International Journal of Food Microbiology 446 (2026) 111509

Contents lists available at ScienceDirect

International Journal of Food Microbiology

journal homepage: www.elsevier.com/locate/ijfoodmicro

Biodiversity of skunaviruses isolated from dairy fermentation facilities employing undefined mesophilic starter cultures

Kelsey White [a,b], Giovanni Eraclio [c], Gabriele Andrea Lugli [d,e], Marco Ventura [d,e], Natalia Biere [f], Frank Hille [f], Federica Volonté [c], Charles M.A.P. Franz [f], Christian Cambillau [a,g], Fabio Dal Bello [c], Jennifer Mahony [a,b,*], Douwe van Sinderen [a,b,*]

[a] School of Microbiology, University College Cork, Cork, T12 Y337, Ireland
[b] APC Microbiome Ireland, University College Cork, Cork, T12 YT20, Ireland
[c] Sacco Srl, Via Manzoni 29/A, 22071, Cadorago, Co, Italy
[d] Laboratory of Probiogenomics, Department of Chemistry, Life Sciences, and Environmental Sustainability, University of Parma, 12 - I 43121, Parma, Italy
[e] Microbiome Research Hub, University of Parma, 12 - I 43121, Parma, Italy
[f] Max Rubner-Institut, Federal Research Institute of Nutrition and Food, Department of Microbiology and Biotechnology, Hermann-Weigmann-Straße 1, D-24103, Kiel, Germany
[g] Laboratoire d'Ingénierie des Systèmes Macromoléculaires (LISM), Institut de Microbiologie, Bioénergies et Biotechnologie (IMM), Aix-Marseille Université – CNRS, UMR 7255, Marseille, France

ARTICLE INFO

Keywords:
Lactococcus
Bacteriophage
Skunavirus
Receptor binding protein
Carbohydrate binding domains
Phageome

ABSTRACT

Mesophilic starter cultures rely on the activity of lactic acid bacteria such as *Lactococcus lactis* and *Lactococcus cremoris* to acidify milk in dairy fermentations. Bacteriophage infection of starter bacteria remains one of the most significant threats to a successful fermentation, as they may cause costly disruptions and loss of production. The exclusively virulent skunaviruses represent one of the most problematic lactococcal phage genera in dairy fermentation facilities and have consequently been studied extensively in recent decades. In the present study, the diversity of lactococcal phages in whey samples originating from fermentations employing undefined mesophilic starter cultures was assessed by culture-dependent phage screening, targeted isolation based on predicted receptor binding protein specificity, and **ph**age **R**BP-**a**ctivated **c**ell **s**orting (PhRACS). Through these approaches, 26 distinct *Skunavirus* isolates were characterized and their genomes were sequenced. A comparative genomic analysis and structure predictions of the proteins encoded by these 26 phages identified genes that encode tail-associated proteins with novel carbohydrate-binding domains. Furthermore, host range analysis revealed a clear correlation between specific *Skunavirus* receptor binding protein phylogroups and the cell wall polysaccharide-associated genotype of the corresponding host strain. Finally, phageome-derived *Skunavirus* contigs were analysed to determine the diversity of phages present in the dairy fermentation facilities.

1. Introduction

Bacteriophages (or phages) are viruses that infect and replicate within bacteria, and that occupy practically every ecological niche on Earth (Hendrix et al., 1999). One niche where they interfere with human food production activity is the dairy fermentation environment. Dairy fermentations primarily rely on the activity of members of the lactic acid bacteria (LAB), such as particular species of *Lactococcus*, *Leuconostoc*, *Lactobacillus*, *Pediococcus*, *Enterococcus* and *Streptococcus*, to acidify milk or to contribute to flavour formation during the maturation and ripening processes (Gobbetti et al., 2015; Tamang et al., 2016). For large-scale industrial fermentations, LAB typically form the basis of starter cultures which can possess a defined (small number of strains in a known composition) or undefined (unknown number of strains in an unknown combination) composition (Daly et al., 1996). Fermentations are also characterized based on their optimal growth temperature and can be classified as either thermophilic or mesophilic (Johnson, 2013). Mesophilic starter cultures tend to incorporate various lactococcal strains, typically *Lactococcus lactis* and *Lactococcus cremoris*, while they may also contain *Leuconostoc* spp. (Li et al., 2021).

* Corresponding authors at: School of Microbiology & APC Microbiome Ireland, University College Cork, Cork, Ireland.
E-mail addresses: j.mahony@ucc.ie (J. Mahony), d.vansinderen@ucc.ie (D. van Sinderen).

https://doi.org/10.1016/j.ijfoodmicro.2025.111509
Received 28 July 2025; Received in revised form 16 October 2025; Accepted 28 October 2025
Available online 3 November 2025
0168-1605/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Extensive efforts focused on characterising the diversity of dairy lactococcal phages have led to the identification of ten distinct taxonomic groups (Deveau et al., 2006). In fermentation environments, the three most frequently encountered lactococcal phage groups are the exclusively virulent skunaviruses (previously termed the 936 group phages) and ceduoviruses (previously termed the c2 group phages), and the P335 phages whose members can either be virulent or temperate (Mahony et al., 2017b, 2012; Millen and Romero, 2016). Infection of starter cultures by virulent lactococcal phages causes slow or failed fermentations, resulting in substantial economic losses (Samson and Moineau, 2013). Despite extensive research efforts to solve "the phage problem" over the past four decades, they remain a considerable challenge in the fermentation sector. The dominance of skunaviruses in global mesophilic dairy fermentations is reflected by the dominance of genomes of these lactococcal phages in public databases, i.e. there are currently 217 complete *Skunavirus* genomes published in GenBank (https://www.ncbi.nlm.nih.gov/labs/virus/; as of June 3, 2025) (Jolicoeur et al., 2023; Miklič and Rogelj, 2003; Moineau et al., 1996, 1992; Murphy et al., 2013; Oliveira et al., 2018; Szczepańska et al., 2007). However, studies investigating phage populations associated with artisanal cheese whey samples have found higher instances of so-called "rare" lactococcal phage groups such as *Audreyjarvisvirus* (previously the 949 group phages) and *Teubervirus* (previously termed the P087 group phages) (Mahony et al., 2017a).

The dominance of lactococcal skunaviruses in industry, paired with the availability of extensive genomic data, has facilitated thorough analysis of their genome architecture, functional assignments, host interactions and 3-D structure reconstructions of individual proteins and protein complexes (Bebeacua et al., 2013; Chmielewska-Jeznach et al., 2018; Mahony et al., 2013; Murphy et al., 2016; Sciara et al., 2010). Skunaviruses, like other siphovirus phages, possess an isometric capsid and a long, non-contractile tail (Casey et al., 1993). These are double-stranded (ds)DNA-containing phages with genomes ranging in size from ~26-32 kb, with a modular genetic architecture comprised of regions involved in DNA packaging, capsid and tail morphogenesis, and cell lysis (late-expressed genes) and replication (early and middle-expressed genes) (Bebeacua et al., 2013; Bidnenko et al., 1998; Bouchard and Moineau, 2004; Chandry et al., 1997; Haaber et al., 2009; Scaltriti et al., 2013, 2010). The genome architecture of the *Skunavirus* tail morphogenesis module is well conserved, and typically incorporates genes encoding the major tail protein, tail tape measure protein (TMP), distal tail protein (Dit), tail-associated lysin (Tal), and receptor binding protein (RBP). However, there is notable sequence diversity within confined regions of certain genes encoding the host recognition device within the morphogenesis module (Castro-Nallar et al., 2012). Within the *Skunavirus* genome, three distinct regions have been identified that exhibit relatively high genetic diversity: the first (~2 kb) corresponds to the late-expressed region encompassing the neck passage structure protein-encoding gene; the second region (~2 kb) contains the late/early gene expression junction (encoding the TMP, RBP, holin, lysin, and early expressed genes of unknown function) with the highest level of sequence variation within the gene encoding the RBP, while the third region (~2.5 kb) encompasses the start of the early-expressed region and entire middle-expressed region (Castro-Nallar et al., 2012). In addition, these regions of genetic diversity have been correlated with a higher recombination rate suggesting the important role of recombination (in addition to point mutations) in generating genetic novelty and diversity of, and associated adaptation by skunaviruses (Castro-Nallar et al., 2012).

The *Skunavirus* host adsorption apparatus or baseplate, located at the distal end of the phage tail, consists of a heteropolymeric assembly of Dit, Tal, and RBP subunits, with Dit forming a hexameric ring in which an RBP trimer attaches to each Dit monomer (Sciara et al., 2010). In skunaviruses, the C-terminal "head" domain of the RBP is responsible for the specific recognition of and binding to a surface-exposed host receptor moiety thereby commencing phage infection (Bebeacua et al., 2013). The RBP may be supported by auxiliary carbohydrate binding module (CBM)-containing proteins, located either at the distal end of the phage tail or proximal end as so-called "whiskers" that can form a neck passage structure or as translationally programmed protein extensions of the major tail protein (Hayes et al., 2019; Sciara et al., 2010). One example of these proteins includes the so-called "evolved" Dit proteins which possess an insertion encoding a CBM which has previously been shown to promote host binding (Hayes et al., 2018; Millen et al., 2023). CBMs present in the RBP and auxiliary binding protein facilitate binding to the lactococcal cell surface-exposed saccharidic receptor, also known as the **c**ell **w**all **p**oly**s**accharide (CWPS), which consists of a relatively conserved rhamnan backbone and a variable side chain, the latter also referred to as the polysaccharide pellicle or PSP (Ainsworth et al., 2014; Chapot-Chartier et al., 2010; Dupont et al., 2004; Mahony et al., 2013). There is considerable sequence diversity within the *cwps* gene cluster, responsible for the biosynthesis of these two CWPS elements, and this is consistent with the compositional and structural diversity of the associated CWPSs of lactococcal strains (particularly that of the surface-exposed PSP) (Chapot-Chartier et al., 2010; Mahony et al., 2020; Parlindungan et al., 2024; Sadovskaya et al., 2017).

In a previous study, the diversity among phages present in whey samples collected from various European dairy fermentation facilities employing mesophilic, undefined starter cultures was investigated through phageome analysis with a particular focus on RBP genotypes and RBP-associated binding specificities (White et al., 2025). To further expand upon the prevalence and diversity of phages present in these fermentation facilities, we herein report on the isolation, genome sequence analysis, and characterisation of 26 skunaviruses isolated from these whey samples. These phages were isolated through both culture-dependent and -independent techniques. The phylogeny of these novel *Skunavirus* isolates was investigated in relation to publicly available *Skunavirus* genomes. *Skunavirus* RBP head domains were shown to be correlated with the CWPS genotype of phage-sensitive hosts, as validated by binding studies. In addition, phageome-derived *Skunavirus* contigs obtained from the same samples were investigated to determine the presence of additional (partial) phage genomes that may be underrepresented in culture-based screening approaches. Assessing both the diversity of *cwps* (sub)types of strains within a starter culture as well as the RBP diversity of the *Skunavirus* population present within dairy fermentation facilities, will assist in risk management of the potential impact of phage infection on the associated starter cultures.

2. Materials and methods

2.1. Bacterial strains, bacteriophages, and growth conditions

Bacterial strains used in this study are presented in Supplementary Table S1. Lactococcal strains were grown overnight at 30 °C in M17 broth (Millipore Sigma Aldrich, Gillingham, UK) supplemented with 0.5 % lactose (LM17). Phages were propagated on appropriate host strains in LM17 broth supplemented with 10 mM $CaCl_2$ (Millipore Sigma Aldrich, Gillingham, UK) and incubated at 30 °C until lysis occurred. Phage lysates were filtered through a 0.45 μm filter (Sarstedt, Nümbrecht, Germany) and stored at 4 °C. Plaque assays were performed using the previously described double agar method for both phage isolation and host range analysis (Lillehaug, 1997).

2.2. Screening and isolation of phages from dairy-associated whey samples

32 whey samples (Supplementary Table S2) were acquired from dairy fermentation facilities employing undefined mesophilic starter cultures (typically supplemented with additional strains/cultures) where diminished milk acidification activity of starter cultures had been observed, indicating the presence of phage(s). These samples were collected between 2019 and 2023 and originated from dairy facilities

located in seven countries across Europe (Spain, Austria, France, Poland, Sweden, Denmark, and Czech Republic). The whey samples were screened for the presence of phages that infect strain(s) from a collection of 92 isolates (previously isolated from mesophilic undefined starter cultures M1 and M2; isolates listed in Supplementary Table S3) representing four *cwps* genotypes (23 isolated of each of the C_1-C_4 subtypes). Whey samples were also screened against representative strains (from an in-house strain collection) of *cwps* B and D genotypes (Supplementary Table S1). Whey samples were filtered using 0.45 μm filters prior to phage screening to remove residual bacterial debris and stored at -20 °C until use. In a 96-well plate (Sarstedt), 20 μl of a fresh overnight culture was added to a well containing 150 μL LM17 broth supplemented with 10 mM $CaCl_2$ and 20 μl of a given filtered whey sample diluted 1:20 in SM buffer [50 mM Tris-HCl, 100 mM NaCl, 10 mM $MgSO_4$, 10 mM $CaCl_2$]. A second plate using 20 μl SM buffer in place of filtered whey was prepared as a positive control for growth. Plates were incubated at 30 °C and were visually examined for lysis (by comparison to the uninfected control), initially after 6 h and following overnight incubation for the presence of presumed phage infection.

Twenty one phage-host combinations, which allowed attainment of a high titre (sufficient for DNA extraction and genome sequencing, typically $>10^7$ pfu/ml), and which represented a spread of geographic sampling locations (six countries: Denmark, France, Poland, Spain, Sweden, and Austria) and hosts and host *cwps* genotypes (17 distinct hosts covering four host *cwps* genotypes, C_1 to C_4), were selected for further analysis. Lysates of potential phages (in which culture lysis was observed either during the first 6 h or complete lysis following overnight incubation) were diluted in SM buffer (1:10), filtered (0.45 μm), and spotted on its corresponding host to validate the presence of phages through double layer plaque assays (Lillehaug, 1997). Phages were propagated from a single plaque against an appropriate lactococcal host. A single plaque was added to 10 mL of LM17 (supplemented with 10 mM $CaCl_2$) concurrently inoculated with 100 μl of the appropriate host (from a fresh overnight culture) and incubated at 30 °C until lysis had occurred. Phage lysates were filtered through a 0.45 μm filter and stored at 4 °C. The selected phages were genotyped using a previously described multiplex PCR system capable of identifying *Skunavirus*, *Ceduovirus*, and P335 phages (Labrie and Moineau, 2000).

