homing endonucleases or methyltransferases (MTases)²⁸, which are often

positioned between the genes encoding the large and small terminase units. Among the studied phages, only F14 was found to encode a putative HNH endonuclease in this position (with a partial HNH_3 domain in the N terminus and HNH_ASNC domain in the C-terminal part). The predicted product presented high structural similarity (E-value 6.8e-27) to the HNH catalytic motif of *Bacillus* phage SPO1 I-HmuI. Additionally, all 8 phages encoded a highly conserved protein with putative HNH activity (including an HNHc family domain) directly upstream of the portal function. Interestingly, the two best matches of this sequence (100% identity) were the putative HNH endonucleases of *Lactococcus* phage BIBB29 (GP03) and of the gram-negative bacterium *Porphyromonas macacae* (locus tag: HR11_02520). Moreover, strong structural resemblance (E-value 3.7e-12) to the HNH protein of thermophilic phage E2 of *Geobacillus* sp. (5H0M) was noted.

Genome-based phylogeny of the sequenced phages. To analyse the genetic relationship between the sequenced phages and other 936-type phages, an unrooted phylogenetic tree was built using the maximum-likelihood (ML) method. For comparison, we chose representative 936-type phages, which based on multiple genome alignments were previously reported to form distinct clades²⁸. The obtained results show clustering of the phages examined in this work, which confirms their high genome similarity (Fig. 7). Albeit closely associated, the 8 phages are shown to be divided into two branches and additional sub-branches. Moreover, a clear association with BIBB29 was noted. Conversely, the ML tree indicates a separation of the studied phages from other 936-type phages, including those previously assigned to the same cluster with BIBB29²⁸. Interestingly, several of the sequenced phages that were isolated from different geographical locations (Table 4) are shown to cluster closely together (e.g., 5g1 with 24tp1 and F14 with 77s and 14s). Conversely, phages obtained from the same location at the same time point (F12, F13, F14) were observed on different sub-branches of the phylogenetic tree. Based on comparative analyses of whole-genome sequencing data obtained in this study, the highest level of phage diversity was assigned to the early gene regions. Dot plot analyses confirmed the observed genome-wide sequence homology between all 8 phages (data not shown). Stretches of disturbed nucleotide identity were mapped to three parts of the early gene region (upstream of the putative DNA pol encoding gene, directly downstream of the hypothetical *ssb* gene and approximately 1 kb upstream of the potential lytic transglycosylase gene) and to several sites of the late region. More generally, of the 29 genes reported by Murphy *et al.*²⁸ to constitute the core genome of 936-type phages, half of them were identical between the studied phages. Despite the high homology of the genome sequences, several features (described earlier) were recognized as differentiating factors among the studied phages. These included (i) the varying number of *orfs* upstream of the putative DNA polymerase gene and (ii) the presence and position of *orfs* encoding putative HNHEs. With regards to the latter, only phage pairs Eg1 and 24tp1 and 77s and F13 exhibited the same genetic organization, encoding either one or no HNHE homologues, respectively. The distinct identity of these phages was determined by running BLASTn analysis on their genome sequences, which revealed high, but not identical, sequence similarity (for 24tp1 and Eg1: 99% identity with 99% coverage, for 77s and F13: 97% identity and 99% coverage). The close clustering of the studied phages indicates strong conservation of phage DNA sequences (>90% identity) in the dairy environment. This finding is quite intriguing, given that the phages sequenced in this study were isolated at varying points in time from different dairy plants and geographical regions across Poland.

