



Article

Prophage-Derived Regions in *Curtobacterium* Genomes: Good Things, Small Packages

Peter Evseev ^{1,*}, Anna Lukianova ¹, Rashit Tarakanov ², Anna Tokmakova ^{1,3}, Anastasia Popova ⁴, Eugene Kulikov ^{3,5}, Mikhail Shneider ¹, Alexander Ignatov ⁶ and Konstantin Miroshnikov ^{1,*}

- ¹ Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Miklukho-Maklaya Str., 117997 Moscow, Russia
 - ² Department of Plant Protection, Russian State Agrarian University—Moscow Timiryazev Agricultural Academy, Timiryazevskaya Str. 49, 127434 Moscow, Russia
 - ³ School of Biological and Medical Physics, Moscow Institute of Physics and Technology National Research University, Institutskiy Per, 9, 141701 Dolgoprudny, Russia
 - ⁴ State Research Center for Applied Microbiology and Biotechnology, 142279 Obolensk, Russia
 - ⁵ Research Center of Biotechnology, Winogradsky Institute of Microbiology, Russian Academy of Sciences, Prosp. 60-letia Oktyabrya, 7-2, 117312 Moscow, Russia
 - ⁶ Agrobiotechnology Department, Agrarian and Technological Institute, RUDN University, Miklukho-Maklaya Str. 6, 117198 Moscow, Russia
- * Correspondence: evseev@ibch.ru (P.E.); kmi@ibch.ru (K.M.)

Abstract: *Curtobacterium* is a genus of Gram-positive bacteria within the order *Actinomycetales*. Some *Curtobacterium* species (*C. flaccumfaciens*, *C. plantarum*) are harmful pathogens of agricultural crops such as soybean, dry beans, peas, sugar beet and beetroot, which occur throughout the world. Bacteriophages (bacterial viruses) are considered to be potential curative agents to control the spread of harmful bacteria. Temperate bacteriophages integrate their genomes into bacterial chromosomes (prophages), sometimes substantially influencing bacterial lifestyle and pathogenicity. About 200 publicly available genomes of *Curtobacterium* species, including environmental metagenomic sequences, were inspected for the presence of sequences of possible prophage origin using bioinformatic methods. The comparison of the search results with several ubiquitous bacterial groups showed the relatively low level of the presence of prophage traces in *Curtobacterium* genomes. Genomic and phylogenetic analyses were undertaken for the evaluation of the evolutionary and taxonomic positioning of predicted prophages. The analyses indicated the relatedness of *Curtobacterium* prophage-derived sequences with temperate actinophages of siphoviral morphology. In most cases, the predicted prophages can represent novel phage taxa not described previously. One of the predicted temperate phages was induced from the *Curtobacterium* genome. Bioinformatic analysis of the modelled proteins encoded in prophage-derived regions led to the discovery of some 100 putative glycopolymer-degrading enzymes that contained enzymatic domains with predicted cell-wall- and cell-envelope-degrading activity; these included glycosidases and peptidases. These proteins can be considered for the experimental design of new antibacterials against *Curtobacterium* phytopathogens.

Keywords: viral genomics; analysis of genomic data; *Curtobacterium*; phages; prophages; endolysins; glycan depolymerases



Citation: Evseev, P.; Lukianova, A.; Tarakanov, R.; Tokmakova, A.; Popova, A.; Kulikov, E.; Shneider, M.; Ignatov, A.; Miroshnikov, K. Prophage-Derived Regions in *Curtobacterium* Genomes: Good Things, Small Packages. *Int. J. Mol. Sci.* **2023**, *24*, 1586. <https://doi.org/10.3390/ijms24021586>

Academic Editors: Malgorzata Lobočka and Jakub Barylski

Received: 30 November 2022

Revised: 10 January 2023

Accepted: 11 January 2023

Published: 13 January 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Bacteria belonging to the genus *Curtobacterium* are of great interest. These actinomycetes have been found in many of Earth's microbiomes. In spite of being unable to form spores, *Curtobacterium* spp. are nevertheless abundant in soil [1,2], marine sediments [3] and in the air up to the stratosphere [4]. *Curtobacteria* demonstrate enhanced tolerance to drought, salinity, UV irradiation and metal ions, and have been suggested to play an important role in plant adaptation to stress conditions [5–7]. Being typical inhabitants of the

rhizosphere of many plants, *Curtobacterium* spp. have been shown to promote seed germination [8,9] and plant growth [10,11], and to suppress the growth of phytopathogens [12,13]. A pathovar, *Curtobacterium flaccumfaciens* pv. *flaccumfaciens*, is an economically important plant pathogen [14] and, occasionally, a human pathogen [15,16]. *Curtobacterium* spp. are often found in plant microbiomes relevant to bacterial diseases [17], and may promote the development of pathogenic fungi [18].

Genetic determinants of environmental adaptation and plant pathogenicity among *Curtobacterium* sp. are not completely understood [19], and may be influenced by mobile elements such as plasmids as well as prophages or their remnants. Bacteriophages are viruses that infect bacteria. Lytic or virulent phages reproduce, lyse the cell and release progeny phages upon infecting bacteria. Lysogenic or temperate phages may multiply via the lytic cycle or adopt a latent state in the cell. Prophages are comprised of phage DNA in the integrated state. Fully functional prophages are capable of excising from the bacterial chromosome, either spontaneously or in response to specific signals; they then lyse the host cells at some subsequent generation upon induction [20]. As a result of mutations, prophages can lose the ability for induction, but still play a role in bacterial adaptation, introducing new fitness factors for the host cell [21–23].

Prophages can constitute as much as 10–20% of a bacterial genome and contribute to interstrain variability. Prophages may harbour virulence factors and pathogenicity islands, thereby playing an important role in the emergence of pathogens [24,25]. The purpose of the present study was to ascertain the repertoire, and possible role, of prophage sequences in the genomes of *Curtobacterium* sp., to propose the origin and phylogeny of initial phages, considering recent changes in bacteriophage taxonomy, and to reveal and systematise information about cytolytic enzymes encoded in prophage sequences which may have biotechnological potential.

2. Results

2.1. Predicted Prophages in *Curtobacterial* and Other Bacterial Genomes

In early summer 2022, genomes of 197 strains assigned to the genus *Curtobacterium* were identified in the NCBI Genome Database [26]. Thirty-six strains were classified as *C. flaccumfaciens* and 28 strains were classified as pathovars of *C. flaccumfaciens*. Previous research [19] revealed inconsistencies in the taxonomy of genus *Curtobacterium* and called for taxonomic re-evaluation. Thus, all 197 genomes have been used for analyses.

The search for genomic regions of prophage origin has been performed using the online server PHASTER [27] and the pipeline PhiSpy [28]. PHASTER currently seems to be the most widely used prophage prediction tool [29], and a recently published comparative study demonstrated the high levels of accuracy, precision, recall and f_1 score of PhiSpy [30]. PHASTER has identified 161 possible prophage regions, while only one of them has been defined as an intact phage; PhiSpy has found 75 prophage regions. The most substantial part of similar predicted prophages referred to 32 genomic loci containing the genes, encoding UDP-N-acetylglucosamine transferase, N-acetylglucosaminyl-diphospho-decaprenol L-rhamnosyltransferase, teichoic acid biosynthesis protein F, dTDP-4-dehydrorhamnose 3,5-epimerase, glycosyltransferase of PMT family, undecaprenyl-phosphate 4-deoxy-4-formamido-L-arabinose transferase, 4-amino-4-deoxy-L-arabinose transferase or related glycosyltransferase, dTDP-glucose 4,6-dehydratase (RmlB) and dTDP-4-dehydrorhamnose reductase (RfbD), according to the results of PHASTER annotation, BLAST [31] and HH-pred [32] searches. These enzymes participate in producing cell wall polysaccharides (CWPSs) and exopolysaccharide (EPS) in Gram-positive bacteria, and lipopolysaccharides in Gram-negative bacteria [33–35]. Some genes encoding enzymes for CWPS and EPS could have been acquired via horizontal gene transfer [35]. Several predicted prophages have not contained phage structural genes, but have been shown to possess genes typical for plasmids and other mobile elements such as transposases and conjugative transfer genes [36]. Thus, the results of these bioinformatic tools need additional post-processing,

but they can be useful for the assessment of the level of mobile elements' presence in bacterial genomes.

To estimate the comparative level of saturation of genomes of different bacterial groups with mobile elements, 12 additional groups of pathogenic bacteria, including phytopathogens, were analysed using PHASTER. Each of these groups included 100 complete and draft genomes belonging to genera *Bacillus*, *Clavibacter*, *Clostridium*, *Microbacterium*, *Nocardia*, *Ralstonia*, *Streptomyces*, *Synechococcus* and *Xanthomonas*, and species *Escherichia coli*, *Mycobacterium tuberculosis* and *Pseudomonas aeruginosa*. The results of PHASTER analysis indicated a smaller number of predicted prophage-derived regions in genus *Curtobacterium* and closely related genus *Clavibacter*, compared with other taxa (Figure 1).