2.3. Phage DNA extraction, genome sequencing, assembly, and annotation

Genomic DNA of selected phages (Table 1) was extracted from 2 mL of fresh phage lysate ($>10^8$ PFU/mL) using the Norgen Biotek Phage DNA isolation kit (Norgen Biotek, Ontario, Canada) according to the manufacturer's instructions. Phage genome sequences were determined by GenProbio srl (Parma, Italy) using a MiSeq platform (Illumina Inc., San Diego, CA, USA) according to the supplier's protocol. Phage genome libraries were prepared using an Illumina Nextera XT DNA Library Preparation Kit (Illumina Inc.). Libraries were quantified using a fluorometric Qubit quantification system (Life Technologies, USA), and loaded on a 2200 TapeStation instrument (Agilent Technologies, USA). Sequencing was performed using the Illumina MiSeq platform with a 600-cycle flow cell version 3 (Illumina Inc.).

Fastq files of Illumina paired end reads (250 bp) obtained from phage sequencing efforts were used as input for genome assembly. The SPAdes program v3.14 was used for de novo assembly of phage sequences with the pipeline option "–careful" and a list of k-mer sizes of 21,33,55,77,99,127 (Bankevich et al., 2012). Contigs were then employed by MEGAnnotator2 (Lugli et al., 2023) for the prediction of protein-encoding open reading frames (ORFs) using Prodigal (Hyatt et al., 2010). Predicted ORFs were functionally annotated by means of DIAMOND against the NCBI RefSeq database and INTERPRO (Jones et al., 2014) against the PFAM database. Furthermore, tRNA genes were identified using tRNAScan-SE version 2.0 (Chan et al., 2021).

2.4. Comparative analysis of phage genomes and characterisation of adhesion proteins

Comparison of individual phage proteins was performed by all-against-all, bi-directional BLASTp using default parameters (Altschul et al., 1990). Multiple alignment of the nucleotide sequences (or protein sequence) of isolated skunaviruses genomes or proteins in this study and previously sequenced Skunavirus members were performed using

Table 1
Characteristics of the skunaviruses isolated from whey samples as part of this study.

Phage	Host *cwps* type	Year of collection (whey)	Genome size (bp)	GC%	No. ORFs	Best Hit (% coverage, % nt ID)	GenBank Accession #
AD1-ME6	C_2	2019	29,359	34.84	52	37,201 (79, 94.02)	PV941749
AM2-MF12	C_4	2020	29,999	34.58	57	Phi129 (82, 95.42)	PV941750
AN3-MA5	C_2	2020	29,514	34.65	52	FB3 (84, 94.49)	PV941751
AN4-MF5	C_2	2020	30,334	34.58	52	38,503 (75, 90.44)	PV941752
AO-4†	C_6	2020	30,371	35.31	54	16,802 (89, 94.36)	PV941753
AO-MA1	C_1	2020	32,986	34.80	56	fd13 (81, 93.31)	PV941754
AO-MH4	C_2	2020	27,071	35.00	48	66,901 (90, 93.26)	PV941755
AV6-MC5	C_2	2020	31,100	34.42	56	PhiG (80, 95.08)	PV941756
BA6-B*	B	2020	29,963	35.08	57	p272 (84, 93.73)	PV941757
BA6-ME1	C_1	2020	29,849	34.46	55	p272 (85, 94.96)	PV941758
BA6-MF12	C_4	2020	29,899	34.59	58	62,601 (81, 92.30)	PV941759
BB4-D*	D	2020	29,715	34.84	53	38,507 (78, 88.52)	PV941760
BB4-MB5	C_2	2020	29,444	34.88	51	Phi4 (82, 94.45)	PV941761
BB4-MG12	C_4	2020	29,565	35.13	53	Phi29 (86, 95.12)	PV941762
BB4-SC34†	B	2020	31,874	35.20	57	P475 (84, 95.08)	PV941763
BD-MC7	C_3	2021	30,707	35.14	51	645 (72, 95.17)	PV941764
BD-MH5	C_2	2021	30,681	34.53	54	fd13 (82, 93.94)	PV941765
BQ-1†	C_3	2022	30,717	35.10	52	79,201 (84, 95.62)	PV941766
BR-MC7	C_3	2022	30,590	35.10	50	645 (72, 95.16)	PV941767
BZ-MH2	C_1	2022	30,909	34.66	53	PhiG (94, 96.39)	PV941768
CA-MA6	C_2	2023	30,197	34.61	53	PhiG (83, 94.98)	PV941769
T1-MF3	C_1	2019	29,716	34.90	51	P4565 (76, 94.78)	PV941770
T1-MH4	C_2	2019	29,002	35.26	52	66,901 (88, 92.19)	PV941771
V1-MA6	C_2	2019	30,346	34.84	57	16,802 (79, 93.18)	PV941772
V2-MB6	C_2	2019	30,279	34.66	51	37,301 (76, 93.92)	PV941773
V2-MF8	C_3	2019	29,823	35.25	50	P680 (77, 95.55)	PV941774

* Isolated from laboratory strains, this study.
† Isolated by PhRACS, previous studies (Casey et al., 2022) and (White et al., 2025).

ClustalW V2.1 software (Thompson et al., 1994). A phylogenetic tree was generated using the neighbour-joining method and bootstrapped with 1000 replicates. Both rooted and unrooted phylogenetic trees were visualised using iTOL v6 software (https://itol.embl.de/).

Additionally, 216 publicly available complete *Skunavirus* genomes were retrieved from the NCBI virus database (https://www.ncbi.nlm.nih.gov/labs/virus/) and annotated using the Pharokka v1.7.5 pipeline (Bouras et al., 2023) alongside the 26 *Skunavirus* genomes sequenced within the context of this study. Based on the gff3 annotation output files generated by Pharokka, the Roary (https://github.com/sanger-pathogens/Roary) pipeline (Page et al., 2015) was then used to calculate the pangenome of the 242 *Skunavirus* genomes with the following options employed the roary script: -s (don't split paralogs) and -i 50 (minimum percentage identity for BLASTp was set to 50 %). A binary matrix containing the presence/absence of each gene cluster was computed and used as an input for the MeV suite (v 4.9.0) (Saeed et al., 2003). The produced matrix was analysed using a two-way hierarchical clustering (HCL) to group phage genomes according to their similarity at protein level.

CAGECAT version 1.0 web platform (https://cagecat.bioinformatics.nl/tools/clinker) was used to generate alignment of phage genomes (Gilchrist and Chooi, 2021). Putative orphan methyltransferases (MTases) were analysed using REBASE (http://rebase.neb.com/rebase/rebase.html) (Roberts et al., 2009).

Skunavirus RBPs and auxiliary host binding proteins that contain CBMs (i.e. NPS, major tail protein extension (TpeX), and Dit) were identified from the *Skunavirus* genomes based on BLASTp analysis using default parameters, with an alignment (*E*-value) cut-off value of 0.0001 and greater than 50 % identity across at least 50 % of the amino acid sequence (Altschul et al., 1990). Protein sequence and structure homology analysis was performed using HHpred (Homology detection and structure prediction by HMM-HMM comparison) using default setting, in order to identify CBMs within *Skunavirus* RBPs and auxiliary binding proteins (Söding et al., 2005). Sequence alignments of the auxiliary host binding proteins were performed using Multalin (http://multalin.toulouse.inra.fr/multalin/) (Corpet, 1988) using sequences identified in this study and previously sequenced phages retrieved from the NCBI virus data base (https://www.ncbi.nlm.nih.gov/labs/virus/).

2.5. Structure prediction of phage structural proteins

Structural predictions of select auxiliary host binding proteins (NPS and Dits) of representative skunaviruses were performed by AlphaFold3 via google servers (https://golgi.sandbox.google.com) (Abramson et al., 2024). The final predicted protein or domain structures were submitted to Foldseek or Dali servers to identify closest structural homologs in the protein database (PDB) (Holm, 2022; van Kempen et al., 2024). Visual representations of predicted structures were prepared with ChimeraX (Pettersen et al., 2021).

2.6. Phageome analysis of cheese whey samples

The presence or absence of closely related *Skunavirus* genomes in 27 whey viromes [previously sequenced (White et al., 2025)] was determined through reads mapping. CoverM v0.4.0 (https://github.com/wwood/CoverM; "contig" mode) was employed to determine the covered fraction of a corresponding *Skunavirus* genome- or phageome-derived contig. The reads mapped to the contigs with identity >95 % and coverage >95 %, and only the contigs/genomes where >90 % of their length was covered by reads are reported. Hadamard scores, which represent a quantitative evaluation of the average nucleotide identity (ANI) and coverage between DNA sequences, were calculated using the ANIm method from pyANI (https://github.com/widdowquinn/pyani) (Pritchard et al., 2016) in order to assess overlaps between isolated/sequenced phages and phageome-derived contigs.

2.7. Transmission electron microscopy (TEM) of selected skunaviruses

Phage lysates of isolates T1-MH4 and V2-MF8 were purified on a discontinuous 3 M & 5 M cesium chloride (CsCl) gradient and were dialysed against SM buffer as previously described (Sambrook et al., 1989). Following cesium chloride purification, the phage samples were fixed to a carbon film on a grid for 20 min, washed twice with distilled water and stained with 2 % uranyl acetate. The images were captured using a Talos L120C transmission electron microscope (ThermoFisher Scientific, Eindohoven, The Netherlands) and a CETA 16 M camera (ThermoFisher Scientific) at an accelerating voltage of 80 kV. Length/width measurements of the phages were performed using ThermoFisher Scientific Velox software (v3.9.0).

2.8. Determination of host range and correlation with cwps genotype

To investigate possible correlations between the RBP phylogroups and corresponding host *cwps* genotypes, the host range of the 26 skunaviruses was determined using a panel of 45 strains whose speciation and *cwps* genotype had previously been determined (Supplementary Table S1). The host range strain panel was made up of a selection of 26 strains used in the original 92 isolate screening panel, while the remaining strains were sourced from our in-house strain collection (the *cwps* type and species of each strain are listed in Supplementary Table S1). Host range analysis was performed using spot assays, where 10 μl of each phage lysate (diluted 1:10 to achieve titres of between 10^8 and 10^9 PFU/mL) was spotted onto LM17 double-layer agar plates (supplemented with 10 mM $CaCl_2$) containing 200 μl of the selected lactococcal strains (from fresh overnight cultures). After incubation overnight at 30 °C, plates were examined for the presence of a zone of lysis and individual plaques were counted. The efficiency of plaquing (EOP) was calculated by dividing the titre of the phage (PFU/ml) on the tested strain by the phage titre on its respective propagating strain.

3. Results

3.1. Skunaviruses dominate whey-derived phage populations

To assess the presence and diversity of lactococcal phages in fermentation facilities employing undefined mesophilic starter cultures, phage screening was conducted on 32 whey samples (Supplementary Table S2) collected from facilities where the presence of phage(s) had been suspected due to slowed fermentation (milk acidification) issues. These 32 whey samples were collected from factories in Spain (four samples), Austria (one sample), France (four samples), Poland (three samples), Sweden (one sample), Denmark (18 samples), and Czechia (one sample) between 2019 and 2023. Although the exact composition of the undefined starter culture(s) employed in each factory at the time of sample collection is not known, two mesophilic starter cultures (M1 and/or M2) used for the production of various semi-hard cheeses, fermented milk (e.g., viili, kefir, and twaróg) and fresh cheeses (e.g., quark) may have been employed in some of these facilities as part of a starter culture rotation system. In a previous study, M1 and M2 had been shown to be primarily composed of lactococcal strains (White et al., 2024) and the *cwps* genotypic composition of these starter cultures was determined through culture-based analysis, metagenomics, and qPCR. A panel of 92 lactococcal isolates [as previously described (White et al., 2024)], representing the four most dominant *cwps* types (23 isolates for each of the C_1-C_4 subtypes) present in these undefined cultures.

Through this analysis, it was established that 29 of the 32 whey samples assessed (representing 90.6 % of samples) exhibited presumptive phage activity against at least one lactococcal strain in the panel based on lysis observed in wells inoculated with isolates compared to uninfected controls (Fig. 1). This phage screening yielded a total of 268 (of 2944 combinations tested) potential phages based on the observed culture lysis: 16 against C_1 isolates, 94 against C_2 isolates, 53 against C_3

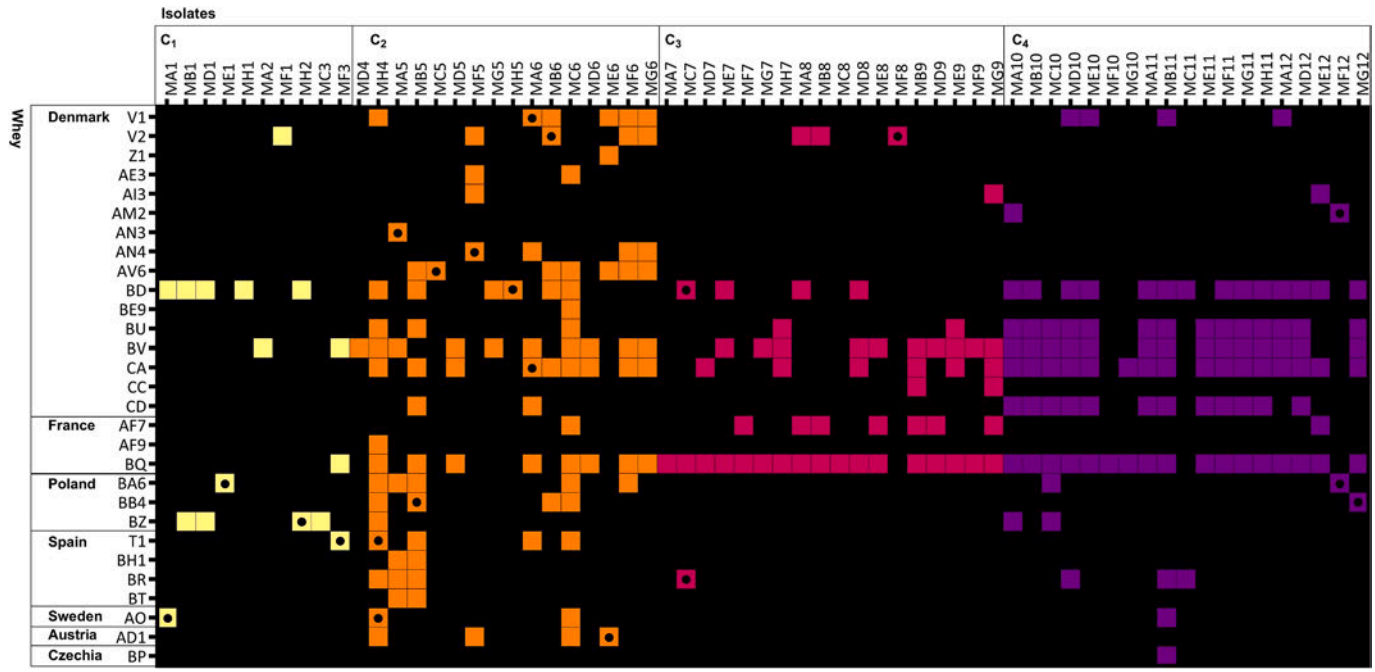


Fig. 1. Phage screening results of whey samples against a panel of lactococcal strains. Presumptive phage activity is indicated by coloured squares or absence of activity by black squares. The 22 phage-positive wells (yielding 26 distinct *Skunavirus* isolates) selected for further analysis are indicated by a black circle within a coloured square. Colouring of the squares is indicative of the *cwps* genotype of the isolate the phage activity was against; yellow for C_1, orange for C_2, pink for C_3, and purple for C_4 isolates.

isolates, and 105 against C_4 isolates. Of the 92 bacterial isolates employed in the screening panel, 63 (68 %) were infected by at least one presumptive phage. The three whey samples causing lysis of the largest number of isolates (Fig. 1) were samples BQ (43 of 92 isolates), BV (36 of 92 isolates), and CA (31 of 92 isolates); all of which were isolated in 2022 from factories in Spain (BQ) or Denmark (BV and CA).