Discussion

Phage contamination of dairy factories remains one of the main causes of fermentation failures. Infections of *L. lactis* strains, used mainly in the production of cheese, buttermilk and sour cream, constitute the largest part of the technical problems experienced in this field. The majority of phages infecting the *L. lactis* species in milk plants identified thus far are 936-type phages⁸ (and references within). The persistence of 936-type phages in this environment is due at least in part to their high thermal resistance and volatility and the use of phage-sensitive starter strains⁴⁷. The characterization of lactococcal phages in terms of their morphological and molecular features, and more recently, their genetic content, is regarded as a key step in understanding phage persistence and exploring the means to reduce their presence in milk plants. Our study focused on industrially prevalent phages isolated from 8 different dairies across Poland during a 3-year period. The occurrence and diversity of phages in dairy plants from various geographical locations have been analysed in a number of works^{10,17,21,22,24}. The common features of industrial phage infection kinetics are short latent periods and relatively high burst sizes⁴⁸. The results of our one-step growth assays are consistent with this observation. Previous studies point to the importance of constant screening of phage populations to minimize their impact on fermentation processes and to permit the use of appropriate strain rotation strategies. While host range and DNA RFLP analyses are the principal techniques used in phage typing, phage characterization through whole-genome sequencing is far more valuable for providing information on newly emerged phages and for understanding the population dynamics of phages in the dairy environment^{8,9,28,49}. In this work, we sequenced 8 distinct phages identified among 90 isolates derived from Polish dairies. Comparative analysis of resultant phage genome sequences revealed several interesting traits. The most noticeable feature is the low diversity among the studied phages, regardless of the region or time of isolation. This finding is quite intriguing; however, several other studies have also reported the identification of highly homologous 936-type phages isolated from geographically separated regions. A study by Moineau *et al.*¹³ showed that 936-type phages isolated from buttermilk factories in over 20 US states presented similar DNA restriction patterns. This feature seemed to be strain independent, as different starter cultures were employed in each dairy plant. Another study by Raiski & Belyasova²² noted a widespread distribution of 936-type phages presenting similar restriction profiles. The same work showed that phages with identical DNA digestion patterns may infect different hosts. This implies a lack of strong correlation between host range and RFLP profile. Our observations seem to be in line with this hypothesis—all 8 phages essentially infected the same starter strains, yet varied in their DNA restriction patterns. This clearly shows that the determination of relationships among phages based on this factor alone may be largely misleading. In this regard, the generally accessible whole-genome sequencing

shown). We hypothesize that this determinant allows the formation of an extended 520-aa GP11-GP12 product. The significance of this genetic determinant is being evaluated; however, it can be speculated that the formation and sequence of an extended MTP protein may contribute to the host infection patterns and the dominance of these phages under specific conditions, such as those of the dairy environment. Arguments for this notion can be found when examining putative TpeX sequences of phages possessing group III RBPs (645, 340, P475, bIBB29 and phages examined in this study). Compared with the 8 sequenced phages, the TpeX protein of phage bIBB29 is highly homologous ($\geq 98\%$ identity), while that of phage P475 exhibits low ($\geq 37\%$) identity. In turn, phages 645 and 340 seem to lack the *tpeX*-encoding gene. It is plausible that sequence variations in TpeX can account also for the discrepancies in infection patterns among phages representing the same RBP group. Evidence for this hypothesis are supplied by previous studies, in which phages 645 and P475 were shown to exhibit a limited host range towards *L. lactis* subsp. *lactis* biovar. *diacetylactis* strains in comparison to bIBB29^{29,50}. This observation contradicts the results obtained by Mahony *et al.*⁴³, in which phages 645 and P475 were shown to successfully infect the same biovar. *diacetylactis* strains tested earlier. An explanation for this may be that formation of MTP extension occurs only under specific conditions. Determination whether +1 frameshift demands specific factors or is an incidental process will need further studies. Only detailed mutational analyses may fully elucidate the role of TpeX in host tropism. Other potential determinants that may contribute to the dominance of phages in a specific setting are HNH homing endonucleases. Arguments for this can be found in a study by Goodrich-Blair & Shub⁵¹ involving the I-HmuI of *B. subtilis* phage SPO1, one of the structural homologues identified for the HNHEs of the currently sequenced phages. The results of this work showed that I-HmuI has higher preference for heterologous phage DNA, which may be especially advantageous during co-infection. Whether this also applies for the HNHEs of *Lactococcus* 936-type phages has yet to be determined. The overall high similarity among the sequenced phages and bIBB29 may be suggestive of a general genome conservation among the Polish dairy phages. Also previous studies report a high homology among dairy phages isolated from other countries²⁸. To determine whether this holds true also in Poland, further studies on a larger group of phages are required. The close relatedness of the sequenced phages could signify their quite recent evolution, most likely due to point mutations or recombination events that took place within the host. It may be that the analysed phages derived from a single, common ancestral phage, which was carried over to and evolved differentially in each dairy plant. A comprehensive study by Castro-Nallar *et al.*⁴⁹ posited that phages isolated from dairy plants are most likely disseminated in the environment and between individual factories by human carriers. Other previously reported potential sources of contamination, such as milk powder or whey protein concentrates, cannot be ruled out⁷. The growing number of sequenced 936-type phage genomes allows comparative analyses of their genetic content and the identification of variable regions. The current study is the first to examine and compare a set of genomes of 936-type phages isolated from the Polish dairy settings. The obtained results bring us closer to understanding the basis of the prevalence of 936-type phages in specific environments and facilitate the search for host range determinants. In conclusion, our work shows the significance of whole-genome data collected from industrially isolated phages in providing valuable insights for applied microbiology and the development of strategies for combating phage infections in dairy environments.