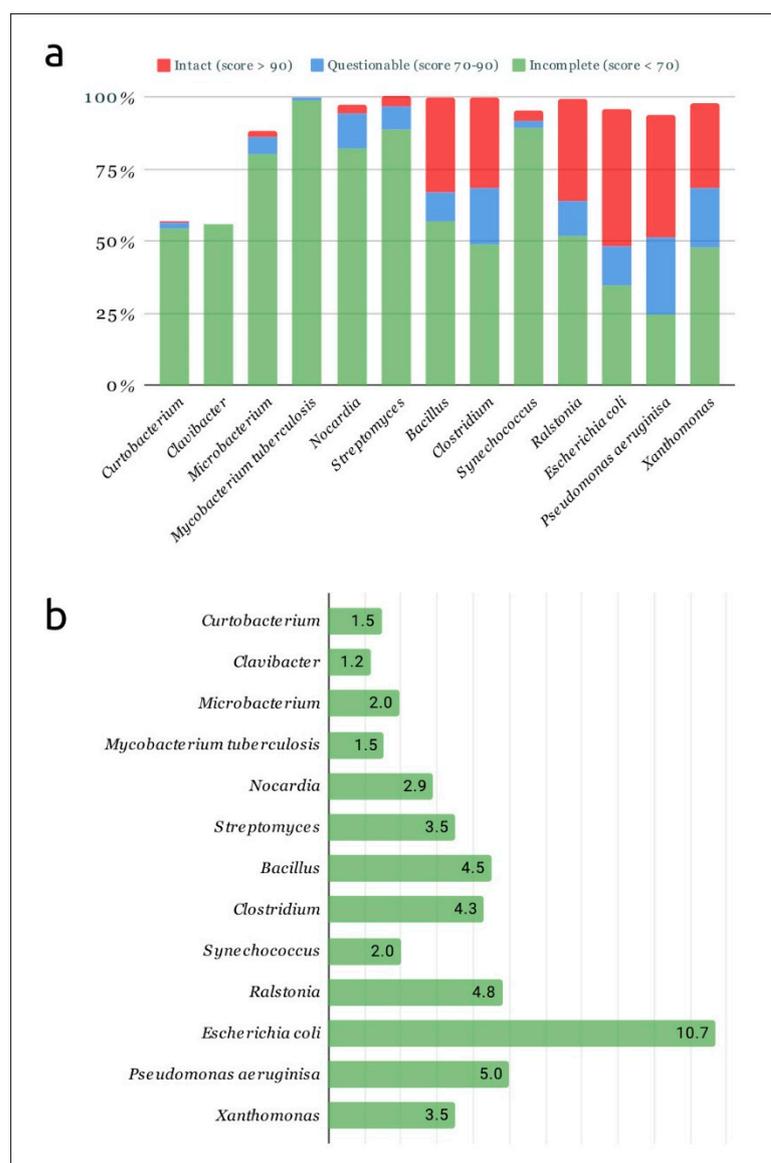


Figure 1. (a) Percentage of genomes of different analysed taxonomic groups where PHASTER predicted the presence of prophages. Scoring was carried out according to the PHASTER criteria [37]. (b) Average number of prophages per genome predicted with PHASTER, where the prophages were detected.

2.2. Post-Processing of Phaster and PhiSpy Results

Predictions by PHASTER and PhiSpy often yielded similar overlapping genomic regions, but manual inspection of putative prophage-derived regions (PDRs) revealed

some discrepancies in the predictions of these two tools, and possible inaccuracies in the definition of the borders of the prophage-derived genomic region. PHASTER predicted potential sites for prophage insertion for only 22 putative prophages and the prediction did not look reliable, so the borders of PDRs were suggested on the basis of the phage origin of genes and comparisons with known phage genomes. Post-processing and manual curation of the prediction were conducted in ways similar to those described in [38] (Figure 2).

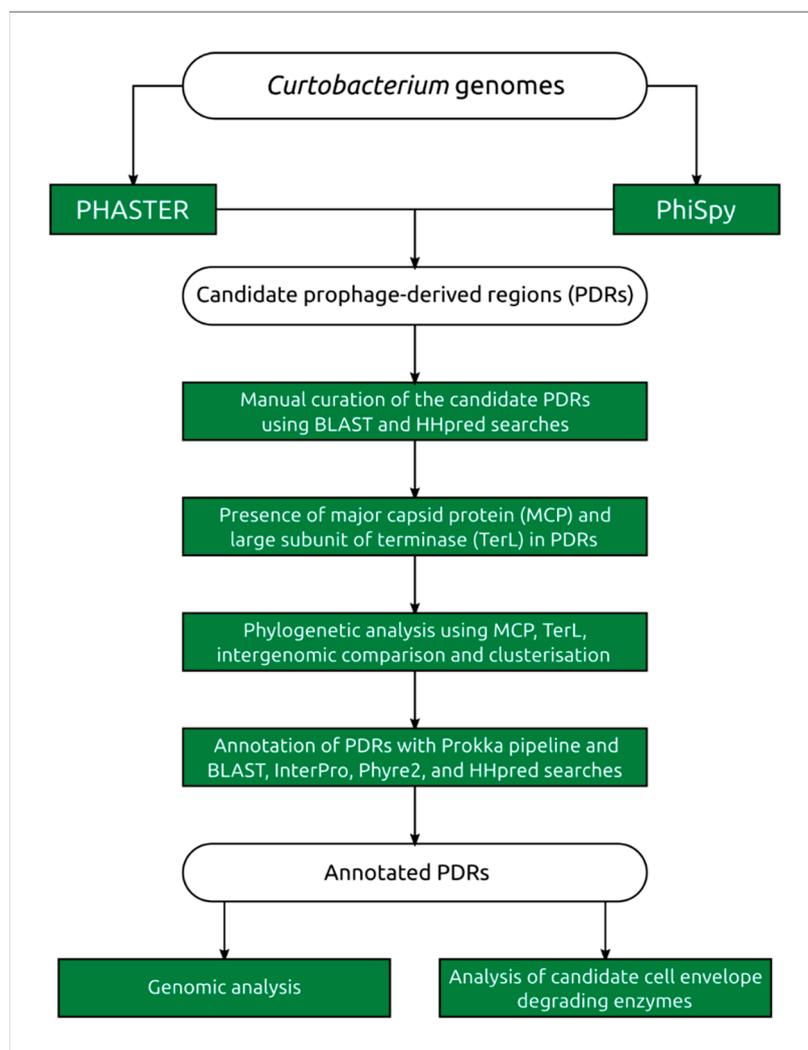


Figure 2. Flow chart of the identification and characterisation of *Curtobacterium* prophage-derived regions. The rectangles represent the tools and methods used for identification of PDRs; the rounded rectangles represent the data used for the analyses.

All of the predicted regions were checked through an analysis of the gene content of predicted regions and their possible prophage origin, with a BLAST search using the NCBI and custom phage databases. The genomic content of these 64 predicted prophage regions was additionally checked through comparisons with genomes of sequenced bacteriophages using an HHpred search. Putative genes of holins and spanins were also checked by the prediction of transmembrane regions. This post-processing revealed 70 prophage-derived regions (PDRs) possessing phage structural genes. Sixty-four regions were found to contain the genes encoding major capsid protein (MCP) and terminase large subunit (TerL). They might represent recently grounded or intact prophages and have been considered for further analysis. General features of these PDRs are listed in Table 1, with their order corresponding to the clustering described below. Genomic regions of putative prophages corresponding to the PDRs

C_sp_UCD-KPL2560||2, C_sp_MCLR17_036||1, C_luteum_NS184||1, C_sp_MCBA15_004||2, C_sp_MCBD17_003||1, C_luteum_NS184||1, C_sp_MCBA15_004||2 and C_sp_MCBD17_003||1 were trimmed according to the contigs' limits.

Table 1. General features of genomes of curated, predicted, prophage-derived regions (PDRs).