Of these 268 presumptive phage-positive whey sample-host combinations, 21 were selected as a representation of the assessed geographic locations and host *cwps* genotypes for phage screening, purification, and subsequent genome sequencing. Following confirmation of phage presence through plaque assays, individual plaque isolates were propagated on their corresponding sensitive lactococcal host to a high titre (at least 10^7 pfu/mL). The phages were initially characterized through multiplex PCR typing capable of identifying *Skunavirus*, *Ceduovirus*, and P335 phages, revealing that all 21 selected phages represent *Skunavirus* members.

Previous analysis of the diversity of RBP genotypes present in the phageome of these whey samples (White et al., 2025), indicated the presence of phages capable of infecting strains possessing a wide variety of *cwps* genotypes (such as B and D-types), although such strains had not been isolated from the starter cultures (White et al., 2024). Therefore, to isolate such phages, targeted phage detection was conducted using whey samples BA6 and BB4 against strains possessing B and D *cwps* types from our in-house strain collection. BA6 was used in screening against four different B-type lactococcal strains, allowing the isolation of a phage capable of infecting IL1403 (phage BA6-B). Two D-type strains, *L. lactis* 184 and 18-16, were used for the targeted detection of a D-type infecting phage leading to the isolation of a phage capable of infecting *L. lactis* 18-16 from whey BB4 (phage BB4-D).

The so-called **ph**age **R**BP-**a**ctivated **c**ell **s**orting (PhRACS) approach had previously been employed for the isolation of three low abundance (Casey et al., 2022) and novel phage-host combinations (White et al., 2025). The genomes of the associated three skunaviruses, i.e. BB4-SC34 (B-type), AO-4 (C_6), and BQ-1 (C_3), were sequenced in addition to the 23 skunaviruses isolated above yielding a total of 26 sequenced *Skunavirus* genomes.

3.2. Modular differences observed across Skunavirus genomes

The genomes of the 26 selected *Skunavirus* members ranged in length from 27.1 to 33.0 kbp (30.1 kbp on average) and possessed between 48 and 58 ORFs (53.2 on average) with a consistent genome architecture. The genome sequences were first compared to publicly available genomes (available in NCBI as of June 2024) through BLASTn analysis (using default parameters) to determine their relatedness/novelty (Table 1).

A phylogenetic tree was generated from the nucleotide sequences of 212 publicly available *Skunavirus* genomes (those indicated as a complete nucleotide sequence on submission), combined with the 26 novel *Skunavirus* genomes from this study (Fig. 2). With the inclusion of the dairy streptococcal *Moineauvirus* SW8 as an outgroup, five major branches of skunaviruses can be discerned (Fig. 2). The majority (14 of 26) of the skunaviruses isolated in this study cluster within one of these five major branches (which also includes previously described skunaviruses PhiG, PhiJF1, PhiF0139, and PhiF.17 isolated from dairy facilities in the Netherlands (Murphy et al., 2016, 2013)) with the remaining 12 genomes being spread among the remaining four branches, thus reflecting the genetic diversity of skunaviruses isolated from the various dairy fermentation facilities. In addition, the skunaviruses isolated from the whey samples appear to cluster independently of the geographical origin of the phages (Fig. 2; Supplementary Fig. S1). Overall, the 26 *Skunavirus* genomes represent a diverse collection of skunaviruses when compared to the nucleotide sequences of other publicly available *Skunavirus* genomes (Fig. 2).

To further investigate the genome content and diversity of skunaviruses, a *Skunavirus* core genome was determined using the *Skunavirus* genomes sequenced within this study and complete genomes available in the NCBI database. The core genome was determined to contain 28 core genes, present in all 242 *Skunavirus* genomes analysed (Supplementary Fig. S1; Supplementary Table S4). This resembles previous MCL analyses of *Skunavirus* genomes which reported a total of 29 core genes (Frantzen and Holo, 2019; Murphy et al., 2016). Slight differences in the reported core genomes may be attributable to the increase in genomes,

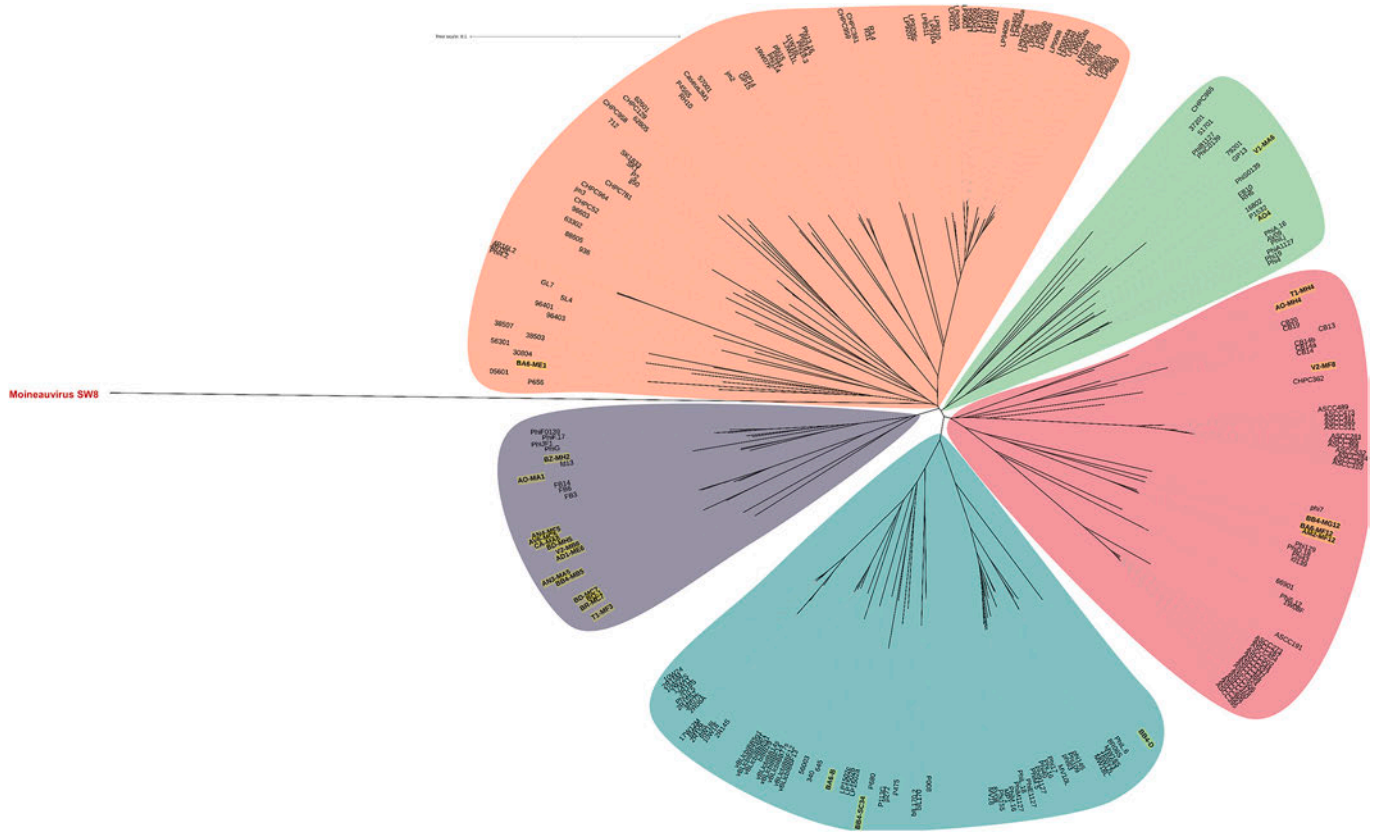


Fig. 2. Phylogenetic tree of 238 *Skunavirus* genomes. Phage phylogeny and maximum-likelihood tree was constructed based on multiple alignments of publicly available whole genome sequences and the 26 skunaviruses sequenced as part of this study (names are highlighted in yellow) using ClustalW and iTOL. The identified, visually apparent five major branches of skunaviruses are highlighted in distinct colours. The dairy streptococcal *Moineauvirus* SW8 was included as an outgroup.

shared assemble parameters in the current study, and/or the sensitivity of the current approach. A notable difference is the lack of the ssDNA binding protein the core genome of the current study, which is missing from a single phage, *Skunavirus* 56,301. However, it was determined that *Skunavirus* 56,301 does encode a ssDNA binding protein, albeit one that is distinct from the other skunaviruses (<50 % amino acid identity).

16 *Skunavirus* genomes, believed to be representative of the five genetic clusters mentioned above, were used to create a schematic overview of the organisation and sequence relatedness of the whey-isolated *Skunavirus* genomes (Fig. 3). The genetic elements involved in DNA packaging, host recognition and binding, virion morphogenesis, and lysis are well conserved among these *Skunavirus* genomes. However, there is considerable genetic diversity within the regions that specify the host binding machinery (RBPs and auxiliary binding proteins) in these phages. Certain core genetic elements (with known functions) found in the early- and middle-expressed regions such as the genes encoding single-stranded DNA binding (SSB) protein, recombinase Sak (sensitivity to AbiK) protein, and a RuvC-like Holliday endonuclease are highly conserved among the genomes, which is consistent with previous studies (Murphy et al., 2016). In addition, the majority of *Skunavirus* genomes (21 out of 26) analysed in this study encode one or more putative orphan methyltransferase(s) (MTases) within the replication region and six possess additional methyltransferase-encoding gene(s) within the DNA packaging module. A previous study found that MTases encoded within the packaging module in skunaviruses (Phi145 and Phi93) provided protection against host-encoded restriction endonucleases (Murphy et al., 2014). However, there is limited similarity between the six MTases encoded within the DNA packaging module identified in the analysed phage genomes and those in the previous study that were determined to provide protection against R-M systems of a particular specificity. Therefore, the REBASE webtool (http://rebase.neb.com/rebase/rebase.html) was used to predict the type and specificity of all putative MTases (Supplementary Table S5). The closely related skunaviruses BD-MC7, BR-MC7, and BQ-1 all share two genes encoding identical MTases within the DNA packaging module and neither shared substantial identity (<32 %) with any known MTase in the database (Supplementary Table S5). The three remaining skunaviruses (AO-MA1, T1-MF3, and BB4-MG12) each encode one distinct MTase (sharing <67.87 % amino acid identity to MTases encoded in the skunaviruses sequenced within this study) in their DNA packaging module. Phages AO-MA1 and BB4-MG12 each encode an MTase (AO-MA1_02 and BB4-MG12_02; Supplementary Table S5) which shares significant sequence identity (97 %) with Type II MTases M.PspTAD18ORF83P (which recognizes the sequence CTSAG) and M.Lph93Dam (which recognizes the sequence GATC), respectively. The remaining MTase (T1-MF3_02) most closely matched to M.PspTAD18ORF83P as well, though with lower (69 %) sequence identity.

3.3. Identification and characterisation of Skunavirus carbohydrate binding modules

Skunavirus binding to the surface-exposed saccharidic host receptor is facilitated by the phage-encoded RBP, and this function is often supplemented by auxiliary binding proteins encompassing CBMs. These auxiliary CBMs may be associated with the distal tail protein (Dit), major tail protein (TpeX), and/or neck passage structure (NPS), and have been found to complement and aid in host cell binding (Hayes et al., 2019, 2018).