Materials and Methods

Bacterial strains and growth conditions. The bacterial strains and phages used in this study are listed in Table 1. *L. lactis* strains were grown in M17 broth (Oxoid) supplemented with 0.5% glucose (GM17) at 30 °C without shaking or on GM17 plates containing 1.5% agar. The phages originated from disturbed milk fermentations produced using mixed starter cultures. Whey samples were collected over a period of 3 years from dairy plants in 8 different locations across Poland. Phage propagation was performed on *L. lactis* strains using GM17 medium supplemented with 10 mM calcium chloride (GM17-CaCl₂), as described below. The obtained phage lysates were filtered and stored at 4 °C until further use.

Preparation of phage lysates and phage DNA extraction. Phage lysates and DNA were obtained from 90 isolates deriving from industrial wheys based on known methods⁵². Briefly, phages were propagated on *L. lactis* IL1403 host strain grown to an optical density (OD₆₀₀) of 0.15–0.2 and subsequently supplemented with 10 mM calcium chloride and infected with phage lysate from stock. The infected cell culture was further incubated at 30 °C or room temperature until lysis occurred. The obtained lysates were filtered using 0.45 µm filters (Stericup Millipore). RNase and DNase (Sigma) were added to the lysates to a final concentration of 20 µg ml⁻¹ each and incubated at 37 °C for 1 h. Then, NaCl and polyethylene glycol (PEG6000) were added to a final concentration of 1 M and 10% (w/v), respectively, and after complete dissolution by brisk shaking, the lysates were stored overnight at 4 °C. Phage particle precipitates were concentrated by centrifugation at 9,000 × g for 20 min at 4 °C. The supernatant was discarded, and the phage pellet was resuspended in TM buffer (10 mM Tris-HCl, pH 8, 10 mM MgCl₂). Chloroform (1:1, v/v) was added to the phage suspension and vortexed vigorously to extract PEG6000 and cell debris. The water phase containing phage particles was recovered after 15 min of centrifugation at 21,000 × g and 4 °C. Then, the following extractions were performed: once with phenol (1:1, v/v) and lithium chloride (1:10, v/v), twice with phenol alone (1:1, v/v) and once with phenol/chloroform solution (1:1, v/v). Subsequently, DNA was precipitated from the water phase with cold 96% isopropanol (1:2, v/v), concentrated by centrifugation at 21,000 × g for 30 min at 4 °C, and finally washed with 70% ethanol. The precipitates were then dried using the SpeedVac concentrator (Eppendorf) and dissolved in 100 µl TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8) overnight at 4 °C, after which the DNA was ready for analysis.

Restriction profiling of phage genomes. Phage DNA digestion and agarose gel electrophoresis were carried out as previously established⁵². Essentially, the genomic DNA of 90 phage isolates was cut separately with

EcoRI, *EcoRV* or *HindIII* restriction endonucleases (Fermentas) overnight at 37 °C as recommended by the manufacturer. Enzymes were selected based on their ability to recognize and cleave within 6 nucleotide sites of low G-C content (33%), typical for lactococcal phages and their host genomes. The digested DNA fragments were separated on a 0.7% (w/v) agarose gel containing ethidium bromide in 1xTAE buffer (Merck). DNA loading dye with 50% formamide was used to load the samples on the agarose gel. Formamide was added to separate the phage cohesive (*cos*) ends⁵³. Comparative analysis based on DNA restriction patterns allowed the separation of the phages into distinct restriction groups (8 in total). Representatives from each group were taken for sequencing.