Group	PDR	Size, kb	GC, %	Bacterial Strain
1	C_sp_C1 1	47.1	69.5	<i>Curtobacterium</i> sp. C1
2	C_sp_MCSS17_015 2	31.5	67.4	<i>Curtobacterium</i> sp. MCSS17_015
3	C_f_S5_26 1	27.2	68.3	<i>Curtobacterium flaccumfaciens</i> S5_26
	C_sp_MCLR17_036 1	26.6	68.6	<i>Curtobacterium</i> sp. MCLR17_036
4	C_sp_MCLR17_032 1	38.3	62.5	<i>Curtobacterium</i> sp. MCLR17_032
5	C_luteum_NS184 1	22.5	62.3	<i>Curtobacterium luteum</i> NS184
	C_sp_VKM_Ac-1376 2	39.4	64.1	<i>Curtobacterium</i> sp. VKM Ac-1376
6	C_sp_MCBD17_030 2	39.4	65.6	<i>Curtobacterium</i> sp. MCBD17_030
7	C_sp_UCD-KPL2560 2	23.2	66.6	<i>Curtobacterium</i> sp. UCD-KPL2560
8	C_sp_MCLR17_034 1	40	61.3	<i>Curtobacterium</i> sp. MCLR17_034
9	C_sp_VKM_Ac-2884 1	43	69	<i>Curtobacterium</i> sp. VKM_Ac-2884
10	C_sp_MCBD17_008 1	35.7	69.7	<i>Curtobacterium</i> sp. MCBD17_008
	C_sp_WW7 3	34.8	69.7	<i>Curtobacterium</i> sp. WW7
11	C_citreum_DSM_20528 1	36	69.6	<i>Curtobacterium citreum</i> DSM 20528
	C_citreum_JCM_1345 1	35.9	69.6	<i>Curtobacterium citreum</i> JCM 1345
	C_f_VKM_Ac-1386 1	38.3	69.3	<i>Curtobacterium flaccumfaciens</i> VKM Ac-1386
	C_sp_Csp1 1	37.5	70.1	<i>Curtobacterium</i> sp. Csp1
	C_sp_Ferrero 1	34.9	65.4	<i>Curtobacterium</i> sp. Ferrero
	C_sp_MCLR17_042 1	35.5	65.8	<i>Curtobacterium</i> sp. MCLR17_042
	C_sp_MCSS17_006 1	35.6	65.7	<i>Curtobacterium</i> sp. MCSS17_006
	C_sp_MCSS17_011 1	36.7	65.4	<i>Curtobacterium</i> sp. MCSS17_011
12	C_sp_PhB172 1	35.9	65.3	<i>Curtobacterium</i> sp. PhB172
	C_f_p_f_BRIP_70607 1	34.6	65.8	<i>Curtobacterium flaccumfaciens</i> pv. <i>flaccumfaciens</i> BRIP_70607
	C_f_p_oortii_CFBP_1384 1	36.3	65.6	<i>Curtobacterium flaccumfaciens</i> pv. <i>oortii</i> CFBP 1384
13	C_sp_MCBD17_028 1	39.3	69.4	<i>Curtobacterium</i> sp. MCBD17_028
	C_sp_PhB137 1	34.1	65.2	<i>Curtobacterium</i> sp. PhB137
	C_f_p_f_CFBP_3417 1	17.5	67.8	<i>Curtobacterium flaccumfaciens</i> pv. <i>flaccumfaciens</i> CFBP 3417
	C_pusillum_ATCC_19096 1	17	68.5	<i>Curtobacterium pusillum</i> ATCC 19096
	C_sp_Csp2 1	16.8	68.9	<i>Curtobacterium</i> sp. Csp2 1
	C_sp_MCBD17_029 1	17.2	69.1	<i>Curtobacterium</i> sp. MCBD17_029 1
	C_sp_MCBD17_032 1	17.7	69.3	<i>Curtobacterium</i> sp. MCBD17_032 1
	C_sp_MCJR17_043 1	17.2	69.2	<i>Curtobacterium</i> sp. MCJR17_043 1
	C_sp_MCJR17_055 1	17.2	69.2	<i>Curtobacterium</i> sp. MCJR17_055 1
	C_sp_MCPF17_015 1	17.2	69.2	<i>Curtobacterium</i> sp. MCPF17_015 1
14	C_sp_PhB146 1	17.8	69.1	<i>Curtobacterium</i> sp. PhB146 1
	C_sp_VKM_Ac-1376 1	17.2	69.5	<i>Curtobacterium</i> sp. VKM Ac-1376
	C_sp_VKM_Ac-2861 1	17.5	69.2	<i>Curtobacterium</i> sp. VKM Ac-2861
	C_sp_9128 1	38.4	69.2	<i>Curtobacterium</i> sp. 9128
15	C_sp_MCBA15_012 1	41.2	70.8	<i>Curtobacterium</i> sp. MCBA15_012
	C_f_p_f_CFBP_8824 1	16.5	69.5	<i>Curtobacterium flaccumfaciens</i> pv. <i>flaccumfaciens</i> CFBP 8824
	C_f_p_oortii_CFBP_1384 2	16.5	68.3	<i>Curtobacterium flaccumfaciens</i> pv. <i>oortii</i> _CFBP 1384
	C_sp_MCBA15_008 1	15.2	64.6	<i>Curtobacterium</i> sp. MCBA15_008
	C_sp_MCBD17_035 1	16.5	66	<i>Curtobacterium</i> sp. MCBD17_035
16	C_sp_MCBD17_040 2	16.2	66.5	<i>Curtobacterium</i> sp. MCBD17_040
	C_sp_MCSS17_011 2	16.5	65.3	<i>Curtobacterium</i> sp. MCSS17_011
	C_sp_YC1 2	17.2	65.2	<i>Curtobacterium</i> sp. YC1
16	C_sp_MCSS17_007 1	37.9	62.1	<i>Curtobacterium</i> sp. MCSS17_007

Table 1. Cont.

Group	PDR	Size, kb	GC, %	Bacterial Strain
17	C_sp_MCBA15_004 1	46.3	65.4	<i>Curtobacterium</i> sp. MCBA15_004
	C_sp_MCBD17_021 1	47.1	67.8	<i>Curtobacterium</i> sp. MCBD17_021
	C_sp_MCSS17_015 1	45	63.8	<i>Curtobacterium</i> sp. MCSS17_015
	C_sp_VKM_Ac-2884 2	43.9	69.2	<i>Curtobacterium</i> sp. VKM Ac-2884
18	C_albidum_DSM_20512 1	38.2	70.2	<i>Curtobacterium albidum</i> DSM 20512
	C_sp_Ferrero 3	41	70	<i>Curtobacterium</i> sp. Ferrero
	C_sp_MCBA15_004 2	14	69.8	<i>Curtobacterium</i> sp. MCBA15_004
19	C_sp_ISL-83 1	36.7	66.9	<i>Curtobacterium</i> sp. ISL-83
	C_sp_MCBD17_003 1	24.7	68.5	<i>Curtobacterium</i> sp. MCBD17_003
20	C_f_p_f_CFBP_3417 2	37	67.3	<i>Curtobacterium flaccumfaciens</i> pv. <i>flaccumfaciens</i> CFBP 3417
	C_sp_MCBA15_009 3	36.8	67.8	<i>Curtobacterium</i> sp. MCBA15_009
	C_sp_MCBD17_026 1	36.8	67	<i>Curtobacterium</i> sp. MCBD17_026
	C_sp_MCPF17_003 2	37.3	68.5	<i>Curtobacterium</i> sp. MCPF17_003
	C_sp_PhB25 1	36.6	69	<i>Curtobacterium</i> sp. PhB25
	C_sp_VKM_Ac-1796 1	37.2	67.3	<i>Curtobacterium</i> sp. VKM Ac-1796
	C_sp_VKM_Ac-2889 1	37.2	67.3	<i>Curtobacterium</i> sp. VKM Ac-2889
	C_sp_WW7 2	36.4	66.6	<i>Curtobacterium</i> sp. WW7

2.3. Intergenomic Comparison, Phylogenetic Analysis, Annotations and General Genomic Features

Intergenomic comparisons of 64 predicted and curated PDRs conducted using the Virus Intergenomic Distance Calculator (VIRIDIC) [39] (Figure 3) indicated both the relatedness of putative prophages, and a complicated picture of their relationships. Applying the 95% similarity threshold, two groups consisting of two and four PDRs can be considered as the same species. At least 46 PDRs could be grouped in several clusters, but intergenomic distances between different species were higher than the 70% genus cut-off for only two species. This assumes that either most PDRs represent distant taxonomic groups of functional phages, or that these PDRs represent defective prophages, including those inherited vertically before quickly diverging after the loss of the possibility to be induced. Some prophage regions have shown similarities to PDRs belonging to different clusters. This might be the result of genetic mosaicism, featuring phage evolution, especially the evolution of temperate phages [40–42]. Thus, clustering of the predicted PDRs using intergenomic similarities can be inconsistent.

Phylogenetic analysis was conducted using the major capsid protein (MCP) and large subunit of terminase (TerL) amino acid sequences encoded in the curated prophage regions and close homologous sequences found with a BLAST search using complete phage genomes available in the NCBI Genome database as of July 2022. The search did not yield the same results for these two proteins, and the topologies of the trees were not identical (Figure 4, Supplementary Figures S1–S3), even though the compositions of clades containing the predicted prophage regions were similar. The dissimilarities might also be the result of the mosaic evolution of phage genomes.

Clustering was performed using the results of MCP phylogenetic analysis. Representatives of each prophage cluster, shown in Figure 4, have been annotated manually using the procedure described in Section 4.2 (Figure 5); the remaining PDRs were annotated using the Prokka pipeline and a custom database constructed with manually annotated PDRs. All PDRs (Supplementary File S1) contained the genes of the tail tape measure protein and did not contain the genes of the tail sheath protein, indicating the siphoviral morphology of corresponding prophages [43]. A total of 197 *Curtobacterium* genomes used for the analyses were also checked for the presence of tail sheath protein homologues using a BLAST search, with the set of tail sheath protein sequences that were used earlier for the characterisation of the evolutionary history of proteins of this class [44]; no such homologues have been found. All PDRs not trimmed by contig limits contained tyrosine integrase genes and DNA-binding protein genes similar to the genes of phage λ lysogeny decision-making proteins

Cro and CII. These findings indicate the temperate lifestyle of putative prophages [45]. Interestingly, the groups 1, 6, 9, 10, 11, 18 and 19 (Table 1) contained peptidase family M78 genes neighbouring the integrase genes (Figure 5a). Similar metallo-endopeptidases participate in the regulation of the excision of the ICEBs1 transposon of *Bacillus subtilis*, cleaving the immunity repressor [46].