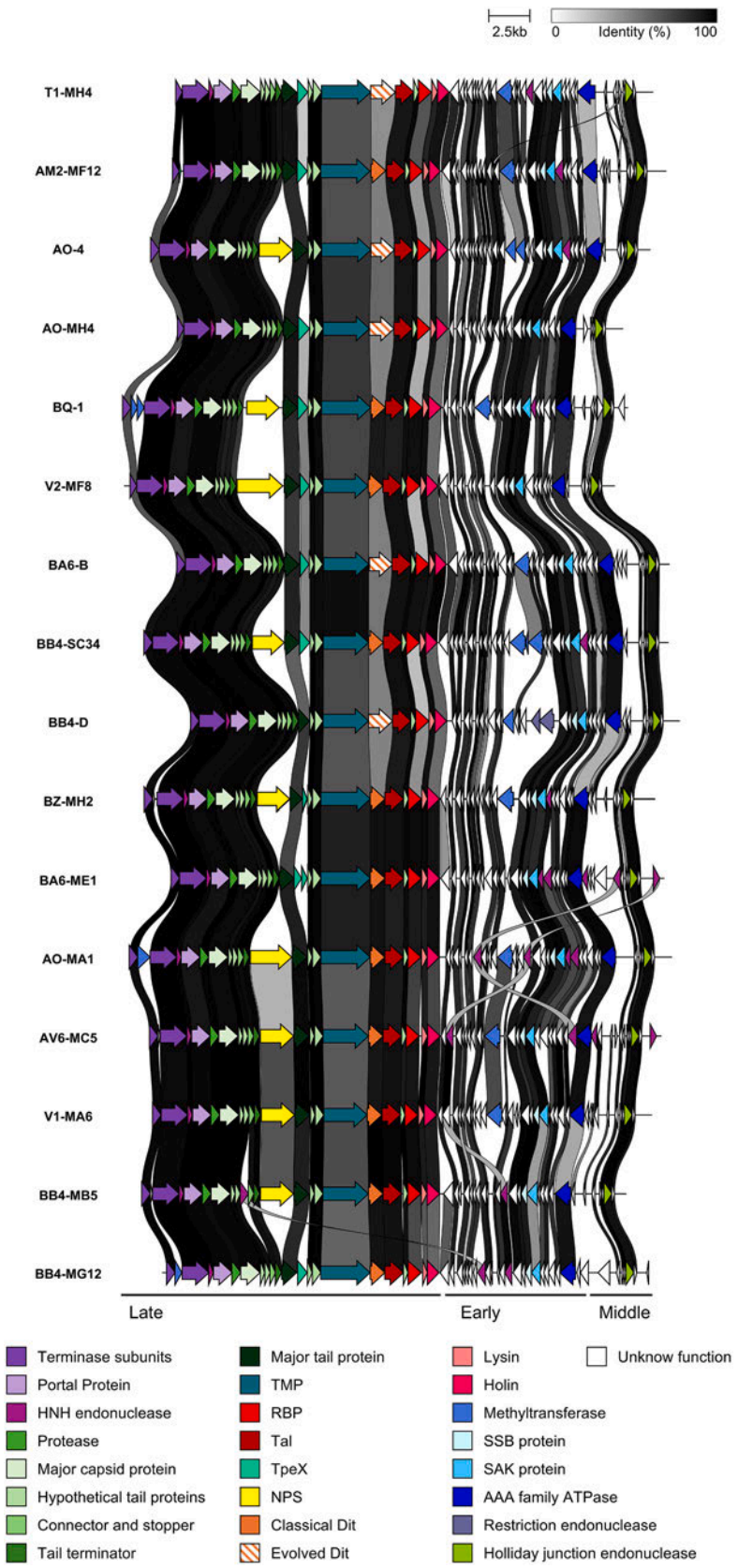


Fig. 3. Schematic representation of the genetic architecture of 16 newly sequenced skunaviruses. Arrows represent predicted genes which are colour-coded based on the presumed functions of the encoded proteins indicated in the panel at the bottom of the figure. Genome alignments were generated using Clinker (https://cagecat.bioinformatics.nl/tools/clinker).

3.4. Diversity of phage-encoded receptor binding proteins

Skunavirus RBPs are comprised of head (which incorporates a carbohydrate binding site that facilitates receptor binding), neck, and shoulder domains. These RBPs have been categorised into distinct phylogenetic groups, which are based on varying levels of amino acid sequence similarity within the C-terminal region of the RBP (which encompasses the head domain), and which correspond to the CWPS of the strain to which the phage can bind, thus being an important host range determinant (Mahony et al., 2013; Murphy et al., 2016; White et al., 2025). To categorise the *Skunavirus* RBPs identified in this study, the C-terminal ~110 residues of the newly identified RBPs were aligned against the head domains of all available *Skunavirus* RBP sequences (a total of 359 RBP sequences; Fig. 4). The majority of the newly isolated skunaviruses (18 of 26) encode group I RBPs (11 subgroup IA, four subgroup IB, and three subgroup IC RBPs). Additionally, two RBPs were found to belong to group III, four to group X, one to group VI, and one to group VIII. How the RBP group of a particular *Skunavirus* relates to its host range will be discussed further below.

3.4.1. Novel carbohydrate binding modules are present in certain distal tail proteins

The presence and diversity of auxiliary CBMs present in the phage genomes sequenced in this study were evaluated (Table 2). Lactococcal Dits can be classified as classical or evolved, where evolved Dits are larger and typically contain an additional CBM. Evolved Dit-associated CBMs have been found to possess the same host specificity as their corresponding RBP and were previously found to belong to one of four classes (Hayes et al., 2018). Of the 26 skunaviruses, five were found to possess evolved Dits (Table 2). Three phages (BA6-B, BB4-D, and AO-4) possess Dits that are similar to one of the previously described evolved Dit classes (Table 2; Supplementary Fig. S2). However, two phages (T1-MH4 and AO-MH4) both possess the same CBM insertion in their respective Dit proteins, though this CBM shares limited amino acid sequence identity (<70 %) with other lactococcal Dit proteins and we therefore propose this to represent a novel class (Class 5 Dit) of evolved Dit proteins (Supplementary Fig. S2). AlphaFold3 was used to predict the baseplate structure of T1-MH4 which includes the novel Class 5 evolved Dit (Fig. 5). Structure predictions of the T1-MH4 Dit revealed that it is most closely related to an uncharacterised protein encoded in the genome of a non-motile facultative anaerobic *Leptotrichia* spp. Transmission electron microscopy images of T1-MH4 did not yield any discernible feature associated with the distal tail region of this phage (Supplementary Fig. S3).

3.4.2. Structural diversity of Skunavirus neck passage structures

Many skunaviruses have been found to encode neck passage structures (NPS), which form a collar-whisker complex attached to the phage portal and which (may) contain CBM(s) that aid in host binding (Hayes et al., 2019; Vegge et al., 2006). Previous studies have indicated the

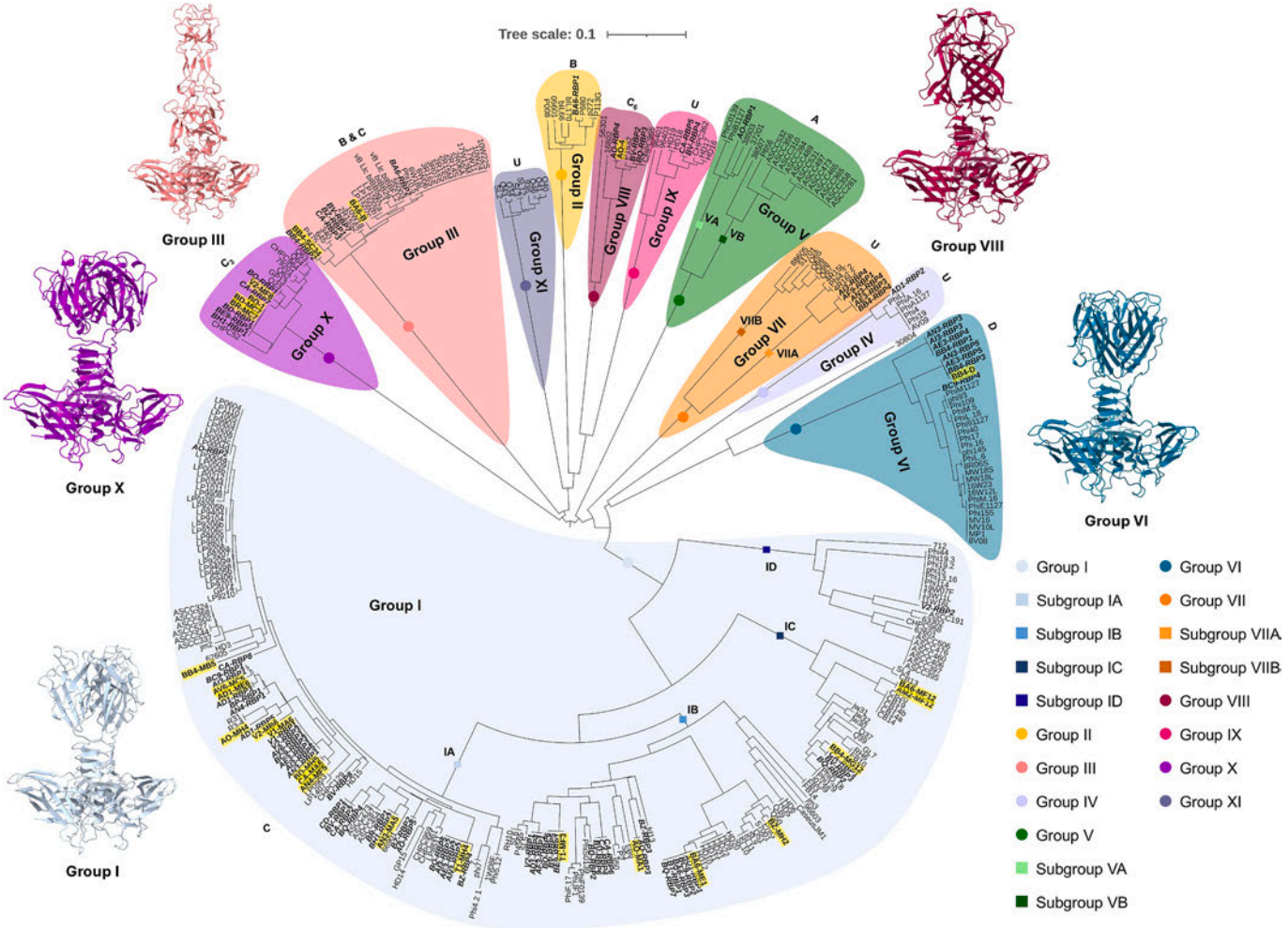


Fig. 4. Phylogenetic tree of the C-terminal domain of the 359 *Skunavirus* RBP sequences. Relevant RBP sequences were aligned and grouped together based on sequence similarity and correspond to host *cwps* (where possible). Structural representations of the RBP groups (I, III, VI, VIII, and X) the phageome-derived phage isolates belong to were generated using AlphaFold3.

Table 2
Types of auxiliary carbohydrate domains observed in the genomes of 26 sequenced skunaviruses.

Phage	Host CWPS Type	RBP Group	NPS		TpeX		Dit	
			Group	Size (aa)	Subgroup	Size (aa)	Type	Size (aa)
BA6-ME1	C_1	IB			***VI***	224	Classical	298
T1-MF3	C_1	IB	B	672			Classical	298
BZ-MH2	C_1	IB	A	656			Classical	298
AO-MA1	C_1	IB	***I***	804			Classical	298
V2-MB6	C_2	IA	A	676			Classical	298
AN3-MA5	C_2	IA	B	666			Classical	298
AN4-MF5	C_2	IA	A	676			Classical	298
V1-MA6	C_2	IA	B	672			Classical	298
T1-MH4	C_2	IA			***V***	208	***Class 5***	498
AV6-MC5	C_2	IA	B	676			Classical	298
BB4-MB5	C_2	IA	B	666			Classical	298
AD1-ME6	C_2	IA	A	676			Classical	298
BD-MH5	C_2	IA	A	676			Classical	298
AO-MH4	C_2	IA			***V***	227	***Class 5***	498
CA-MA6	C_2	IA	A	676			Classical	298
BB4-MG12	C_4	IC			I	193	Classical	298
AM2-MF12	C_4	IC			I	206	Classical	298
BA6-MF12	C_4	IC			I	206	Classical	298
BA6-B	B	III			I	211	Class 3	470
BB4-SC34	B	III	B	641	I	204	Classical	298
BB4-D	D	VI					Class 2	486
AO-4	C_6	VIII	B	672			Class 3	474
V2-MF8	C_3	X	***H***	924	I	187	Classical	298
BQ-1	C_3	X	***J***	658	I	175	Classical	298
BD-MC7	C_3	X	***J***	658	I	175	Classical	298
BR-MC7	C_3	X	***J***	658	I	175	Classical	298

Novel CBM groups, subgroups, and types identified in this study are in bold-face text and italicised.

presence of seven (group A-G) NPS groups corresponding to distinct neck passage structures and/or limited amino acid sequence identity (Hayes et al., 2019). 13 (of 18 possessing an NPS) skunaviruses were found to possess Group A or B NPS structures (Table 2). The NPSs of the remaining five phages share limited amino acid identity with other *Skunavirus* NPS proteins and were, therefore, considered novel NPSs. These novel NPS groups will be referred to as Groups H, I, and J (Supplementary Table S6). The structural predictions (generated using AlphaFold3) of these novel groups, as well as representatives of groups A and B, are shown in Fig. 6 A-E. Structural prediction of *Skunavirus* V2-MF8 novel NPS yielded a unique structure which was found to encode three distinct CBMs at the C-terminal end: a five-blade β-propeller, trimeric TP901-1 RBP-like domain, and Tuc2009 BppA (Fig. 6E). Notably, the β-propeller structurally resembles the Glutaminyl Cyclase from *Xanthomonas campestris* (based on Dali server returns). *Skunavirus* V2-MF8 was therefore assessed by electron microscopy for its morphological characteristics (Fig. 6F). Collar structures and whiskers with small, bobble-like structures were clearly visible resembling the whisker-like appendages with small terminal knots observed in certain skunaviruses (such as Phi15 and PhiG, for example (Murphy et al., 2013)). The observed whiskers were 23.04 ± 4.44 nm ($n = 19$) in length with the associated whisker balls being about 9.474 ± 1.188 nm ($n = 17$) wide (Supplementary Table S7). The Type J NPS of *Skunavirus* BQ-1 contains two distinct domains, one resembling the structure of *Streptococcus pneumoniae* peptidoglycan deacetylase and the other a *S. pneumoniae* fucose-binding module (based on Dali server returns; Fig. 6C). The C-terminal domain of the Type I NPS of *Skunavirus* AO-MA1 contains a unique domain most closely structurally resembling the beta-1,3-glucanase from *Phanerochaete chysosporium* gluconolactone complex (based on Dali server returns; Fig. 6D).

3.4.3. Sequence diversity of the major tail extension protein

Finally, 12 skunaviruses encode a major tail extension (TpeX) protein which results in the presence of additional decorations along the tail, and which have also been found to exhibit the same host specificity as their corresponding RBP, with more favourable binding at the cell poles where division occurs (but with significantly reduced affinity) (Hayes et al., 2019; Murphy et al., 2016). There are currently four defined TpeX subgroups that all correspond to the Tuc2009 BppA CBM, but vary in amino acid composition and/or length (Hayes et al., 2019). 12 of the skunaviruses characterized in this study encode a TpeX, of which the majority (9/12) represent members of subgroup I. The remaining three TpeX proteins, incorporating a Tuc2009 BppA domain, form two novel subgroups. Two TpeX proteins (encoded by phages T1-MH4 and AO-MH4) cluster closely together and form a novel subgroup V, while the remaining TpeX (from phage BA6-ME1) forms subgroup VI (relation to previous TpeX subgroups is summarised in Supplementary Table S8).