Host range determination. The infection specificity of the 8 sequenced phages was tested against a total of 60 *L. lactis* strains, including industrial⁴³, laboratory⁶ and environmental¹¹ isolates from the IBB PAS collection. Individual phage lysates (5 µl) were spotted separately on double-layer GM17-CaCl₂ plates seeded with 200 µl of overnight-grown culture of selected bacterial strains, as described elsewhere, with minor modifications⁵⁰. After overnight incubation at 30 °C, plates were examined for the presence of lytic zones, which were considered evidence of strain sensitivity.

One-step growth assay. One-step growth studies were performed essentially as described by Chow *et al.*⁵⁴. For this, the phage-sensitive *L. lactis* IL1403 indicator strain (10 ml) was grown in GM17 broth to OD₆₀₀ ~ 0.4. Then, the cell culture was centrifuged, and the pellet was resuspended in 2 ml of liquid GM17-CaCl₂ and infected with the phage at a multiplicity of infection (MOI) of 0.01. After a 5-min incubation at 30 °C (adsorption of phages to bacterial cells), the culture was centrifuged (4,500 × g, 4 min), and the cell pellet was resuspended in 10 ml of GM17-CaCl₂ and further incubated at 30 °C. Aliquots were taken at 3-min intervals until lysis of bacterial culture was complete. Serial dilutions of samples from each time point were immediately plated on the sensitive *L. lactis* host strain using the two-layer plates method, as described elsewhere⁵⁵. The phage concentration was calculated as the number of PFU per ml, determined by counting the number of visible lysis zones after overnight incubation at 30 °C. The phage burst size was calculated by dividing the average final count of liberated phage particles (at plateau) by the initial average count of infected bacterial cells. The latent period was determined as the time period prior to the release of infection particles. Experiments were repeated in triplicate.

Morphology analysis by electron microscopy. Electron microscopy analysis was performed for the 8 sequenced phages as described elsewhere²⁹. Lysates of each phage were concentrated by centrifugation (20,000 × g, 4 °C for 4 h). Then, the pellet was washed with 100 mM ammonium acetate, centrifuged (20,000 × g, 4 °C for 1 h) three times and stained using phosphoric potassium tungstate (2% w/v) on a Formvar-coated copper grid. Electron micrographs were taken in the Laboratory of Electron Microscopy at the Nencki Institute of Experimental Biology (Warsaw, Poland) using the high-performance JEM 1400 transmission electron microscope (JEOL Co., Japan, 2008) equipped with an energy-dispersive full range X-ray microanalysis system (EDS INCA Energy TEM, Oxford Instruments, UK), tomographic holder and 11 megapixel MORADA G2 TEM Camera (EMSIS GmbH, Germany).

Phage classification by multiplex PCR analysis. The confirmation of phage classification to 936-type phages was performed on phage lysates (1 µl) as templates using multiplex PCR with 936-(5'-TCAATGGAAGACCAAGCGGA-3'/5'-GTAGGAGACCAACCCAAGCC-3') and c2-specific (5'-CAGGTGTAAAAGTTTCGAGA-3'/5'-CAGATAATGCACCTGAATCA-3') primers, as described earlier⁵⁶. PCR products were separated on a 0.7% (w/v) agarose gel in 1xTAE buffer (Merck), stained with ethidium bromide, and visualized under UV light on a G:BOX instrument (Syngene).

Genome sequencing and annotation. Sequencing of bacteriophage DNA samples was performed in the Laboratory of DNA Sequencing and Oligonucleotide Synthesis (IBB PAS) using Illumina Sequencing Technology. Open reading frames (ORFs) were automatically predicted using RAST (Rapid Annotation using Subsystem Technology) software. The BLASTp algorithm⁵⁷ was run manually to assist in the functional annotation of detected *orfs* based on comparison with proteins in the non-redundant (nr) database of NCBI (<http://ftp.ncbi.nih.gov/blast/db>).

Bioinformatic analyses. Multiple sequence alignments were generated using MUSCLE (v3.7) as implemented within the Phylogeny.fr online platform (<http://www.phylogeny.fr/>)^{58,59}. Functional sequence analyses of the putative phage products of unknown function were performed using the HMMER and HHPRED programs (<https://toolkit.tuebingen.mpg.de/#/>)^{60,61}. The tRNA scanning SE and ARAGON software programs were used to search for tRNAs and tmRNAs^{62,63}. Phage genomes were visualized using Easyfig, 2.1 software⁶⁴.