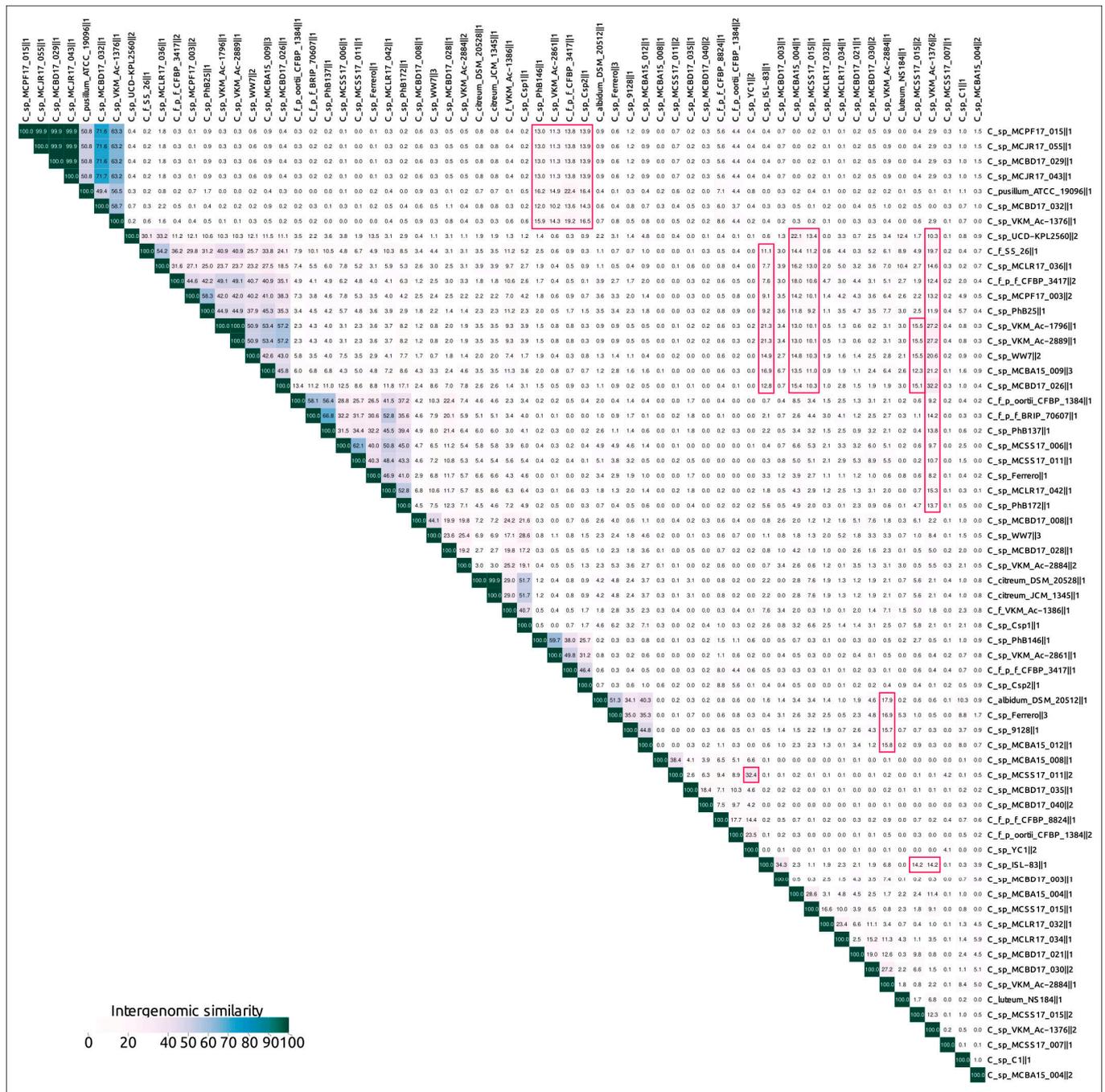


Figure 3. VIRIDIC-generated heatmap of 64 curated PRDs containing major capsid protein and terminase genes. The colour coding indicates the clustering of the phage genomes based on intergenomic similarity. The numbers represent the similarity values for each genome pair, rounded to the first decimal. Examples of heatmap islands indicating similarities between prophages not belonging to main clusters are outlined in red.

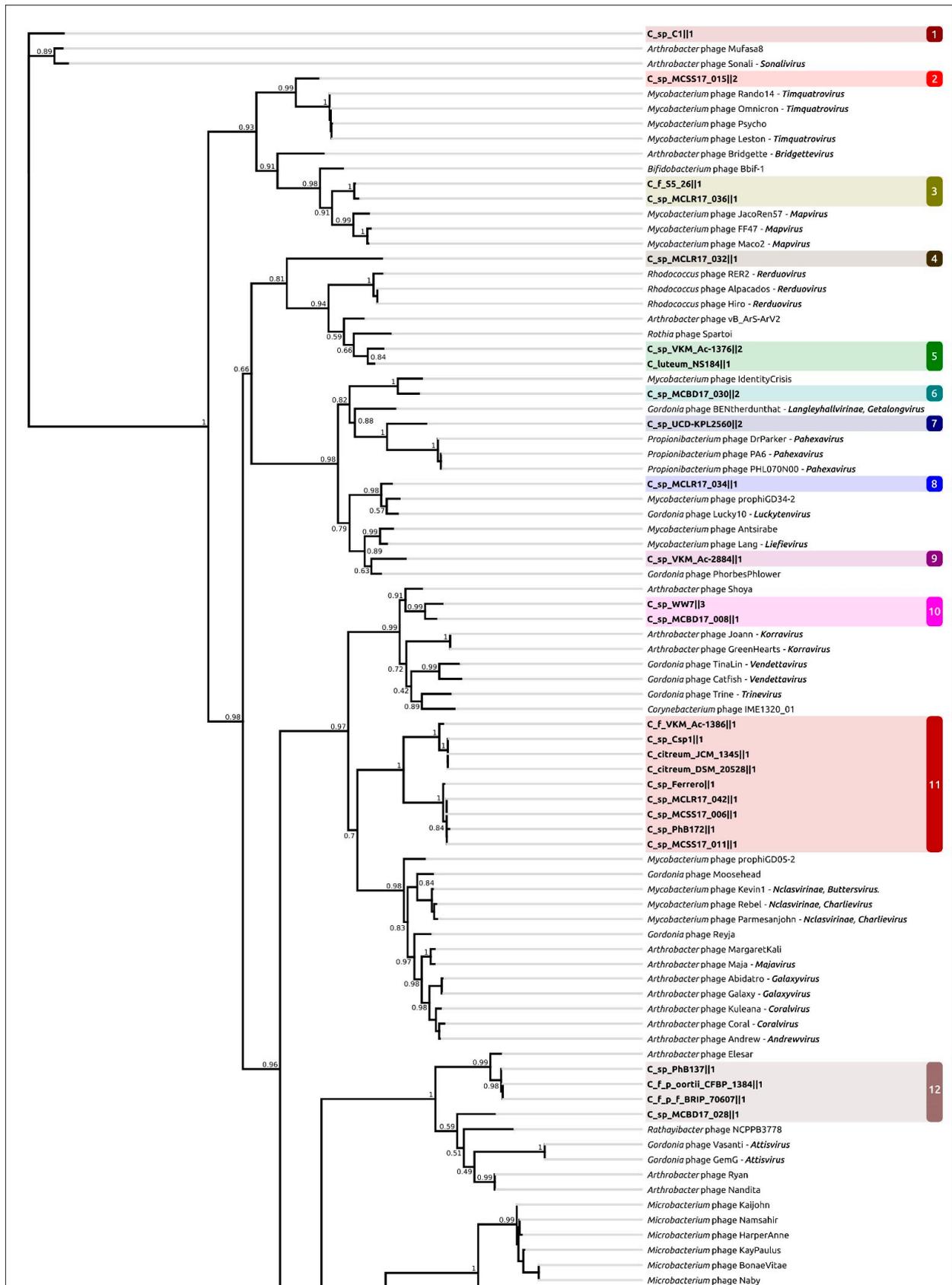


Figure 4. Cont.