3.5. Linking Skunavirus RBPs to host range

To further understand the link between the host range of skunaviruses and their corresponding RBP, host range analysis of the 26 isolated skunaviruses was undertaken against a panel of 45 lactococcal strains (26 coming from the original 92 isolate screening panel, the remaining from in-house strain collections; strains listed in Supplementary Table S1) possessing various *cwps* genotypes (A, B, C_1-C_6, and D-types) and correlated to their defined RBP groups (Fig. 7; Supplementary Table S9). A strict correlation was observed between the *Skunavirus* RBP (sub)group and the *cwps*-type of the lactococcal strain they infect, being consistent with previous studies (Hayes et al., 2019; Mahony et al., 2013; White et al., 2025). For example, all skunaviruses with group I RBPs were found to infect C-type strains; with subgroup IA infecting C_2-type strains, subgroup IB mainly infecting C_1-type strains, and IC infecting C_4-type strains (White et al., 2025). However, a *Skunavirus* with a subgroup IB RBP was also capable of infecting three C_2-type strains (Fig. 7). Therefore, this host range analysis supports and supplements the most recent study by White et al. on RBP specificity as determined through fluorescent binding assays, where the binding specificity of subgroup IB RBPs could not be elucidated through fluorescent binding assays (White et al., 2025).

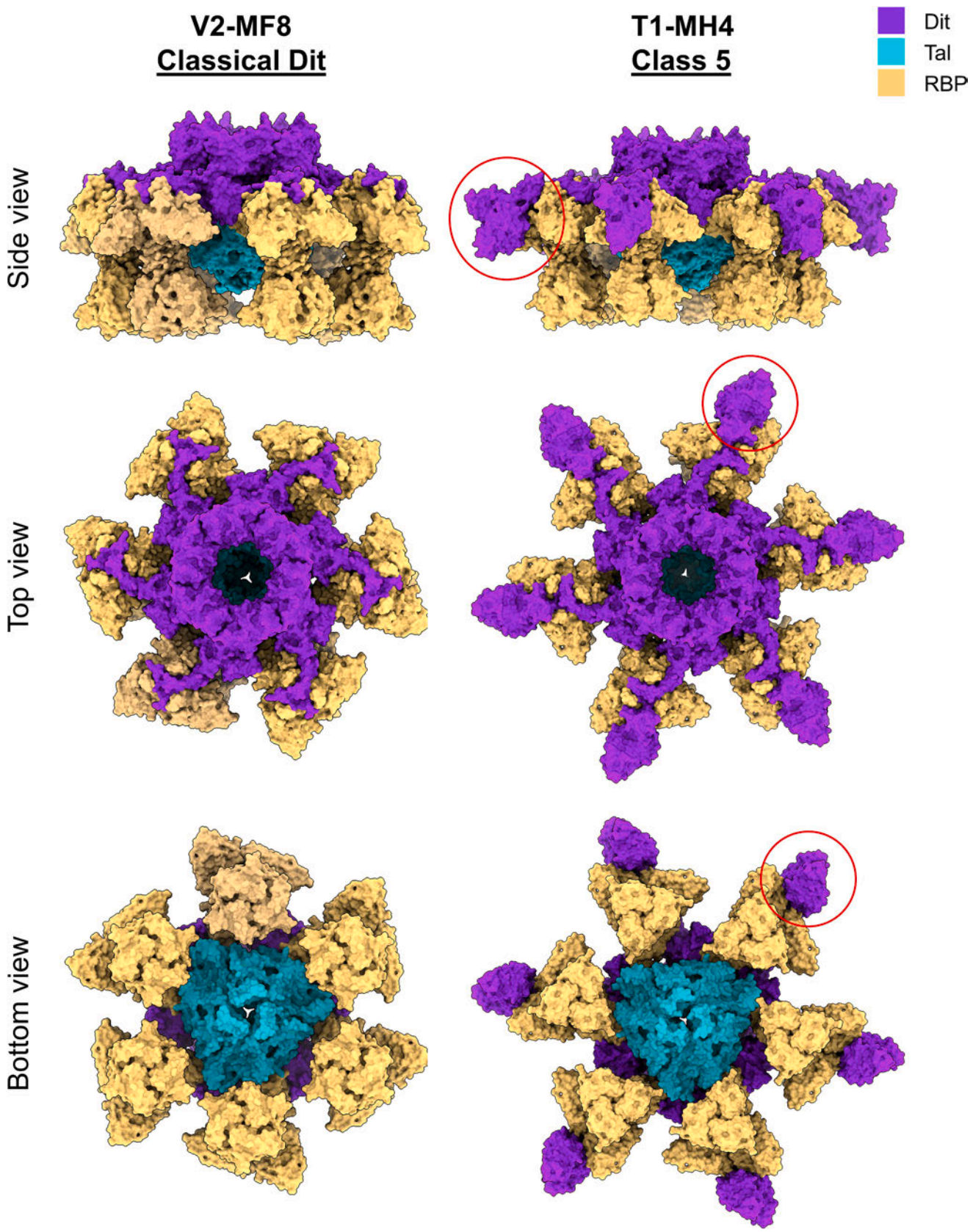


Fig. 5. Predicted structures of the Dit-Tal-RBP assemblies associated with the baseplate of two whey-derived skunaviruses (V2-MF8 and T1-MH4) as predicted with AlphaFold3. The baseplates of *Skunavirus* V2-MF8 and T1-MF8 corresponding to a classical Dit and a novel Class 5 evolved Dit. The Dit is coloured purple, the Tal blue, and the RBP yellow. The CBM of the evolved Dit of T1-MF8 is circled in red.

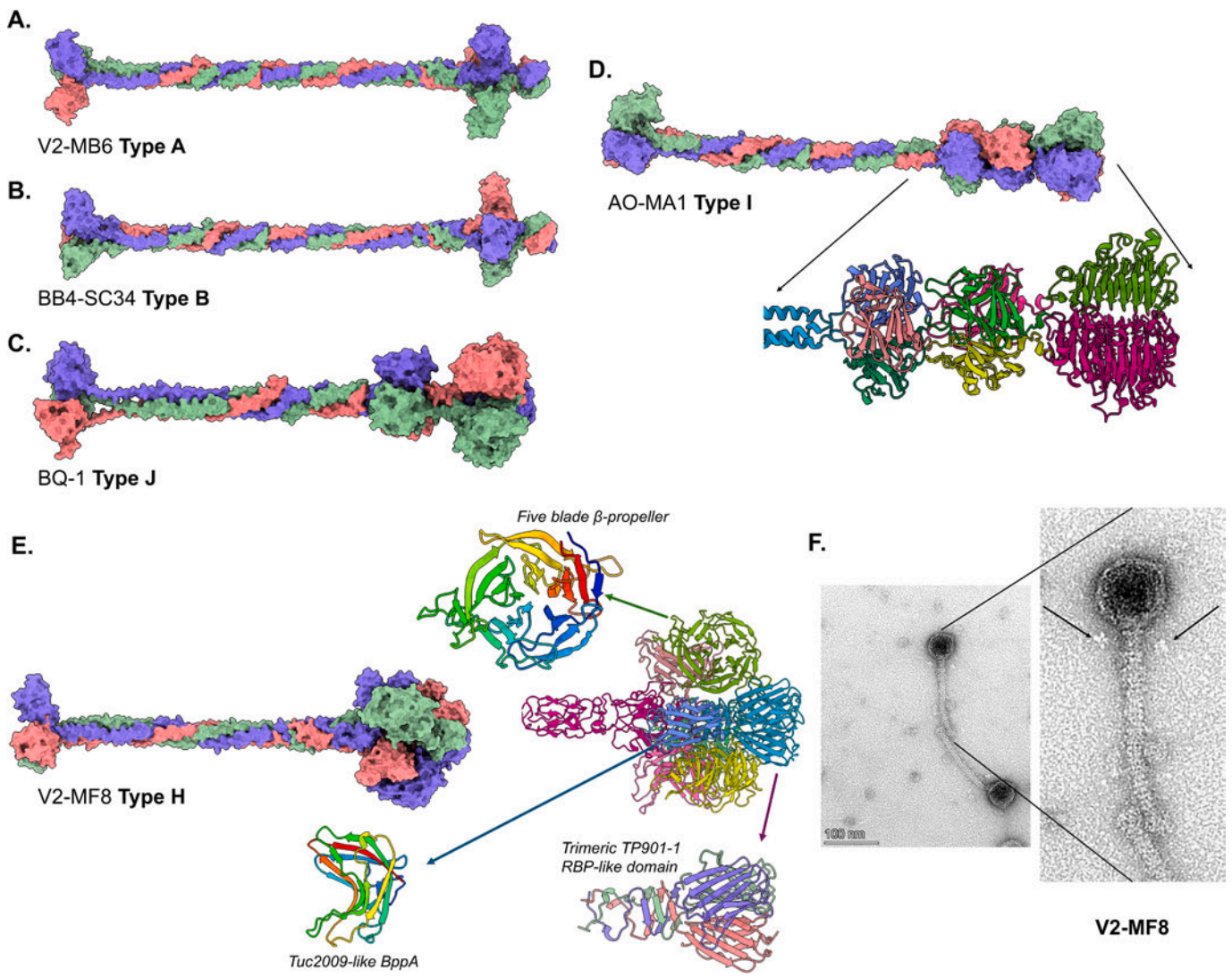


Fig. 6. Predicted neck passage structure proteins of isolated whey-derived skunaviruses as determined with AlphaFold3. Structural predictions of representative NPS proteins encoded by sequenced skunaviruses: (A) Type A (V2-MB6), (B) Type B (BB4-SC34), (C) Type J (BQ-1), (D) Type I (AO-MA1), and (E) Type H (V2-MF8). (E) A close-up view of three distinct CBMs found within the Type H NPS. (F) Micrograph of *Skunavirus* V2-MF8 (Type J) with arrows marking the whiskers with small bobble-like structures that alongside collar structures make up the neck passage structure of the phage.

3.6. Combinatorial approach to investigate phage diversity reveals additional phage diversity

Initial analysis of phage diversity present in these fermentation facilities was conducted through virome sequencing and subsequent phageome analysis of 27 (of the 32) whey samples focused on investigating the diversity and binding specificity of *Skunavirus* RBPs present in the phageomes (White et al., 2025). Here, to complement the traditional approach of phage isolation and sequencing, complete (or nearly complete, i.e. >25.0 kbp) *Skunavirus* genomes from the whey phageome assemblies were identified (Table 3) and aligned against the sequenced *Skunavirus* genomes. From the analysed 27 whey phageome assemblies, only 11 distinct (within a sample) complete or nearly complete *Skunavirus* genomes were identified. These contigs ranged in length from 25.7 to 30.8 kbp and were identified in the phageome obtained from 11 samples (AD1, BP, V2, AE3, AI3, AN3, AV6, BD, BA6, BB4, and BZ). To determine if any of these retrieved contigs closely resembles any of the isolated/sequenced phages, Hadamard scores (product of average nucleotide identity (ANI) and coverage) were generated for each isolated/sequenced *Skunavirus* against the 11 phageome-derived *Skunavirus* contigs (Supplementary Fig. S4). In addition, a phylogenetic tree of the genomes of the microbiologically isolated skunaviruses as well as phageome-derived *Skunavirus* contigs was generated (Supplementary Fig. S4). Validated by manual inspection, it was determined that two phageome-derived *Skunavirus* contigs, Contig_AE3-1 and Contig_AI3-1, are essentially identical to the isolated (and subsequently sequenced) skunaviruses V1-MA6 (Hadamard score 1.00) and CA-MA6 (Hadamard score 0.98), respectively. Additionally, two phageome-derived *Skunavirus* contigs, Contig_AN3-1 and Contig_BD-1, are nearly identical to the genomes of skunaviruses V1-MA6 (Hadamard score 0.90) and BD-MC7 (Hadamard score 0.84), respectively. The virome contigs correspond to identical regions of the respective phages, with the complete assemblies of the two phageome-derived contigs the likely cause of Hadamard scores below 1 (due to reduced query coverage). The seven remaining phageome-derived *Skunavirus* contigs therefore correspond to skunaviruses that were not isolated within this study by means of cultivation-based techniques. This therefore demonstrates that phage diversity was not fully covered by the traditional (i.e. cultivation-based) isolation and sequencing method described herein. Moreover, the identification of 11 near-complete phageome-derived *Skunavirus* contigs in 27 whey phageomes reveals the limitations of a “metagenomics-only” approach to investigate phage diversity. This may be due to a need to increase

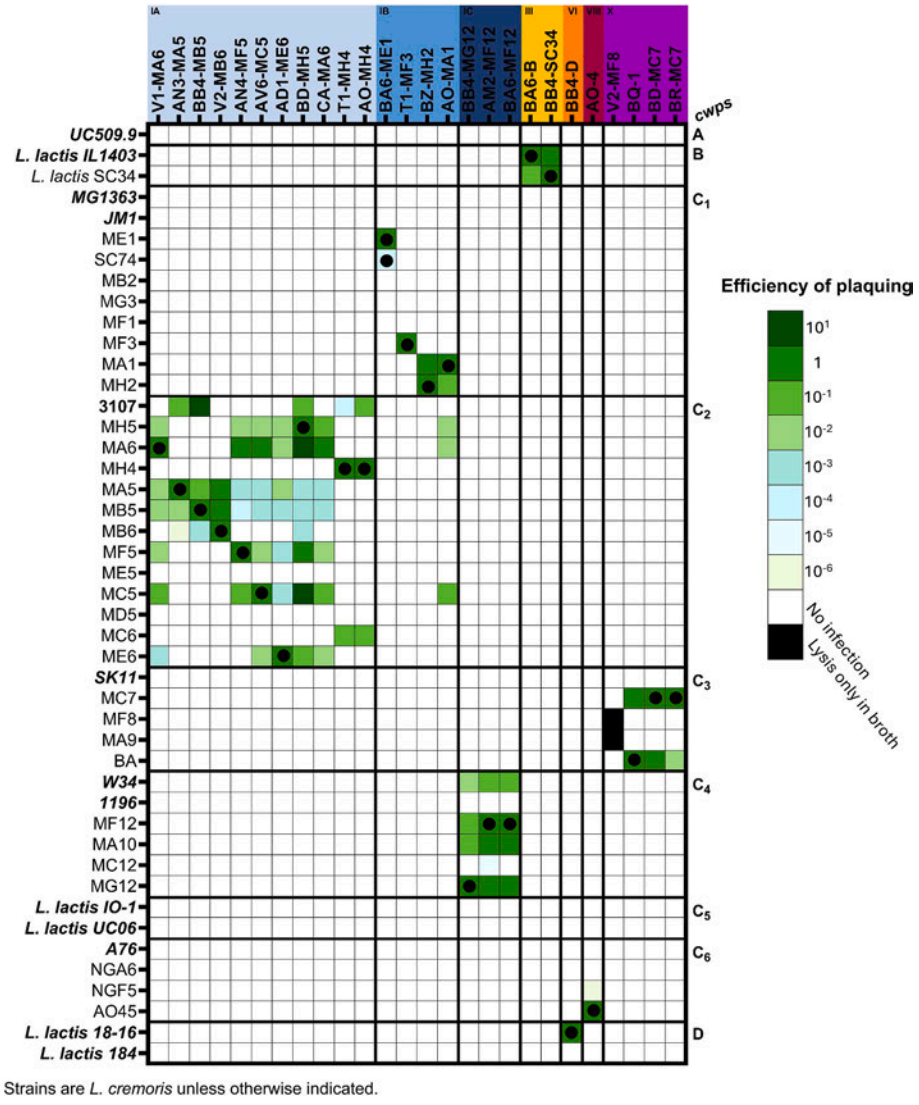


Fig. 7. Host range of the 26 skunavirus isolated within this study. Host range was determined by spot testing against a panel of 45 lactococcal strains (strains belong to the species *L. cremoris*, unless otherwise indicated in the figure). Skunavirus names are coloured according to their RBP group. Squares are colour-coded based on the calculated EOP of a phage against a particular strain (indicated in the figure legend) relative to its propagating host (indicated by a black circle within a square). Detailed EOP information summarised in this figure is presented in Supplementary Table S9. White squares indicate no infection. Black squares indicate infection by a phage determined by lysis in broth that does not produce visible plaques.