Phylogenetic studies. The evolutionary distance between the studied phages (or their respective proteins) was determined using the tools implemented in the Geneious[®] 8.1.8 program⁶⁵. Multiple sequence alignments were generated using MUSCLE with 8 iterations (default settings). A maximum-likelihood tree was constructed using PHYML and the Jucker-Castor substitution model with 100 bootstrap resamplings⁶⁶.

Data availability. The data generated during this study are disclosed in this article (and its Supplementary Materials file). Complete genomic sequences of the phages analysed in this study are deposited in GenBank under the given accession numbers with restricted access until publication of this article and are available from the corresponding author on reasonable request.

References

- Bogossian, G. Control of bacteriophage in commercial microbiology and fermentation facilities in *The bacteriophages* (ed. Calendar, R.) 667–673 (Oxford University Press, New York, 2006).
- Hejnowicz, M. S. & Bardowski, J. Bacteriophage infections of mesophilic bacteria in dairy industry. *Polish Dairy J.* **7**, 2–7 (in Polish) (2005).
- Madera, C., Monjardin, C. & Suarez, J. E. Milk contamination and resistance to processing conditions determine the fate of *Lactococcus lactis* bacteriophages in dairies. *Appl. Environ. Microbiol.* **70**, 7365–7371 (2004).
- Atamer, Z. *et al.* Screening for and characterization of *Lactococcus lactis* bacteriophages with high thermal resistance. *Int. Dairy J.* **19**, 228–235 (2009).
- Marco, M. B., Moineau, S. & Quiberoni, A. Bacteriophages and dairy fermentations. *Bacteriophage* **2**, 149–158 (2012).
- Apra *et al.* Multiplex PCR to detect bacteriophages from natural whey cultures of buffalo milk and characterisation of two phages active against *Lactococcus lactis*, ϕ Apr-1 and ϕ Apr-2. *Vet. Ital.* **53**, 207–214 (2017).
- Szczepankowska, A. K., Górecki, R. K., Kolakowski, P. & Bardowski, J. K. Lactic acid bacteria resistance to bacteriophage and prevention techniques to lower phage contamination in dairy fermentation in *Lactic Acid Bacteria - R & D for Food, Health and Livestock Purposes* (ed. Kongo, M.) 23–71 (InTech, 2013).
- Rousseau, G. M. & Moineau, S. Evolution of *Lactococcus lactis* phages within a cheese factory. *Appl. Environ. Microbiol.* **75**, 5336–5344 (2009).
- Kotsonis, S. E. *et al.* Characterization and genomic analysis of phage ϕ ccphi28, a phage of the family *Podoviridae* infecting *Lactococcus lactis*. *Appl. Environ. Microbiol.* **74**, 3453–3460 (2008).
- Deveau, H., Labrie, S. J., Chopin, M. C. & Moineau, S. Biodiversity and classification of lactococcal phages. *Appl. Environ. Microbiol.* **72**, 4338–4346 (2006).
- Jarvis, A. W. *et al.* Species and type phages of lactococcal bacteriophages. *Intervirology* **32**, 2–9 (1991).
- Josephsen, J. *et al.* 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**, 123–140 (1994).
- Moineau, S. *et al.* Isolation and characterization of lactococcal bacteriophages from cultured buttermilk plants in United States. *J. Dairy Sci.* **79**, 2104–2111 (1996).
- Bissonnette, F., Labrie, S., Deveau, H., Lamoureux, M. & Moineau, S. Characterization of mesophilic mixed starter cultures used for the manufacture of aged cheddar cheese. *J. Dairy Sci.* **83**, 620–627 (2000).
- Brüssow, H. Phages of dairy bacteria. *Annu. Rev. Microbiol.* **55**, 283–303 (2001).
- Verreault, D. *et al.* Detection of airborne lactococcal bacteriophages in cheese manufacturing plants. *Appl. Environ. Microbiol.* **77**, 491–497 (2011).
- Josephsen, J., Petersen, A., Neve, H. & Nielsen, E. W. Development of lytic *Lactococcus lactis* bacteriophages in a Cheddar cheese plant. *Int. J. Food Microbiol.* **50**, 163–171 (1999).
- Jarvis, A. W. Differentiation of lactic streptococcal phages into phage species by DNA-DNA homology. *Appl. Environ. Microbiol.* **47**, 343–349 (1984).
- Moineau, S., Pandian, S. & Klaenhammer, T. R. Evolution of a lytic bacteriophage via DNA acquisition from the *Lactococcus lactis* chromosome. *Appl. Environ. Microbiol.* **60**, 1832–1841 (1994).