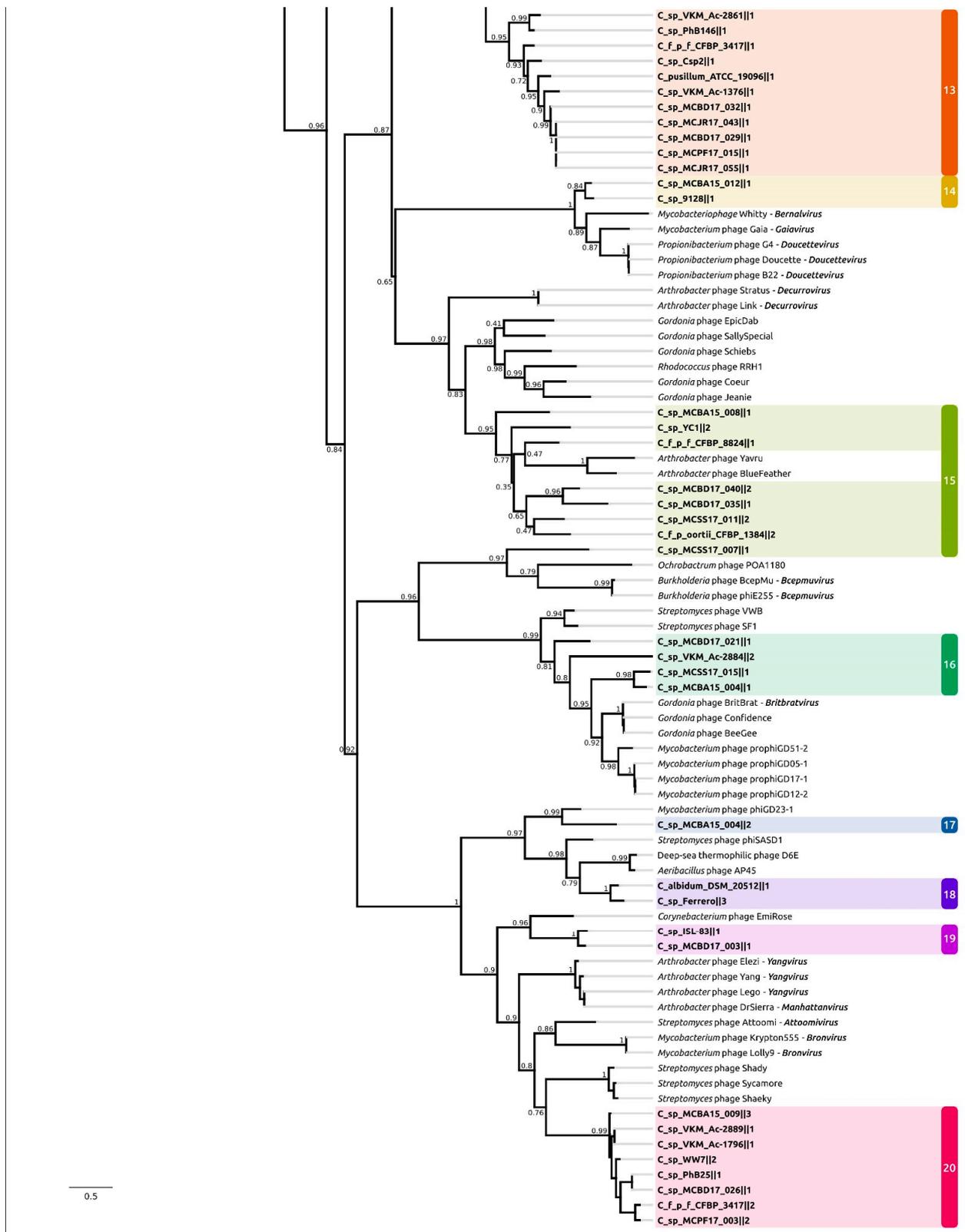
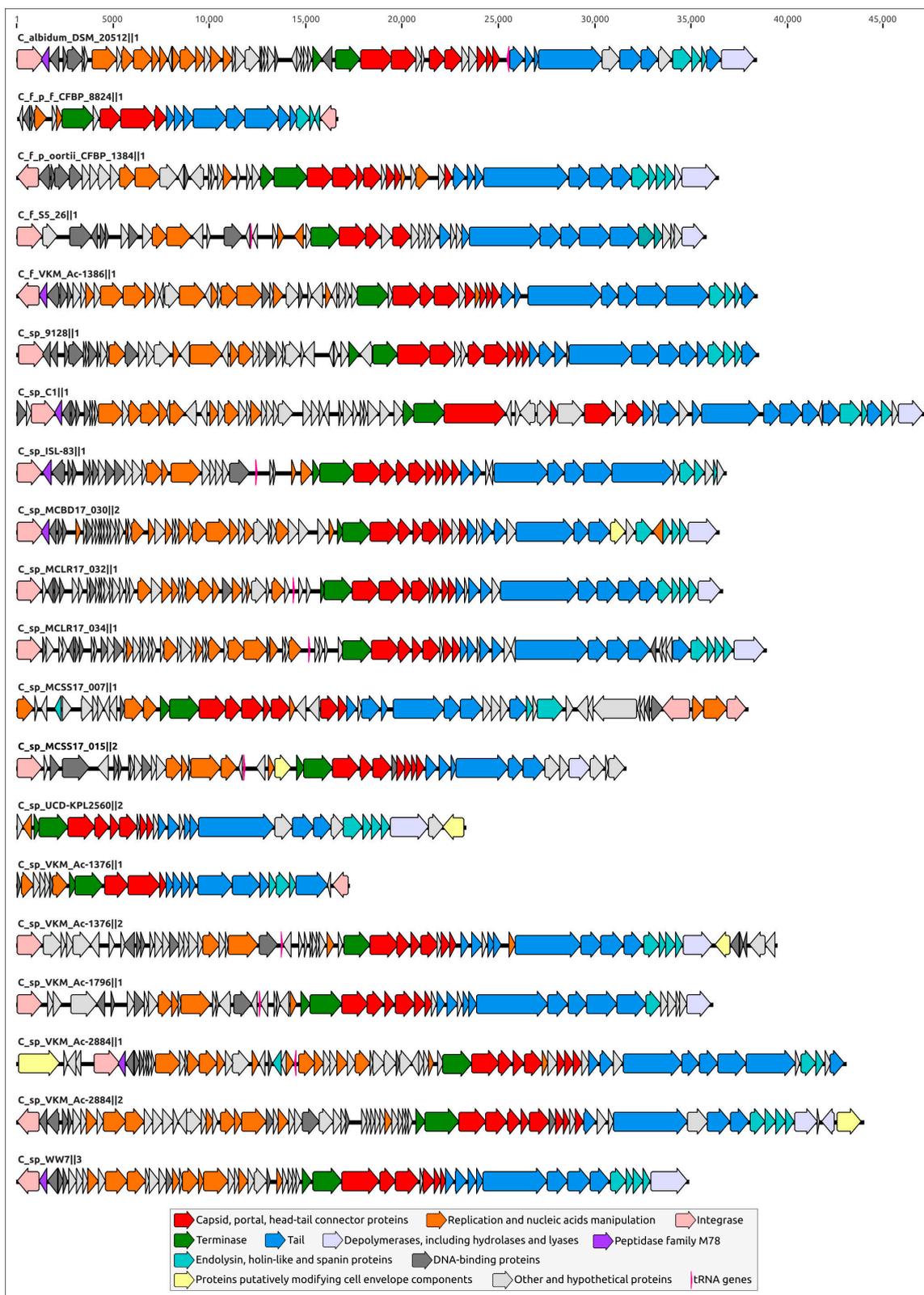
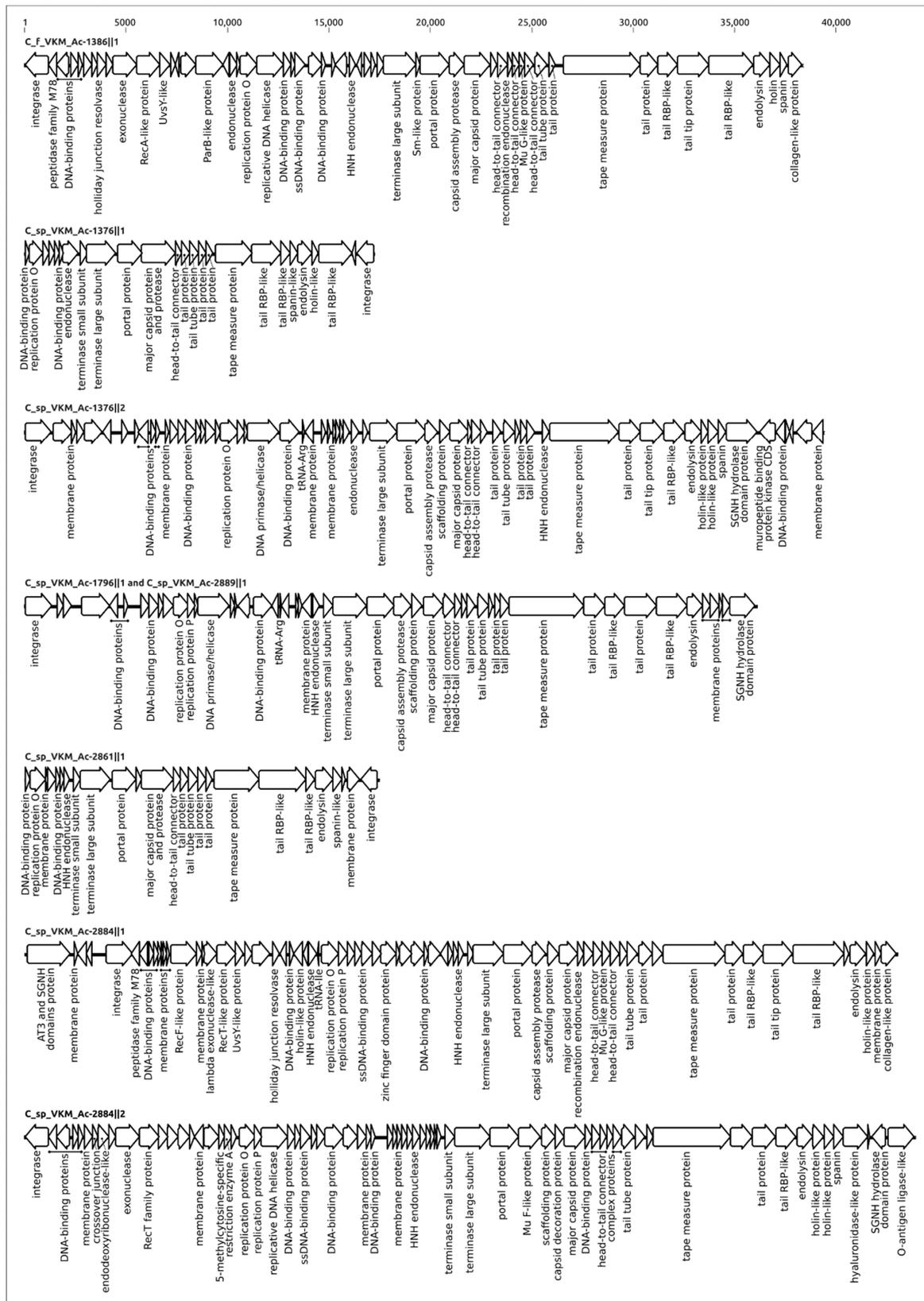


Figure 4. Best-scoring ML phylogenetic tree constructed with 164 amino acid sequences of major capsid protein. The NCBI taxonomy is shown to the right of the phage name. The numbers near the tree branches indicate the TBE values. The total number of bootstrap trees was 1000. The scale bar shows 0.5 estimated substitutions per site and the tree was rooted to C_sp_C1|1. Different background colours and numbers indicate the clustering proposed for analysis.



(a)

Figure 5. Cont.



(b)

Figure 5. (a) Simplified genetic maps of the representatives of each putative prophage cluster shown in Figure 4. (b) Genetic maps of seven putative prophages predicted to be contained in *Curtobacterium* genomes of VKM collection strains. Arrows indicate the direction of transcription.

The replication apparatus of all predicted prophages included the protein similar to the λ replication protein O required for initiation of DNA replication and present in other temperate phages [42,47]. Larger genomes also contained genes similar to another phage λ initiator replication protein P [47]. Most genomes contained DNA primase/helicase genes, as in some lambdoid phages [48]. No DNA polymerase (DNAP) genes were found in the curated PDRs, even though it seemed that, in a few cases, regions of apparently bacterial origin, adjacent to prophages, contained DNAP genes, basically encoding bacterial DNA polymerase III subunits.

Genomic regions of all predicted prophages, not trimmed by contig limits, comprised all other functional modules featuring the tailed phages, including structural and lysis modules. All predicted major capsid proteins featured HK97 fold, typical for *Heunggongvirae* viruses, including tailed bacteriophages [49]. HHpred analysis indicated similarities among the analysed structural proteins and structural proteins of transposable phage Mu [50,51], including Mu F-like and Mu G-like putative capsid assembly proteins (Figure 5b). The lysis cassettes were predicted to encode a three-step lysis system [52] containing the endolysin, holin and spanin genes. The genomes of putative prophages assigned to clusters 1, 4–10, 12 and 17 can contain two adjacent membrane holin-like proteins, possibly possessing the holin–antiholin system, where antiholin controls the timing of host cell lysis by inhibiting holin [53,54]. Most PRDs' genomes have genes encoding depolymerases of different functionality, as discussed below. They can be released during the lytic stage, facilitating phage escape [55].