Table 3
Characteristics of eleven phageome-derived contigs.

Phageome-derived contig	Year of collection (whey)	Genome size (bp)	GC%	Best Hit (% coverage,% nt ID)	No. ORFs				GenBank Accession #
					Late	Early	Middle	Total	
Contig_AD1-5	2019	28,321	35.01	vB_Llc_bIBB24tp1 (76, 94.17)	21	26	4	51	PV941775
Contig_AE3-1	2020	30,747	35.12	16,802 (79, 93.18)	21	31	4	56	PV941776
Contig_AI3-1	2020	30,315	34.86	PhiG (83, 95.04)	22	27	6	55	PV941777
Contig_AN3-1	2020	27,314	35.18	FB6 (87, 89.70)	18	30	4	52	PV941778
Contig_AV6-1	2020	26,492	35.12	2R06A (72, 94.07)	20	27	4	51	PV941779
Contig_BA6-1	2020	29,666	35.25	p272 (84, 92.99)	21	29	5	55	PV941780
Contig_BB4-1	2020	30,105	35.00	RM5 (79, 90.83)	22	25	5	52	PV941781
Contig_BD-1	2021	25,726	35.34	CHPC964 (76, 91.49)	15	25	3	43	PV941782
Contig_BP-1	2021	30,841	34.75	Phi4 (76, 94.83)	23	29	4	56	PV941783
Contig_BZ-1	2022	29,052	35.31	vB_Llc_bIBBF12 (98, 96.24)	21	30	4	55	PV941784
Contig_V2-1	2019	30,670	35.01	GP13 (81, 93.75)	22	30	4	56	PV941785

sequencing depth (an average sequencing output of 3.7 million high quality reads was generated per sample Supplementary Table S2), or due to difficulties in accurately assembling closely related and mosaic *Skunavirus* genomes.

To garner further insights into the phage diversity present in the 27 whey samples for which virome data was available, the isolated/sequence and phageome-derived *Skunavirus* genomes/contigs were mapped against the raw virome reads to determine which skunaviruses

(or very closely related ones) may be present in which dairy fermentation facilities (Fig. 8). Certain skunaviruses such as V1-MA6 and CA-MA6, for example, appear (with >99 % covered bases) to be persistent in the Danish dairy facilities being present in eight and four different whey viromes, respectively, collected between 2019 and 2021. Other skunaviruses, such as T1-MH4 and BB4-MB5 (or very closely related phages), may be associated with the undefined starter cultures as they are present in several whey samples across (five and four of the six countries, respectively) geographical locations. In certain cases, phages (for example BA6-MF12 and AN3-MA5) that were isolated from a particular whey sample during the phage screening process outlined in this study were not identified in the same corresponding whey viromes (but were detected in other virome samples). In addition, two of the isolated phages (BB4-MG12 and BZ-MH2) were not identified in any whey virome. This may be due to the phages being in lower abundance or due to limitations associated with virome extractions and sequencing.

4. Discussion

Phage infection remains one of the main challenges in dairy fermentation systems, typically resulting in slowed fermentations (Samson and Moineau, 2013). In particular, skunaviruses remain the most endemic and problematic lactococcal phage group in facilities utilising mesophilic starter cultures (mainly composed of *L. lactis* and *L. cremoris* strains) (Mahony et al., 2012). Several decades of research has underpinned the current depth of understanding of this group of virulent lactococcal phages in relation to their genetic, morphological, and molecular characteristics (Jolicoeur et al., 2023; Miklič and Rogelj, 2003; Moineau et al., 1996, 1992; Murphy et al., 2013; Oliveira et al., 2018; Szczepańska et al., 2007). Consistent screening and monitoring of phage populations associated with starter cultures and/or fermentation facilities is essential to minimise the impact and assess the risk of phage infection of starter cultures. Phage screening has traditionally been performed through culture-dependent methods (Bissonnette et al., 2000; Chmielewska-Jeznach et al., 2018; Jolicoeur et al., 2023; Murphy et al., 2013; Oliveira et al., 2018), and although this is a widely used

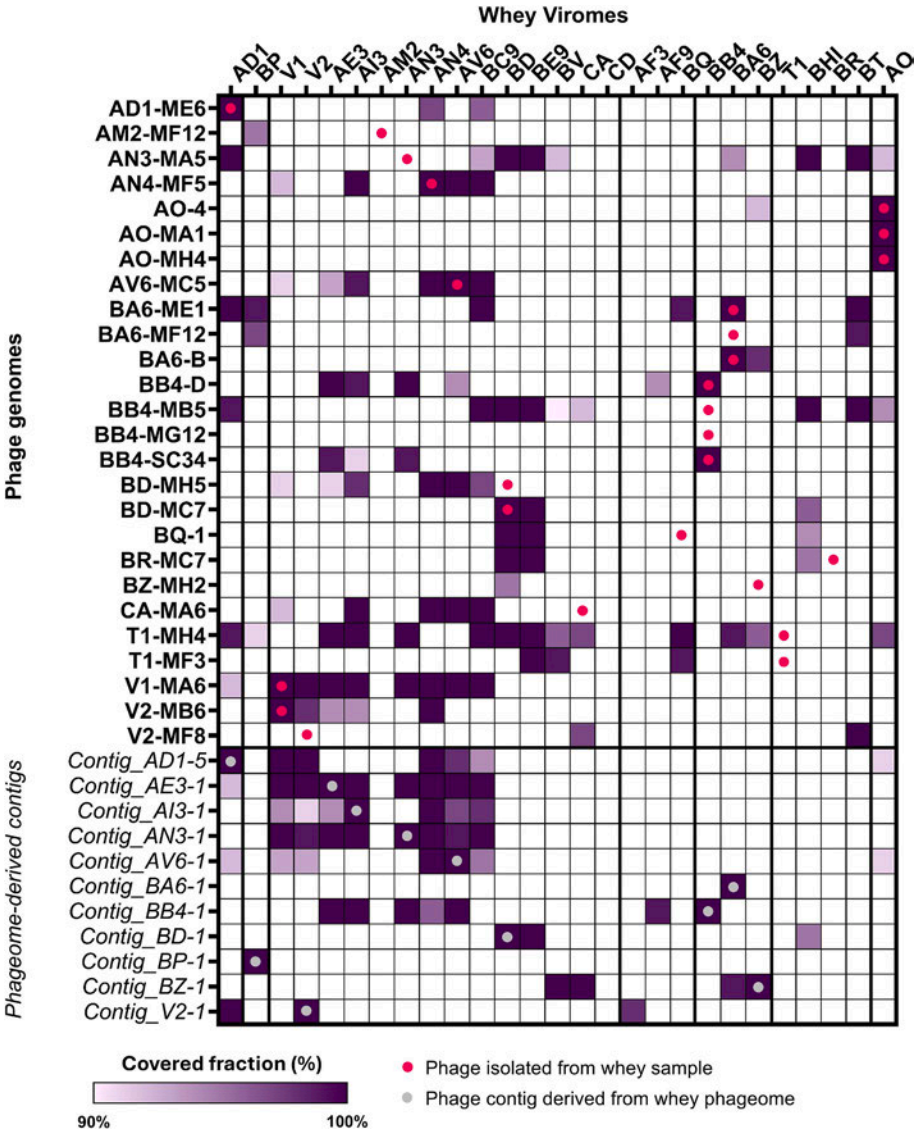


Fig. 8. Presence or absence of a *Skunavirus* genome/contig in the virome reads of 27 previously sequenced whey viromes. The covered fraction of a particular (or closely related) genome/contig was determined through reads mapping of phageome-derived *Skunavirus* contigs and isolated/sequenced *Skunavirus* against raw reads of 27 whey sample viromes. The minimum covered fraction depicted in the figure is 90 %, squares are coloured according to the covered fraction (between 90-100 %).

approach, culture-dependent methods of phage detection only capture the dominant and propagatable phages present in a sample. Metagenomics, and more recently viromics/phageomics, have been utilised to determine phage prevalence and diversity in a number of ecological niches (Colombo et al., 2018; Dugat-Bony et al., 2020; Paillet et al., 2024).

The development of more targeted methods to retrieve viral particles and associated nucleic acids have substantially improved phageome analysis, particularly in dairy settings (Muhammed et al., 2017). Furthermore, the assembly of virome data into high-quality phage genomes is highly challenging due to the modular and mosaic nature of many phage genomes (Roux et al., 2017; Sutton et al., 2019). Viral metagenome assembly tools, such as Phables (Mallawaarachchi et al., 2023), have significantly improved scientific endeavours to uncover viral diversity based on sequence information. In the current study, phage biodiversity in whey samples obtained from several dairy fermentation facilities across Europe employing mesophilic undefined starter cultures was assessed. Twenty-six novel and distinct skunaviruses were identified through traditional culture-dependent phage screening methods, targeted isolation based on predicted RBP specificity, and PhRACS. Based on the RBP diversity uncovered in these whey samples through a previously reported phageome analysis (White et al., 2025), culture-independent analysis was used to guide or facilitate culture-based isolation of phage-host combinations, indicating that a combinational approach to phage screening will yield a more comprehensive view of phage diversity present in a sample. However, as the binding specificity of several RBP groups remains undefined (White et al., 2025), it was not possible to retrieve the complete complement of phage-host combinations present in the dairy fermentation facilities investigated within this study. Nevertheless, as more lactococcal strains are isolated and their genomes sequenced, novel *cwps* gene clusters (and associated structures) will likely be identified. Additionally deeper metagenomic sequencing may allow identification of low abundance lactococcal genomes harbouring rare *cwps* clusters within complex microbial communities (White et al., 2024).

Phylogenetic analysis revealed substantial genomic diversity between the skunaviruses isolated within the study, with the 26 genomes represented in all five major phylogenetic branches of skunaviruses (Fig. 2, Supplementary Fig. S1). Additionally, it is important to note that although the isolated skunaviruses were shown to cluster independently of the geographic location of their corresponding fermentation facility, 14 of these skunaviruses clustered within one of the five major branches. The presence of closely related phages in distinct factories may be the result of similar strains being present in the undefined mesophilic starter cultures used in disparate locations. Alternatively, it is also possible that phages are associated with starter cultures employed within the dairy fermentation facilities, resulting in the isolation of distinct (but similar) phages in disparate locations yet originating from a starter-derived common ancestor. Host range analysis of these phages was confirmed and expanded upon the recently established predicted binding specificity of their associated RBP, by identifying C_1 CWPS as the primary receptor target of RBP subgroup IB (White et al., 2025). In addition, a number of novel CBMs were identified in the NPS, evolved Dit, and major tail extension proteins of several phages. Of note, two skunaviruses were found to carry *nps* genes encoding proteins which assemble as collar and whiskers with small, bobble-like structures similar to those observed in other skunaviruses (Fig. 5 A and B) (Murphy et al., 2013). These collar structures associated with the NPS also resemble the KIS (LL-K insertion sequence) elements previously identified in *Lactobacillus delbrueckii* phages (Casey et al., 2014). A notable difference is that in the case of *Lb. delbrueckii* phages, the apparent NPS is encoded across two open reading frames (Forsman and Alatossava, 1994), whereas they are encoded by a single gene (*nps*) for the lactococcal phages (McGrath et al., 2006). The observed generalist binding capabilities of the NPS of skunaviruses (Hayes et al., 2019) augments the likelihood of a phage being in proximity to a suitable host thereby playing an auxiliary role in phage infection. The diversity of identified *Skunavirus* RBPs continues to increase (White et al., 2025), possibly reflecting the continuous adaptation and evolution of these phage components in order to keep up with diversifying host receptors (Hayes et al., 2019; Mahony et al., 2013; Murphy et al., 2016). In the present study, genetic and subsequent structural diversity was observed in the CBMs of auxiliary binding proteins including Dit and TpeX. In accordance with previous studies, these genome regions (tail morphogenesis, RBP, and NPS) are considered diversity hotspots (Castro-Nallar et al., 2012). *Skunavirus* RBPs represent the primary determinant of a phage's host range, with additional CBMs presumed to contribute to the host-binding process (Hayes et al., 2019, 2018). The auxiliary binding protein TpeX is associated with the presence of additional decorations along the phage tail while NPS proteins form a collar-whisker complex attached at the head-tail interface of the phage (Hayes et al., 2019). CBMs present in both the NPS and TpeX have previously been found to bind favourably towards the polar ends of a host cell (where cell division occurs) maintaining the same host specificity as their corresponding RBP (although with reduced affinity) (Hayes et al., 2019). However, in the current study, skunaviruses encoding subgroup IC, group III and group X RBPs, were found to encode the same subgroup I TpeX (Table 2). In addition, with the emergence of novel CBMs being present in the TpeX, NPS, and Dit it may be necessary to re-examine whether these CBMs always maintain the same host specificity as their RBP or if they may aid in expanding a phage's host range. Interestingly, it has recently been shown that evolved Dits in skunaviruses aid in adsorption to exopolysaccharide (EPS) receptors (Millen et al., 2023).