- Miklič, A. & Rogelj, I. Characterization of lactococcal bacteriophages isolated from Slovenian dairies. *Int. J. Food Sci. Technol.* **38**, 305–311 (2003).
- Suarez, V., Moineau, S., Reinheimer, J. & Quiberoni, A. Argentinean *Lactococcus lactis* bacteriophages: genetic characterization and adsorption studies. *J. Appl. Microbiol.* **104**, 371–379 (2008).
- Raiski, A. & Belyasova, N. Biodiversity of *Lactococcus lactis* bacteriophages in the Republic of Belarus. *Int. J. Food Microbiol.* **130**, 1–5 (2009).
- Kleppen, H. P., Bang, T., Nes, I. F. & Holo, H. Bacteriophages in milk fermentations: diversity fluctuations of normal and failed fermentations. *Int. Dairy J.* **21**, 592–600 (2011).
- Murphy, J. *et al.* Biodiversity of lactococcal bacteriophages isolated from 3 Gouda-type cheese-producing plants. *J. Dairy Sci.* **96**, 4945–4957 (2013).
- Mahony, J. *et al.* Sequence and comparative genomic analysis of lactococcal bacteriophages ϕ 50, 712 and P008: evolutionary insights into the 936 phage species. *FEMS Microbiol. Lett.* **261**, 253–261 (2006).
- Bebeacua, C., Tremblay, D., Faren, C., Chapot-Chartier, M.-P. & Sadovskaya, I. Structure, adsorption to host, and infection mechanism of virulent lactococcal phage ϕ 2. *J. Virol.* **87**, 12302–12312 (2013).
- Beresford, T. P., Ward, L. J. & Jarvis, A. W. Temporally regulated transcriptional expression of the genomes of lactococcal bacteriophages ϕ c2 and ϕ sk1. *Appl. Environ. Microbiol.* **59**, 3708–3712 (1993).
- Murphy, J. *et al.* Comparative genomics and functional analysis of the 936 group of lactococcal *Siphoviridae* phages. *Sci. Rep.* **6**, 21345 (2016).
- Hejnowicz, M. S., Golebiewski, M. & Bardowski, J. Analysis of the complete genome sequence of the lactococcal bacteriophage bIBB29. *Int. J. Food Microbiol.* **131**, 52–61 (2009).
- Parreira, R., Valyasevi, R., Lerayer, A. L. S., Ehrlich, S. D. & Chopin, M. -C. Gene organization and transcription of a late-expressed region of a *Lactococcus lactis* phage. *J. Bacteriol.* **178**, 6158–6165 (1996).
- Szczepanska, A. K. Bacteriophage-encoded functions engaged in initiation of homologous recombination events. *Crit. Rev. Microbiol.* **35**, 197–220 (2009).
- Szczepanska, A. K. *et al.* A distinct single-stranded DNA-binding protein encoded by the *Lactococcus lactis* bacteriophage bIL67. *Virology* **363**, 104–112 (2007).
- Akulenko, N. V., Ivashina, T. V., Shaloiko, L. A., Shliapnikov, M. G. & Ksenzenko, V. N. Novel site-specific endonucleases F-TfII, F-TfIII and F-TfIV encoded by the bacteriophage T5. *Mol. Biol. (Mosk.)* **38**, 632–641 (2004).
- Rydman, P. S. & Bamford, D. H. Bacteriophage PRD1 DNA entry uses a viral membrane-associated transglycosylase activity. *Mol. Microbiol.* **37**, 356–363 (2000).
- Lehnherr, H., Hansen, A.-M. & Ilyina, T. Penetration of the bacterial cell wall: a family of lytic transglycosylases in bacteriophages and conjugative plasmids. *Mol. Microbiol.* **30**, 454–457 (1998).
- Labrie, S., Vučkov, N., Loessner, M. J. & Moineau, S. Distribution and composition of the lysis cassette of *Lactococcus lactis* phages and functional analysis of bacteriophage ϕ ul36 holin. *FEMS Microbiol. Lett.* **233**, 37–43 (2004).
- Young, I., Wang, I. & Roof, W. D. Phages will out: strategies of host cell lysis. *Trends Microbiol.* **8**, 120–128 (2000).
- Samson, J. E., Magadan, A. H., Sabri, M. & Moineau, S. Revenge of the phages: defeating bacterial defences. *Nat. Rev. Microbiol.* **11**, 675–687 (2013).
- Dupont, K., Vogensen, F. K., Neve, H., Bresciani, J. & Josephsen, J. Identification of the receptor-binding protein in 936-species lactococcal bacteriophages. *Appl. Environ. Microbiol.* **70**, 5818–5824 (2004).
- Ricagno, S. *et al.* Crystal structure of the receptor-binding protein head domain from *Lactococcus lactis* phage bIL170. *J. Virol.* **80**, 9331–9335 (2006).
- Mahony, J. *et al.* Phage biodiversity in artisanal cheese wheys reflects the complexity of the fermentation process. *Viruses* **9**, 1–18 (2017).