The tail modules of analysed RDRs vary in size and complexity and can comprise up to ten or more genes, including the head–tail connector complex genes. Several proteins showed structural similarity to known tail spike proteins and were predicted to contain depolymerase domains. Interestingly, some PDRs can contain the genes that can modify cell envelope components (Figure 5a).

2.4. Taxonomy of Related Phages

To define closely related phage taxonomic groups, orthoANI [56] and VIRIDIC tools were used to compare nucleotide sequences of viral genomes. The first step of the analysis included calculations of average nucleotide identity with orthoANI using 20 putative prophage sequences representing 20 clusters (shown in Figure 4) and all 13,477 complete phage genomes available in the NCBI Genome database as of July 2022. Next, 20 putative prophages and the related genomes, found with the orthoANI calculation and having the highest ANI values and average aligned length, were used to estimate intergenomic similarity, using the VIRIDIC tool.

Neither ANI calculations (Supplementary Table S1), nor a VIRIDIC comparison matrix (Supplementary Figure S4), revealed meaningful similarities between *Curtobacterium* PDRs and known complete phage genomes. However, a small likeness of the order of 10% has been detected with some phages infecting *Microbacterium* bacteria, which is phylogenetically close to the genus *Curtobacterium*. In particular, small *Microbacterium* phages with a genome size under 20,000 bp [57], recently assigned to newly established taxa (family *Orlajensenviridae*, subfamily *Pelczarvirinae*, genus *Paopuvirus*) [58], were shown to share distant intergenomic similarities with putative *Curtobacterium* prophages, with a similar genome size, assigned to Group 13.

Phylogenetic analysis using amino acid sequences of conservative proteins can reveal more distant evolutionary and taxonomic relationships [59]. The list of related phage taxonomic groups exposed using the MCP and TerL phylogenies (Figure 4, Supplementary Figure S1) basically includes both unclassified and classified temperate actinophages of siphoviral morphology assigned to subfamily *Nclasvirinae* and several dozen genera not assigned to any subfamily or family, including *Bridgettevirus*, *Britbratovirus*, *Bronvirus*, *Coralvirus*, *Decurrovirus*, *Fromanvirus*, *Mapvirus*, *Timquatrovirus*, etc. The related phages infect bacteria belonging to genera *Arthrobacter*, *Attivirus*, *Bifidobacterium*, *Corynebacterium*, *Gordonia*, *Microbacterium*, *Mycobacterium*, *Streptomyces*, *Propionibacterium*, *Rathayibacter* and *Rhodococcus*.

CRISPR spacers represent biological records of past phage–bacteria interactions, and can be used for finding related hosts and phages [60,61]. A CRISPR search found CRISPR repeat regions in only 11 genomes of the 197 *Curtobacterium* genomes analysed, namely in strains *Curtobacterium flaccumcaciens* pv. *flaccumfaciens* (*C. fpf*) BRIP_70615, *C. fpf* BRIP 70615, *C. fpf* CFBP 3417, *C. luteum* JCM 1480, *C. luteum* ATCC 15830, *C. luteum* DSM 20542, *C. sp.* 9128 DE0339, *C. sp.* MCB17_028, *C. sp.* MCP17_018, *C. sp.* MCSS17_007 and *C. sp.* WW7. The total number of spacers was 72. BLAST, using the spacers' sequences, found similar regions, mainly in the genomes of actinophages of siphoviral morphology, including phages belonging to subfamilies *Arquatrovirinae*, *Bclavirinae*, *Guernseyvirinae*, *Mclavirinae*, *Nclavirinae*, *Nymbaxtervirinae*, *Weiservirinae*, unclassified phages and different genera not assigned to subfamilies or families.

2.5. Prophage Induction

Six strains of *Curtobacterium* sp. (VKM Ac-2098, VKM Ac-2884, VKM Ac-2861, VKM Ac-1796, VKM Ac-1376 and VKM Ac-2889) were assessed for the presence of inducible prophages through induction using different mitomycin C concentrations, as described. Bioinformatic analysis suggests the presence of prophages integrated into genomes of these strains, and strains were available. The prophage-free *Curtobacterium* strain CFBP 3418 was used as a control for the induction experiments.

It was shown that applying filtrates from bacterial cultures of *Curtobacterium* strains VKM Ac-2098, VKM Ac-2884, VKM Ac-2861, VKM Ac-1796, VKM Ac-1376 and VKM Ac-2889 that were treated with mitomycin C to a final concentration of at least 1 µg/mL resulted in the formation of a lysis zone on the bacterial lawns of all tested *Curtobacterium* sp. strains, except for CFBP 3418 (Supplementary Figure S5). Furthermore, the addition of mitomycin C at a concentration of 1 µg/mL did not lead to growth inhibition of the prophage-free strain CFBP 3418. Thus, this concentration of mitomycin C was chosen as the optimal concentration for prophage induction from bacterial cultures of *Curtobacterium* strains.

The siphoviral morphology of phage particles induced from the bacterial culture of the *Curtobacterium* strain VKM Ac-2884 was revealed using transmission electron microscopy (TEM) (Figure 6). The induced prophage was characterised with a flexible ~190-nm-long tail and an isometric capsid with a diameter of ~60 nm.

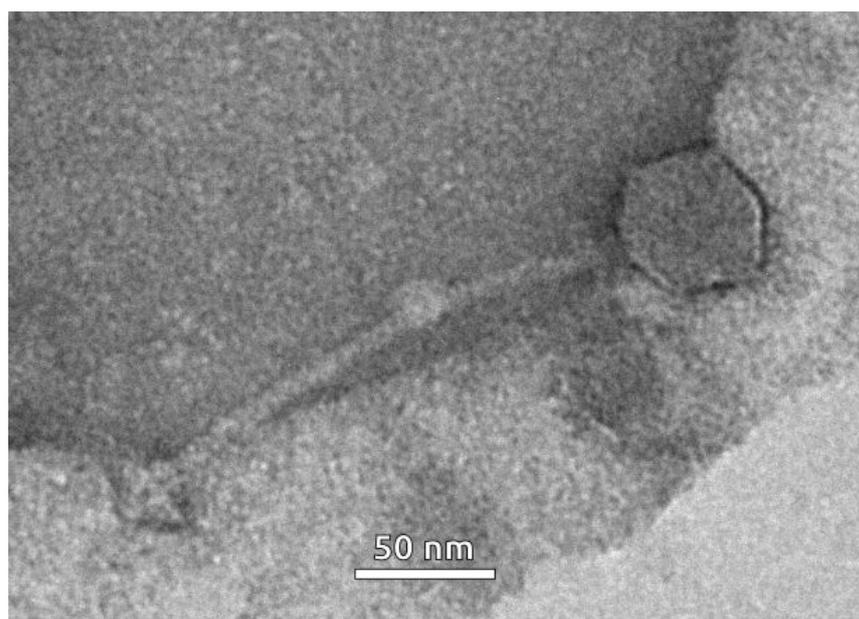


Figure 6. Electron microscopy image of phage particle induced by mitomycin C from the bacterial culture of *Curtobacterium* sp. VKM Ac-2884. The scale bar is 50 nm.

The genome of the strain *Curtobacterium* sp. VKM Ac-2884 was predicted to contain two prophages. Both of them were siphoviruses, and it is impossible to distinguish between them using TEM imaging. A PCR analysis of total DNA isolated from concentrated phage particles after induction was conducted. Amplification was observed only with a set of primers constructed for the detection of phage **C_sp_VKM_Ac-2884||2** (Supplementary Figure S6).

A phylogenetic analysis using the major capsid protein (Figure 4) indicated the relatedness of phage **C_sp_VKM_Ac-2884||2** to unclassified phages *Mycobacterium* prophGD12-2, *Mycobacterium* prophGD05-1, *Streptomyces* SF1 and *Streptomyces* SF1. The closest classified *Gordonia* phage BritBrat (*Britbratovirus britbrat*) belongs to the *Britbratovirus* genus not assigned to a subfamily or family. The VIRIDIC Intergenomic Distance Calculator failed to indicate any meaningful intergenomic nucleotide similarity between the induced prophage and the related phages listed above. ANI calculations using all phage sequences deposited in NCBI GenBank also failed to find closely related phages with any meaningful average nucleotide identity and coverage. Thus, the induced prophage can represent a new viral genus or a higher-ranked taxon.

2.6. Analysis of Phage Endolysins Encoded in PDRs

A search for peptidoglycan hydrolase (lysin) genes in the predicted prophage regions indicated the presence of homologues of phage lysins in all PDRs that were not trimmed by contig borders. Fifty-eight lysins found by the search (Supplementary File S2) were clustered using an ML phylogenetic analysis (Figure 7). An HHpred analysis indicated a similar structure and domain architecture within the clusters.

A domain architecture and putative functional assignments of proteins and domains have been suggested using the results of HHpred and InterProScan [62] searches, and clarified using an analysis of the results of protein structural modelling (Figure 8).

Cluster 1 is represented by a single 444 amino acid residue (aa)-long multidomain endolysin. HHpred HMM-HMM comparisons showed the closeness of the N-terminal domain of this protein (approximately 1–160 aa) to lysins belonging to the γ -glutamyl D,L-endopeptidase (NlpC/P60) family [63]. This domain can be involved with enzymatic activity.

Predicted Domain 2 (161–270 aa) contains putative amino acid residues forming the substrate entrance channel groove [63] and is proposed to be essential for substrate recognition. Domain 3 was not predicted as a compact structure by either AlphaFold 2 [64] or RoseTTAFold [65]. Hypothetically, Domain 3 can facilitate the folding of functional prophage lysin. Domain 4 was predicted to contain transmembrane regions and was modelled to include three α -helices. This domain can assist translocation across the membrane into the peptidoglycan.

Endolysins assigned to **Cluster 2** include eight proteins featuring the two-domain structure with the CHAP catalytic domain putatively arranged in the N-terminal part. It is impossible to predict the catalytic function of these enzymes (amidase or endopeptidase) confidently.