To complement a culture-based phage isolation and characterisation as described here, 11 complete (or nearly complete) phageome-derived *Skunavirus* contigs generated from previously described phageome assemblies (White et al., 2025) were compared to the 26 isolated/sequenced *Skunavirus* genomes. The depth of sequencing used in the current study (average 3.7 M quality filtered reads per sample) is consistent with other fermentation viromes studied recently (Dugat-Bony et al., 2020; Muhammed et al., 2017). However, deeper virome sequencing, perhaps complemented by long read sequencing, is likely to improve the accuracy and completeness of assembled phage genomes from virome data. Additionally, a major limitation of phageome sequencing is the difficulty in assembling genomes of closely related phages within a population. Nonetheless, our analysis highlights both the limitations and benefits of applying culture-dependent or -independent analysis methods in isolation. Culture-based analysis of phage diversity is biased towards the strains used during screening and isolation. Whereas culture-independent analysis does not allow for further biological investigation of phages (i.e. host range analysis) and is constrained by sequencing depth and available assembly pipelines (Mallawaarachchi et al., 2023; Roux et al., 2017; Sutton et al., 2019).

Overall, there is a significant evolutionary benefit to the apparent elasticity of *Skunavirus* genomes; with genetic acquisitions, alterations, and modifications (particularly within their tail morphogenesis module) ensuring their continued success (Castro-Nallar et al., 2012). The intertwined evolutionary relationship between skunaviruses and their cognate hosts also drives the ever-expanding diversity observed within lactococcal *cwps* gene clusters (Mahony et al., 2020; Parlindungan et al., 2024). Therefore, accurate and continuous monitoring of phage populations within dairy fermentation systems is likely one of the most important practices necessary to define both the dominant, low abundance and emerging phages present in a sample. Additionally, the isolation of lactococcal strains possessing diverse *cwps* clusters will also be essential to circumvent phage attacks, particularly in fermentation facilities employing defined starter cultures. Dissecting the genetic and functional diversity of lactococcal phages and their hosts is crucial to improve phage-related risk assessment and consequent predictive and rational planning capabilities in dairy fermentations.

5. Conclusion

By investigating the diversity of lactococcal phages present in whey samples originating from various fermentation facilities employing undefined mesophilic starter cultures, 26 *Skunavirus* were isolated and sequenced. An extensive host range analysis employing strains corresponding to 11 different cell wall polysaccharide structures further resolved the specificity of a *Skunavirus* RBP to a corresponding host *cwps* genotype. We conclude that the combination of phageome analysis and conventional culture-based approaches is likely the most effective route to developing a holistic view of the phage diversity in fermentation samples or indeed of phages present in a variety of ecological niches. Furthermore, it is likely that the benefits of a combinational approach will increase as nucleic acid extraction, sequencing (depth), and assembly methods of virome/phageome data continue to improve. By establishing a deeper understanding of both current and potentially problematic phage populations present in dairy fermentation facilities, more informed, robust, and efficient culture rotation strategies may be developed.

CRediT authorship contribution statement

Kelsey White: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Giovanni Eraclio:** Writing – review & editing, Resources. **Gabriele Andrea Lugli:** Writing – review & editing, Methodology, Investigation, Formal analysis, Data curation. **Marco Ventura:** Writing – review & editing, Supervision, Software, Resources. **Natalia Biere:** Writing – review & editing, Visualization, Resources, Data curation. **Frank Hille:** Writing – review & editing, Visualization, Resources, Data curation. **Federica Volonté:** Writing – review & editing, Resources. **Charles M.A.P. Franz:** Writing – review & editing, Visualization, Supervision, Formal analysis. **Christian Cambillau:** Writing – review & editing, Visualization, Formal analysis. **Fabio Dal Bello:** Writing – review & editing, Funding acquisition. **Jennifer Mahony:** Writing – review & editing, Validation, Supervision, Funding acquisition, Conceptualization. **Douwe van Sinderen:** Writing – review & editing, Validation, Supervision, Resources, Funding acquisition, Conceptualization.

Funding

This publication has emanated from research conducted with the financial support of Science Foundation Ireland under Grant numbers 20/FFP-P/8664, SFI/12/RC/2273-P1 and SFI/12/RC/2273-P2. K.W. is supported by the Irish Research Council (https://research.ie/) Enterprise Partnership Scheme (No. EPSPG/2020/12) in partnership with Sacco Srl.

Declaration of competing interest

Fabio dal Bello, Giovanni Eraclio and Federica Volonté are employees of Sacco Srl.

Acknowledgments

We acknowledge UCSF ChimeraX for the molecular graphics developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from the National Institute of Health R01-GM129325 and the Office of Cyber Infrastructure and Computational Biology, and the National Institute of Allergy and Infectious Diseases.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijfoodmicro.2025.111509.

Data availability

The genome sequences of the phages sequenced within this study and phageome-derived contigs have been deposited in the GenBank database (https://www.ncbi.nlm.nih.gov/genbank/) under the accession numbers PV941749-PV941785 (individual accession numbers for each genome can be found in Tables 1 and 3).

References

Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J., Bodenstein, S.W., Evans, D.A., Hung, C.-C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Žemgulytė, A., Arvaniti, E., Beattie, C., Bertolli, O., Bridgland, A., Cherepanov, A., Congreve, M., Cowen-Rivers, A.I., Cowie, A., Figurnov, M., Fuchs, F.B., Gladman, H., Jain, R., Khan, Y.A., Low, C.M.R., Perlin, K., Potapenko, A., Savy, P., Singh, S., Stecula, A., Thillaisundaram, A., Tong, C., Yakneen, S., Zhong, E.D., Zielinski, M., Žídek, A., Bapst, V., Kohli, P., Jaderberg, M., Hassabis, D., Jumper, J.M., 2024. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature 630, 493–500. https://doi.org/10.1038/s41586-024-07487-w.

Ainsworth, S., Sadovskaya, I., Vinogradov, E., Courtin, P., Guerardel, Y., Mahony, J., Grard, T., Cambillau, C., Chapot-Chartier, M.-P., van Sinderen, D., 2014. Differences in lactococcal cell wall polysaccharide structure are major determining factors in bacteriophage sensitivity. mBio 5, e00880-14. https://doi.org/10.1128/mBio.00880-14.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V. M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19, 455–477. https://doi.org/10.1089/cmb.2012.0021.

Bebeacua, C., Tremblay, D., Farenc, C., Chapot-Chartier, M.-P., Sadovskaya, I., van Heel, M., Veesler, D., Moineau, S., Cambillau, C., 2013. Structure, adsorption to host, and infection mechanism of virulent lactococcal phage p2. J. Virol. 87, 12302–12312. https://doi.org/10.1128/jvi.02033-13.

Bidnenko, E., Ehrlich, S.D., Chopin, M.-C., 1998. *Lactococcus lactis* phage operon coding for an endonuclease homologous to RuvC. Mol. Microbiol. 28, 823–834. https://doi.org/10.1046/j.1365-2958.1998.00845.x.

Bissonnette, F., Labrie, S., Deveau, H., Lamoureux, M., Moineau, S., 2000. Characterization of mesophilic mixed starter cultures used for the manufacture of aged cheddar cheese. J. Dairy Sci. 83, 620–627. https://doi.org/10.3168/jds.S0022-0302(00)74921-6.

Bouchard, J.D., Moineau, S., 2004. Lactococcal Phage Genes Involved in Sensitivity to AbiK and their relation to single-strand annealing proteins. J. Bacteriol. 186, 3649–3652. https://doi.org/10.1128/jb.186.11.3649-3652.2004.

Casey, C.N., Morgan, E., Daly, C., Fitzgerald, G.F., 1993. Characterization and classification of virulent lactococcal bacteriophages isolated from a Cheddar cheese plant. J. Appl. Bacteriol. 74, 268–275. https://doi.org/10.1111/j.1365-2672.1993.tb03025.x.

Bouras, G., Nepal, R., Houtak, G., Psaltis, A.J., Wormald, P.-J., Vreugde, S., 2023. Pharokka: a fast scalable bacteriophage annotation tool. Bioinformatics 39, btac776. https://doi.org/10.1093/bioinformatics/btac776.

Casey, E., Mahony, J., O'Connell-Motherway, M., Bottacini, F., Cornelissen, A., Neve, H., Heller, K.J., Noben, J.P., Dal Bello, F., van Sinderen, D., 2014. Molecular characterization of three *Lactobacillus delbrueckii* subsp. *bulgaricus* phages. Appl. Environ. Microbiol. 80, 5623–5635. https://doi.org/10.1128/AEM.01268-14.

Casey, E., McDonnell, B., White, K., Stamou, P., Crowley, T., O'Neill, I., Lavelle, K., Hayes, S., Andrea Lugli, G., Arboleya, S., James, K., Ventura, M., Martinez, I., Gueimonde, M., dal Bello, F., Nally, K., Mahony, J., van Sinderen, D., 2022. Needle in a whey-stack: PhRACS as a discovery tool for unknown phage-host combinations. mBio 13, e03334-21. https://doi.org/10.1128/mbio.03334-21.

Castro-Nallar, E., Chen, H., Gladman, S., Moore, S.C., Seemann, T., Powell, I.B., Hillier, A., Crandall, K.A., Chandr, P.S., 2012. Population genomics and phylogeography of an Australian dairy factory derived lytic bacteriophage. Genome Biol. Evol. 4, 382–393. https://doi.org/10.1093/gbe/evs017.

Chan, P.P., Lin, B.Y., Mak, A.J., Lowe, T.M., 2021. tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. Nucleic Acids Res. 49, 9077–9096. https://doi.org/10.1093/nar/gkab688.

Chandry, P.S., Moore, S.C., Boyce, J.D., Davidson, B.E., Hillier, A.J., 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. https://doi.org/10.1046/j.1365-2958.1997.5491926.x.

Chapot-Chartier, M.P., Vinogradov, E., Sadovskaya, I., Andre, G., Mistou, M.Y., Trieu-Cuot, P., Furlan, S., Bidnenko, E., Courtin, P., Péchoux, C., Hols, P., Dufre, Y.F., Kulakauskas, S., 2010. Cell surface of *Lactococcus lactis* is covered by a protective polysaccharide pellicle. J. Biol. Chem. 285, 10464–10471. https://doi.org/10.1074/jbc.M109.082958.

Chmielewska-Jeznach, M., Bardowski, J.K., Szczepankowska, A.K., 2018. Molecular, physiological and phylogenetic traits of *Lactococcus* 936-type phages from distinct dairy environments. Sci. Rep. 8, 12540. https://doi.org/10.1038/s41598-018-30371-3.

Colombo, S., Arioli, S., Gargari, G., Neri, E., Della Scala, G., Mora, D., 2018. Characterization of airborne viromes in cheese production plants. J. Appl. Microbiol. 125, 1444–1454. https://doi.org/10.1111/jam.14046.

Corpet, F., 1988. Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res. 16, 10881–10890. https://doi.org/10.1093/nar/16.22.10881.

Daly, C., Fitzgerald, G.F., Davis, R., 1996. Biotechnology of lactic acid bacteria with special reference to bacteriophage resistance. Antonie Van Leeuwenhoek 70, 99–110. https://doi.org/10.1007/BF00395928.

Deveau, H., Labrie, S.J., Chopin, M.C., Moineau, S., 2006. Biodiversity and classification of lactococcal phages. Appl. Environ. Microbiol. 72, 4338–4346. https://doi.org/10.1128/AEM.02517-05.

Dugat-Bony, E., Lossouarn, J., De Paepe, M., Sarthou, A.S., Fedala, Y., Petit, M.A., Chaillou, S., 2020. Viral metagenomic analysis of the cheese surface: a comparative study of rapid procedures for extracting viral particles. Food Microbiol. 85, 103278. https://doi.org/10.1016/j.fm.2019.103278.

Dupont, K., Vogensen, F.K., Neve, H., Bresciani, J., Josephsen, J., 2004. Identification of the receptor-binding protein in 936-species lactococcal bacteriophages. Appl. Environ. Microbiol. 70, 5818–5824. https://doi.org/10.1128/AEM.70.10.5818-5824.2004.

Forsman, P., Alatossava, T., 1994. Repeated sequences and the sites of genome rearrangements in bacteriophages of *Lactobacillus delbrueckii* subsp. *lactis*. Arch. Virol. 137, 43–54. https://doi.org/10.1007/BF01311172.

Frantzen, C.A., Holo, H., 2019. Unprecedented diversity of lactococcal group 936 bacteriophages revealed by amplicon sequencing of the portal protein gene. Viruses 11. https://doi.org/10.3390/v11050443.

Gilchrist, C.L.M., Chooi, Y.-H., 2021. Clinker & clustermap.js: automatic generation of gene cluster comparison figures. Bioinformatics 37, 2473–2475. https://doi.org/10.1093/bioinformatics/btab007.

Gobbetti, M., De Angelis, M., Di Cagno, R., Mancini, L., Fox, P.F., 2015. Pros and cons for using non-starter lactic acid bacteria (NSLAB) as secondary/adjunct starters for cheese ripening. Trends Food Sci. Technol. 45, 167–178. https://doi.org/10.1016/j.tifs.2015.07.016.

Haaber, J., Rosseau, G.M., Hammer, K., Moineau, S., 2009. Identification and characterization of the phage gene *sav*, involved in sensitivity to the lactococcal abortive infection mechanism AbiV. Appl. Environ. Microbiol. 75, 2484–2494. https://doi.org/10.1128/AEM.02093-08.

Hayes, S., Vincentelli, R., Mahony, J., Nauta, A., Ramond, L., Lugli, G.A., Ventura, M., van Sinderen, D., Cambillau, C., 2018. Functional carbohydrate binding modules identified in evolved dits from siphophages infecting various Gram-positive bacteria. Mol. Microbiol. 110, 777–795. https://doi.org/10.1111/mmi.14124.

Hayes, S., Mahony, J., Vincentelli, R., Ramond, L., Nauta, A., van Sinderen, D., Cambillau, C., 2019. Ubiquitous carbohydrate binding modules decorate 936 lactococcal siphophage virions. Viruses 11, 631. https://doi.org/10.3390/v11070631.

Hendrix, R.W., Smith, M.C.M., Burns, R.N., Ford, M.E., Hatfull, G.F., 1999. Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. Proc. Natl. Acad. Sci. U. S. A. 96, 2192–2197. https://doi.org/10.1073/pnas.96.5.2192.