42. Spinelli, S. *et al.* Lactococcal bacteriophage p2 receptor-binding protein structure suggests a common ancestor gene with bacterial and mammalian viruses. *Nat. Struct. Mol. Biol.* **13**, 85–89 (2006).
43. Mahony, J. *et al.* Investigation of the relationship between lactococcal host cell wall polysaccharide genotype and 936 phage receptor binding protein. *Appl. Environ. Microbiol.* **79**, 4385–4392 (2013).
44. McGrath, S. *et al.* Anatomy of a lactococcal phage tail. *J. Bacteriol.* **188**, 3972–3982 (2006).
45. Legrand, P. *et al.* The atomic structure of the phage Tuc2009 baseplate tripod suggests that host recognition involves two different carbohydrate binding modules. *MBio* **7**, e01781–15 (2016).
46. Casjens, S. R. The DNA-packaging nanomotor of tailed bacteriophages. *Nat. Rev. Microbiol.* **9**, 647–657 (2011).
47. Mahony, J., Murphy, J. & van Sinderen, D. Lactococcal 936-type phages and dairy fermentation problems: from detection to evolution and prevention. *Front. Microbiol.* **3**, 1–9 (2012).
48. Daly, C., Fitzgerald, G. F. & Davis, R. Biotechnology of lactic acid bacteria with special reference to bacteriophage resistance. *Antonie van Leeuwenhoek*. **70**, 99–110 (1996).
49. Castro-Nallar, E. *et al.* Population genomics and phylogeography of an Australian dairy factory derived lytic bacteriophage. *Gen. Biol. Evol.* **4**, 382–393 (2012).
50. Dupont, K., Vogensen, F. K. & Josephsen, J. 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 (2005).
51. Goodrich-Blair, H. & Shub, D. A. Beyond homing: competition between intron endonucleases confers a selective advantage on flanking genetic markers. *Cell*. **84**, 211–221 (1996).
52. Szczepanska, A. K., Hejnowicz, M. S., Kołakowski, P. & Bardowski, J. Biodiversity of *Lactococcus lactis* bacteriophages in Polish dairy environment. *Acta Biochim. Pol.* **54**, 151–158 (2007).
53. Lillehaug, D., Lindqvist, B. & Birkeland, N. K. Characterization of phiLC3, a *Lactococcus lactis* subsp. *cremoris* temperature bacteriophage with cohesive single-stranded DNA ends. *Appl. Environ. Microbiol.* **57**, 3206–3211 (1991).
54. Chow, J. J., Batt, C. A. & Sinskey, A. J. Characterization of *Lactobacillus bulgaricus* bacteriophage ch2. *Appl. Environ. Microbiol.* **54**, 1138–1142 (1988).
55. Adams, M. D. *Bacteriophages*. Interscience Publishers, Inc., New York (1959).
56. Labrie, S. & Moineau, S. Multiplex PCR for detection and identification of lactococcal bacteriophages. *Appl. Environ. Microbiol.* **66**, 987–994 (2000).
57. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
58. Dereeper, A. *et al.* Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* **36**, W465–469 (2008).
59. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
60. Söding, J. Protein homology detection by HMM-HMM comparison. *Bioinformatics* **21**, 951–960 (2005).
61. Hildebrand, A., Remmert, A., Biegert, A. & Söding, J. Fast and accurate automatic structure prediction with HHpred. *Proteins* **77**, 128–132 (2009).
62. Lowe, T. M. & Eddy, S. R. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **25**, 955–964 (1997).
63. Laslett, D. & Canback, B. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res.* **32**, 11–16 (2004).