The two endolysins assigned to **Cluster 3** are similar to Cluster 2 lysins. The N-terminal catalytic domain has been proposed to function as N-acetylmuramoyl-L-alanine amidase [66], and the C-terminal domain has appeared to be responsible for peptidoglycan binding [66].

Cluster 4 contains a single 372-aa-long endolysin with a predicted three-domain architecture. The N-terminal domain is similar to several lytic enzymes of Gram-positive bacteria corresponding to the lysozyme family with muramidase activity cleaving the β -1,4 glycosidic bonds in the backbone of peptidoglycan.

Central and C-terminal domains can facilitate peptidoglycan binding.

Cluster 5 contains endolysins putatively encoded in 11 PDRs. The N-terminal domains showed a closeness to D-Ala-D-Ala endopeptidases. Structural comparisons of **C_f_VKM_Ac-1386||1** and *Enterococcus* phage IMEEF1 (Supplementary Figure S7a) showed overall similarity not only to the N-terminal domain, but also to the C-terminal peptidoglycan-binding domain.

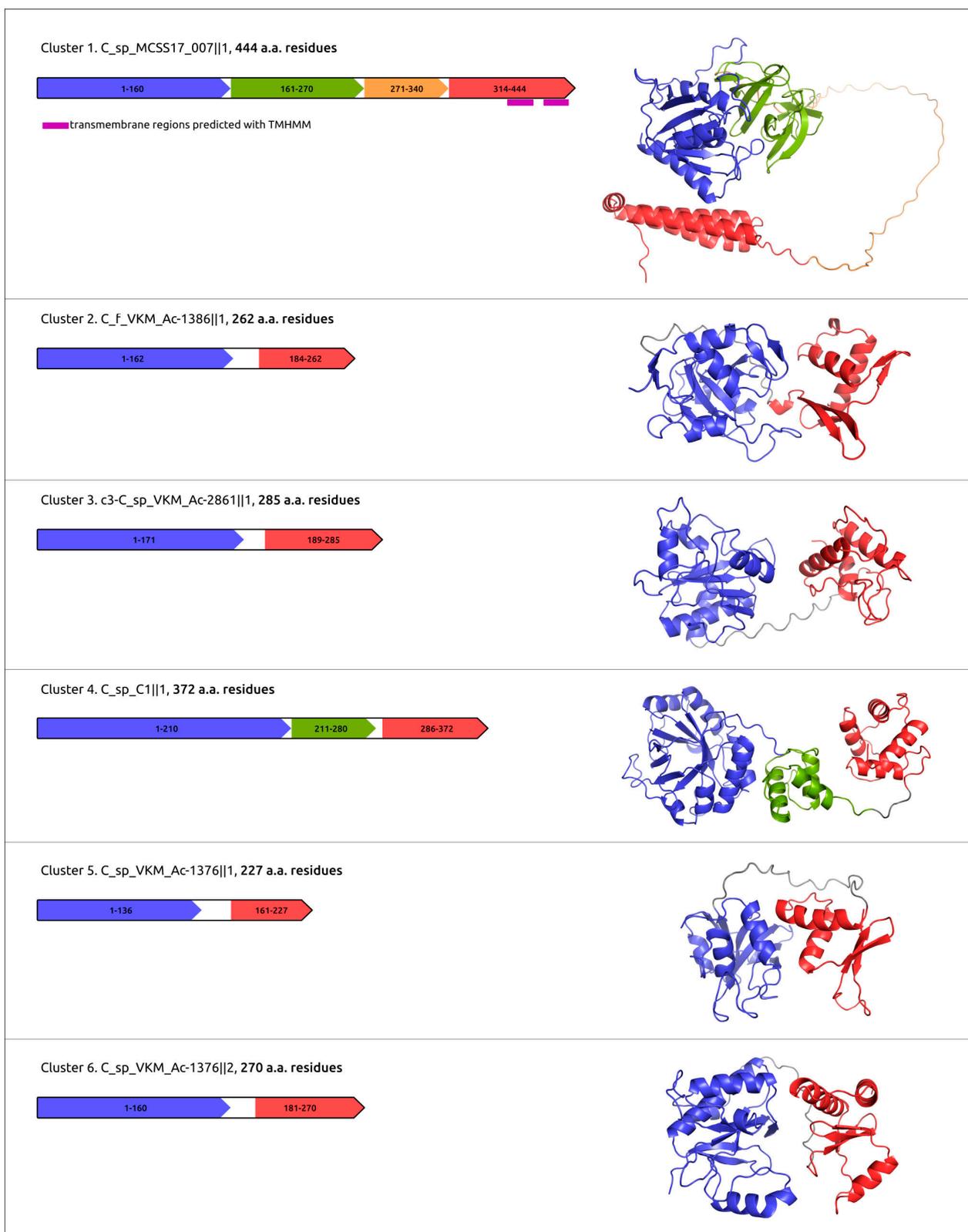


Figure 8. Domain architecture and AlphaFold models of putative endolysins encoded in the putative prophage regions.

Cluster 6, the largest cluster, comprises endolysins putatively encoded in 35 PDRs. They are predicted to contain two structural domains. The N-terminal domain is close to the D,D-dipeptidase/D,D-carboxypeptidase proteins. Using a protein structure comparison,

Dali [67] indicated the likeness of C-terminal domains of Cluster 6 proteins to the cell-binding domain of *Enterococcus* phage ϕ M1EF22 lysin (RMSD 3.0, PDB code 7D55) [68] (Supplementary Figure S7b).

2.7. Analysis of Other Glycopolymer-Degrading Enzymes Encoded in PDRs

Besides endolysins (peptidoglycan-degrading enzymes), the PDRs analysed contained other genes of putative glycopolymer-degrading enzymes (which, for the purposes of this work, will be referred to as depolymerases, DPOs). These genes are regularly located downstream of the lysis module and can be part of both the lysis system and the penetration apparatus. They were present in the most common putative *Curtobacterium* prophages, in at least 45 of the PDRs analysed (Supplementary File S3). The results of a sequence search and structural analysis indicated a great diversity within the DPOs' functions and structures (Figure 9, Supplementary Figure S8). In some cases, the phylogenetic analysis using the DPO amino acid sequences resulted in low bootstraps and arranged in common branches the sequences representing the proteins with different enzymatic activity; therefore, clustering using structural similarity was used (Figure 10). This clustering method showed better consistency with the putative functional assignment of DPOs.

Cluster 1 contains only one α/β -hydrolase (572 aa). The catalytic domain is similar to carboxyl esterase from the oil-degrading bacterium *Oleispira antarctica* (HHpred probability 99.76%, PDB code 3I6Y) [69] and other hydrolases including family S9 peptidases. The catalytic triad Ser⁴³⁷-His⁵⁴⁹-Asp⁵¹⁸ can be easily detected with HHpred and structural alignment using the AF2 model. An enzymatic domain is located in the C-terminal part. The superimposition of **C_albidum_DSM_20512||1** DPO with carboxyl esterase 3I6Y showed RMSD 2.8 Å (Supplementary Figure S9a).

Cluster 2 includes two DPOs where an HHpred search indicated a similarity with N-acetylglucosamine-1-phosphodiester α -N-acetylglucosaminidases (NAGPA), which removed the terminal GlcNAc residues [70]. Both DPOs contain three domains including N-terminal α -helical domain and central catalytic domain. The superimposition of **C_luteum_NS184||1** (427 aa) with NAGPA 6PKU shows RMSD 2.4 Å (Supplementary Figure S9b).

Cluster 3 comprises two DPOs that are similar to several polysaccharide lyases. Comparison of **C_sp_Ferrero||1** (675 aa) with the alginate lyase from *Defluviitalea phaphyphila* (PDB code 6JP4) [71] shows a topology that is similar to the C-terminal α/α -barrel domain of 6JP4 and the putative prophage DPO (RMSD 5.9 Å) (Supplementary Figure S12c).

Putative depolymerase domain-containing proteins assigned to **Cluster 4** and **Cluster 5** demonstrated a structural architecture typical of tail fibre (spike) proteins [72,73], including those found in prophage regions [74]. These proteins contained a parallel β -structured pyramidal central part, formed upon trimerisation (Supplementary Figure S10a). The function of such proteins may be assigned as being hyaluronidase, pectate lyase or other enzymes, including the enzymes degrading cell wall components via a lyase mechanism. Analysis of the structures of Cluster 4 DPOs indicated the presence of Asp and Tyr residues located similarly to the well-studied streptococcal phage-encoded hyaluronidase HylP1 [75] (Supplementary Figure S10b). The proteins assigned to Cluster 5 show a similarity with different phage tail spike proteins (TSPs) involved in the degradation of polysaccharides [76,77].

Cluster 6 contains only one DPO showing a strong similarity with α -L-fucosidase from *Paenibacillus thiaminolyticus* (Supplementary Figure S11a) (HHpred probability about 100.00%, PDB code 6GN6) [78]. This glycoside hydrolase, belonging to the family GH29, cleaves L-fucose from oligosaccharide non-reducing termini.

Cluster 7 also consists of a single depolymerase similar to GDSL/SGNH hydrolase from *Bacteroides thetaiotaomicron* (probability 99.84%, PDB code 7BR2) [79]. This protein has been suggested to function as an oligosaccharide deacetylase [79]. The topology of the predicated prophage DPO is similar (Supplementary Figure S11b), comprising the N-terminal catalytic domain and C-terminal binding part.