Holm, L., 2022. Dali server: structural unification of protein families. Nucleic Acids Res. 50, W210–W215. https://doi.org/10.1093/nar/gkac387.

Hyatt, D., Chen, G.-L., LoCascio, P.F., Land, M.L., Larimer, F.W., Hauser, L.J., 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11, 119.

Johnson, M.E., 2013. Mesophilic and thermophilic cultures used in traditional cheesemaking. Microbiol. Spectr. 1. https://doi.org/10.1128/microbiolspec.cm-0004-2012, 10.1128/microbiolspec.cm-0004–2012.

Jolicoeur, A.P., Lemay, M.L., Beaubien, E., Bélanger, J., Bergeron, C., Bourque-Leblanc, F., Doré, L., Dupuis, M.È., Fleury, A., Garneau, J.E., Labrie, S.J., Labrie, S., Lacasse, G., Lamontagne-Drolet, M., Lessard-Hurtubise, R., Martel, B., Menasria, R., Morin-Pelchat, R., Pageau, G., Samson, J.E., Rousseau, G.M., Tremblay, D.M., Duquenne, M., Lamoureux, M., Moineau, S., 2023. Longitudinal study of *Lactococcus* phages in a Canadian cheese factory. Appl. Environ. Microbiol. 89, e0042123. https://doi.org/10.1128/aem.00421-23.

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A.F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.-Y., Lopez, R., Hunter, S., 2014. InterProScan 5: genome-scale protein function classification. Bioinformatics 30, 1236–1240. https://doi.org/10.1093/bioinformatics/btu031.

Labrie, S., Moineau, S., 2000. Multiplex PCR for detection and identification of lactococcal bacteriophages. Appl. Environ. Microbiol. 66, 987–994. https://doi.org/10.1128/AEM.66.3.987-994.2000.

Li, T.T., Tian, W.L., Gu, C.T., 2021. Elevation of *Lactococcus lactis* subsp. *cremoris* to the species level as *Lactococcus cremoris* sp. nov. and transfer of *Lactococcus lactis* subsp. *tructae* to *Lactococcus cremoris* as *Lactococcus cremoris* subsp. *tructae* comb. nov. Int. J. Syst. Evol. Microbiol. 71. https://doi.org/10.1099/ijsem.0.004727.

Lillehaug, D., 1997. An improved plaque assay for poor plaque-producing temperate lactococcal bacteriophages. J. Appl. Microbiol. 83, 85–90.

Lugli, G.A., Fontana, F., Tarracchini, C., Milani, C., Mancabelli, L., Turroni, F., Ventura, M., 2023. MEGAnnotator2: a pipeline for the assembly and annotation of microbial genomes. Microbiome Res. Rep. 2. https://doi.org/10.20517/mrr.2022.21.

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, 1–9. https://doi.org/10.3389/fmicb.2012.00335.

Mahony, J., Kot, W., Murphy, J., Ainsworth, S., Neve, H., Hansen, L.H., Heller, K.J., Sørensen, S.J., Hammer, K., Cambillau, C., Vogensen, F.K., van Sinderen, D., 2013. Investigation of the relationship between lactococcal host cell wall polysaccharide genotype and 936 phage receptor binding protein phylogeny. Appl. Environ. Microbiol. 79, 4385–4392. https://doi.org/10.1128/AEM.00653-13.

Mahony, J., Moscarelli, A., Kelleher, P., Lugli, G.A., Ventura, M., Settanni, L., van Sinderen, D., 2017a. Phage biodiversity in artisanal cheese wheys reflects the complexity of the fermentation process. Viruses 9. https://doi.org/10.3390/v9030045.

Mahony, J., Oliveira, J., Collins, B., Hanemaaijer, L., Lugli, G.A., Neve, H., Ventura, M., Kouwen, T.R., Cambillau, C., van Sinderen, D., 2017b. Genetic and functional characterisation of the lactococcal P335 phage-host interactions. BMC Genomics 18. https://doi.org/10.1186/s12864-017-3537-5.

Mahony, J., Frantzen, C., Vinogradov, E., Sadovskaya, I., Theodorou, I., Kelleher, P., Chapot-Chartier, M.-P., Cambillau, C., Holo, H., van Sinderen, D., 2020. The CWPS Rubik's cube: linking diversity of cell wall polysaccharide structures with the encoded biosynthetic machinery of selected *Lactococcus lactis* strains. Mol. Microbiol. 114, 582–596. https://doi.org/10.1111/mmi.14561.

Mallawaarachchi, V., Roach, M.J., Decewicz, P., Papudeshi, B., Giles, S.K., Grigson, S.R., Bouras, G., Hesse, R.D., Inglis, L.K., Hutton, A.L.K., Dinsdale, E.A., Edwards, R.A., 2023. Phables: from fragmented assemblies to high-quality bacteriophage genomes. Bioinformatics 39, btad586. https://doi.org/10.1093/bioinformatics/btad586.

McGrath, S., Neve, H., Seegers, J.F.M.L., Eijlander, R., Vegge, C.S., Brøndsted, L., Heller, K.J., Fitzgerald, G.F., Vogensen, F.K., Douwe, van, S., 2006. Anatomy of a Lactococcal Phage Tail. J. Bacteriol. 188, 3972–3982. https://doi.org/10.1128/jb.00024-06.

Miklič, A., Rogelj, I., 2003. Characterization of lactococcal bacteriophages isolated from Slovenian dairies. Int. J. Food Sci. Technol. 38, 305–311. https://doi.org/10.1046/j.1365-2621.2003.00676.x.

Millen, A.M., Romero, D.A., 2016. Genetic determinants of lactococcal C2viruses for host infection and their role in phage evolution. J. Gen. Virol. 97, 1998–2007. https://doi.org/10.1099/jgv.0.000499.

Millen, A.M., Magill, D., Romero, D., Simdon, L., 2023. Evolved distal tail protein of skunaviruses facilitates adsorption to exopolysaccharide-encoding lactococci. Microbiome Res. Rep. 2. https://doi.org/10.20517/mrr.2023.29.

Moineau, S., Fortier, J., Ackermann, H.-W., Pandian, S., 1992. Characterization of lactococcal bacteriophages from Quebec cheese plants. Can. J. Microbiol. 38, 875–882. https://doi.org/10.1139/m92-143.

Moineau, S., Borkaev, M., Holler, B.J., Walker, S.A., Kondo, J.K., Vedamuthu, E.R., Vandenbergh, P.A., 1996. Isolation and characterization of lactococcal bacteriophages from cultured buttermilk plants in the United States. J. Dairy Sci. 79, 2104–2111. https://doi.org/10.3168/jds.S0022-0302(96)76584-0.

Muhammed, M.K., Kot, W., Neve, H., Mahony, J., Castro-Mejía, J.L., Krych, L., Hansen, L. H., Nielsen, D.S., Sørensen, S.J., Heller, K.J., van Sinderen, D., Vogensen, F.K., 2017. Metagenomic analysis of dairy bacteriophages: extraction method and pilot study on whey samples derived from using undefined and defined mesophilic starter cultures. Appl. Environ. Microbiol. 83. https://doi.org/10.1128/AEM.00888-17.

Murphy, J., Royer, B., Mahony, J., Hoyles, L., Heller, K., Neve, H., Bonestroo, M., Nauta, A., van Sinderen, D., 2013. Biodiversity of lactococcal bacteriophages isolated from 3 Gouda-type cheese-producing plants. J. Dairy Sci. 96, 4945–4957. https://doi.org/10.3168/jds.2013-6748.

Murphy, J., Klumpp, J., Mahony, J., O'Connell-Motherway, M., Nauta, A., van Sinderen, D., 2014. Methyltransferases acquired by lactococcal 936-type phage provide protection against restriction endonuclease activity. BMC Genomics 15, 831. https://doi.org/10.1186/1471-2164-15-831.

Murphy, J., Bottacini, F., Mahony, J., Kelleher, P., Neve, H., Zomer, A., Nauta, A., van Sinderen, D., 2016. Comparative genomics and functional analysis of the 936 group of lactococcal Siphoviridae phages. Sci. Rep. 6, 21345. https://doi.org/10.1038/srep21345.

Oliveira, J., Mahony, J., Hanemaaijer, L., Kouwen, T.R.H.M., van Sinderen, D., 2018. Biodiversity of bacteriophages infecting *Lactococcus lactis* starter cultures. J. Dairy Sci. 101, 96–105. https://doi.org/10.3168/jds.2017-13403.

Page, A.J., Cummins, C.A., Hunt, M., Wong, V.K., Reuter, S., Holden, M.T.G., Fookes, M., Falush, D., Keane, J.A., Parkhill, J., 2015. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics 31, 3691–3693. https://doi.org/10.1093/bioinformatics/btv421.

Paillet, T., Lamy-Besnier, Q., Figueroa, C., Petit, M.-A., Dugat-Bony, E., 2024. Dynamics of the viral community on the cheese surface during maturation and persistence across production years. mSystems 9, e0020124. https://doi.org/10.1128/msystems.00201-24.

Parlindungan, E., Sadovskaya, I., Vinogradov, E., Lugli, G.A., Ventura, M., van Sinderen, D., Mahony, J., 2024. Novel cell wall polysaccharide genotypes and structures of lactococcal strains isolated from milk and fermented foods. Int. J. Food Microbiol. 110840. https://doi.org/10.1016/j.ijfoodmicro.2024.110840.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J. H., Ferrin, T.E., 2021. UCSF ChimeraX: structure visualization for researchers, educators, and developers. Protein Sci. 30, 70–82. https://doi.org/10.1002/pro.3943.

Pritchard, L., Glover, R.H., Humphris, S., Elphinstone, J.G., Toth, I.K., 2016. Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. Anal. Methods. https://doi.org/10.1039/c5ay02550h.

Roberts, R.J., Vincze, T., Posfai, J., Macelis, D., 2009. REBASE-A database for DNA restriction and modification: enzymes, genes and genomes. Nucleic Acids Res. 38. https://doi.org/10.1093/nar/gkp874.

Roux, S., Emerson, J.B., Eloe-Fadrosh, E.A., Sullivan, M.B., 2017. Benchmarking viromics: An in silico evaluation of metagenome-enabled estimates of viral community composition and diversity. PeerJ 2017. https://doi.org/10.7717/peerj.3817.

Sadovskaya, I., Vinogradov, E., Courtin, P., Armalyte, J., Meyrand, M., Giaouris, E., Palussière, S., Furlan, S., Péchoux, C., Ainsworth, S., Mahony, J., van Sinderen, D., Kulakauskas, S., Guérardel, Y., Chapot-Chartier, M.-P., 2017. Another brick in the wall: a rhamnan polysaccharide trapped inside peptidoglycan of *Lactococcus lactis*. mBio 8. https://doi.org/10.1128/mbio.01303-17, 10.1128/mbio.01303-17.

Saeed, A.I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., Braisted, J., Klapa, M., Currier, T., Thiagarajan, M., 2003. TM4: a free, open-source system for microarray data management and analysis. Biotechniques 34, 374–378.

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

Samson, J.E., Moineau, S., 2013. Bacteriophages in food fermentations: new frontiers in a continuous arms race. Annu. Rev. Food Sci. Technol. 4, 347–368. https://doi.org/10.1146/annurev-food-030212-182541.

Scaltriti, E., Moineau, S., Launay, H., Masson, J.-Y., Rivetti, C., Ramoni, R., Campanacci, V., Tegoni, M., Cambillau, C., 2010. Deciphering the function of lactococcal phage ul36 Sak domains. J. Struct. Biol. 170, 462–469. https://doi.org/10.1016/j.jsb.2009.12.021.

Scaltriti, E., Polverini, E., Grolli, S., Eufemi, E., Moineau, S., Cambillau, C., Ramoni, R., 2013. The DNA binding mechanism of a SSB protein from *Lactococcus lactis* siphophage p2. Biochim. Biophys. Acta 1834, 1070–1076. https://doi.org/10.1016/j.bbapap.2013.02.014.

Sciara, G., Bebeacua, C., Bron, P., Tremblay, D., Ortiz-Lombardia, M., Lichière, 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. https://doi.org/10.1073/pnas.1000232107.

Söding, J., Biegert, A., Lupas, A.N., 2005. The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res. 33, W244–W248.

Sutton, T.D.S., Clooney, A.G., Ryan, F.J., Ross, R.P., Hill, C., 2019. Choice of assembly software has a critical impact on virome characterisation. Microbiome 7, 12. https://doi.org/10.1186/s40168-019-0626-5.

Szczepańska, A.K., Hejnowicz, M.S., Kołakowski, P., Bardowski, J., 2007. Biodiversity of *Lactococcus lactis* bacteriophages in Polish dairy environment. Acta Biochim. Pol. 54, 151–158. https://doi.org/10.18388/abp.2007_3281.

Tamang, J.P., Watanabe, K., Holzapfel, W.H., 2016. Review: diversity of microorganisms in global fermented foods and beverages. Front. Microbiol. 7, 377. https://doi.org/10.3389/fmicb.2016.00377.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22, 4673–4680. https://doi.org/10.1093/nar/22.22.4673.

van Kempen, M., Kim, S.S., Tumescheit, C., Mirdita, M., Lee, J., Gilchrist, C.L.M., Söding, J., Steinegger, M., 2024. Fast and accurate protein structure search with Foldseek. Nat. Biotechnol. 42, 243–246. https://doi.org/10.1038/s41587-023-01773-0.

Vegge, C.S., Neve, H., Brøndsted, L., Heller, K.J., Vogensen, F.K., 2006. Analysis of the collar-whisker structure of temperate lactococcal bacteriophage TP901-1. Appl. Environ. Microbiol. 72, 6815–6818. https://doi.org/10.1128/AEM.01033-06.

White, K., Eraclio, G., McDonnell, B., Bottacini, F., Lugli Gabriele, A., Ventura, M., Volontè, F., Dal Bello, F., Mahony, J., van Sinderen, D., 2024. A multifaceted investigation of lactococcal strain diversity in undefined mesophilic starter cultures. Appl. Environ. Microbiol. 90, e02152-23. https://doi.org/10.1128/aem.02152-23.

White, K., Eraclio, G., McDonnell, B., Lugli, G.A., Crowley, T., Ventura, M., Volonté, F., Cambillau, C., Dal Bello, F., Mahony, J., van Sinderen, D., 2025. Lactococcal phage–host profiling through binding studies between cell wall polysaccharide types and Skunavirus receptor-binding proteins. Microb. Genom. 11. https://doi.org/10.1099/mgen.0.001395.