64. Sullivan, M. J., Petty, N. K. & Beatson, S. A. Easyfig: a genome comparison visualiser. *Bioinformatics* **27**, 1009–1010 (2011).
65. Kearse, M. *et al.* Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649 (2012).
66. Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology*. **59**, 307–321 (2010).
67. Chopin, A., Chopin, M.-C., Moillo-Batt, A. & Langella, P. Two plasmid-determined restriction and modification systems in *Streptococcus lactis*. *Plasmid* **11**, 260–263 (1984).
68. Braun, V., Heretwig, S., Neve, H., Geis, A. & Teuber, M. Taxonomic differentiation of bacteriophages of *Lactococcus lactis* by electron microscopy, DNA-DNA hybridization and protein profiles. *J. Gen. Microbiol.* **135**, 2551–2556 (1989).
69. Gasson, M. J. Plasmid complements of *Streptococcus lactis* NCDO 712 and other lactic streptococci after protoplast-induced curing. *J. Bacteriol.* **154**, 1–9 (1983).
70. Eftathiou, J. D. & McKay, L. L. Inorganic salts resistance associated with a lactose-fermenting plasmid in *Streptococcus lactis*. *J. Bacteriol.* **130**, 257–264 (1977).
71. Radziwill-Bienkowska, J. M. *et al.* Adhesion of the genome-sequenced *Lactococcus lactis* subsp. *cremoris* IBB477 strain is mediated by specific molecular determinants. *J. Appl. Microbiol. Biotechnol.* **100**, 9605–9617 (2016).
72. Chandry, P. S., Moore, S. C., Boyce, J. D., Davidson, B. E. & Hillier, A. J. 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 (1997).
73. Rincé, A., Tangney, M. & Fitzgerald, G. F. Identification of a DNA region from lactococcal phage sk1 protecting phage 712 from the abortive infection mechanism AbiF. *FEMS Microbiol. Lett.* **182**, 185–191 (2000).
74. Crutz-Le Coq, A. M., Cesselin, B., Commissaire, J. & Anba, J. Sequence analysis of the lactococcal bacteriophage bIL170: insights into structural proteins and HNH endonucleases in dairy phages. *Microbiology*. **148**, 985–1001 (2002).
75. Lubbers, M. W., Waterfield, N. R., Beresford, T. P., Le Page, R. W. & Jarvis, A. W. Sequencing and analysis of the prolate-headed lactococcal bacteriophage c2 genome and identification of the structural genes. *Appl. Environ. Microbiol.* **61**, 4348–4356 (1995).

Acknowledgements

We would like to extend our thanks to Szymon Suski from the Laboratory of Electron Microscopy at the Nencki Institute of Experimental Biology (Warsaw, Poland) for help with electron micrographs and Janek Gawor from the Laboratory of DNA Sequencing and Oligonucleotide Synthesis, IBB PAS for carrying out phage genome sequencing. This work was supported by the European Funds Portal Innovative Economy “Centre of medicinal product biotechnology. Package of innovative biopharmaceuticals for human and animal therapy and prophylactics.” POIG.01.01.02-00-007/08-06.

Author Contributions

M.Ch.-J. carried out experimental studies and genome annotations. She also participated in drafting of the manuscript and figure preparation. A.K.S. was responsible for coordinating the experiments, critical assessment of results, physiological and bioinformatic studies, and wrote the main body text of the manuscript. J.K.B. was involved in critical review of the manuscript and supervision of on-going studies. All authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-30371-3>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018