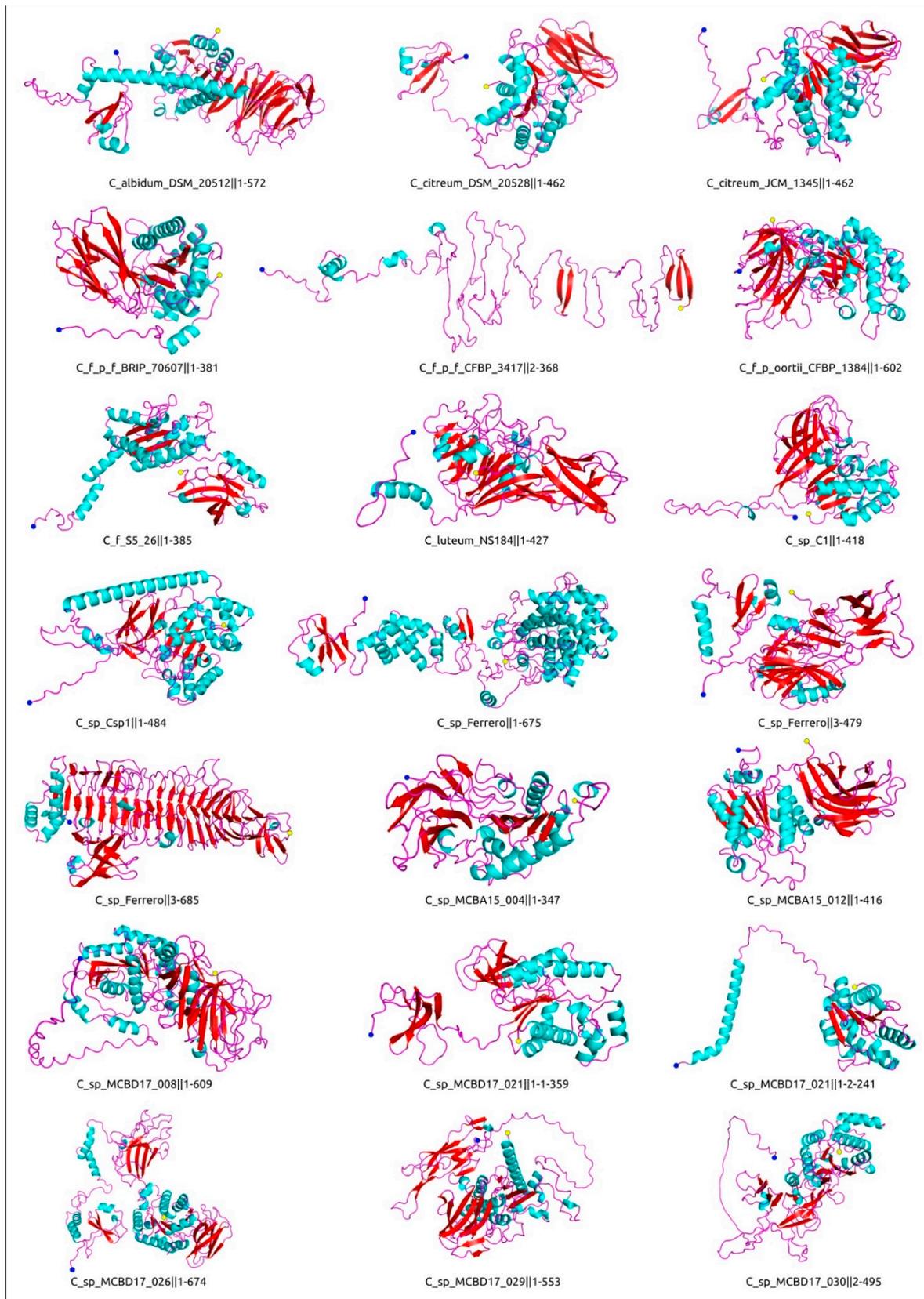


Figure 9. The 21 structural models of depolymerases encoded in the predicted *Curtobacterium* PDRs, coloured by type of secondary structure. The N-terminus is labelled with a blue circle and the C-terminus is labelled with a yellow circle. The remaining models are shown in Supplementary Figure S8.

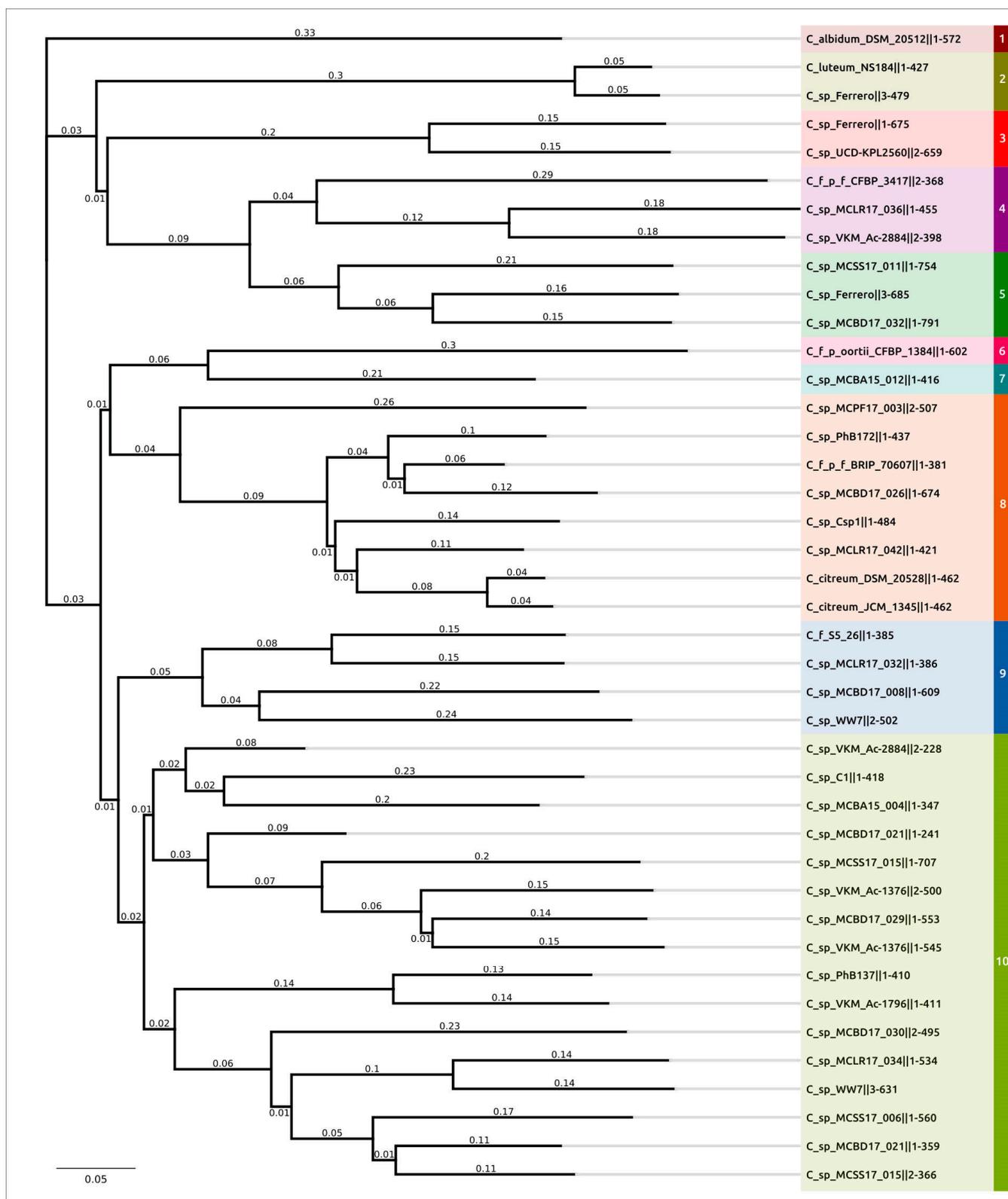


Figure 10. Phylogenetic tree based on amino acid sequences of glycopolymer-degrading enzymes using structural similarity assessed with TM-scores. Branches belonging to different clusters are highlighted using colour. The last three digits in the names of structures indicate the length of protein sequences.

Proteins assigned to Clusters 8, 9 and 10 can also function as polysaccharide deacetylases. **Cluster 8** DPOs show a similarity with the GDSL/SGNH-like lipase/acyl hydrolase family protein from *Neisseria meningitidis* (HHpred probability 99.53%, PDB code 4K7J) (Supplementary Figure S11c). Presumably, the protein functions as an oligosaccharide deacetylase. The central β -barrel domain can function as a carbohydrate-binding part.

Cluster 9 depolymerases share a similarity with the acyl hydrolase family protein from *Parabacteroides merdae* (PDB code 4Q9A) (Supplementary Figure S11d). The domain architectures of cluster 9 DPOs are basically similar. The structures have three parts: the N-terminal part of variable size includes α -helices, the central part contains the catalytic domain and the C-terminus contains β -strands.

Cluster 10 comprises proteins with a GDSL/SGNH hydrolase domain and versatile structural organisation. Members of this cluster have a high degree of structural similarity with a group of structurally related proteins, which belong to the SGNH-hydrolase superfamily involved in carbohydrate metabolism and polysaccharide degradation, and which can function as carbohydrate deacetylases.

Most predicted structures contain two domains, where the enzymatic domain is located in the C-terminal part of the molecule; they include the DPOs from *C_sp_MCJR17_043*||1 (553 aa) and identical sequences from the PDRs of other *Curtobacterium* strains *C_sp_VKM_Ac-1376*||1 (545 aa) (Supplementary Figure S12a), *C_sp_VKM_Ac-1376*||2 (500 aa) (Supplementary Figure S12b), *C_sp_MCBD17_021*||1 (241 aa), *C_sp_MCBA15_004*||1 (347 aa), *C_sp_C1*||1 (418 aa), *C_sp_VKM_Ac-2884*||2 (228 aa) (Supplementary Figure S12c), *C_sp_MCSS17_015*||2 (366 aa), *C_sp_MCBD17_021*||1 (359 aa) and *C_sp_MCBD17_030*||2 (495 aa). Several depolymerases, including those from *C_sp_PhB137*||1 (410 aa), *C_sp_VKM_Ac-1796*||1 (411 aa) (and identical *C_sp_VKM_Ac-2889*||1), *C_sp_MCLR17_034*||1 (534 aa) and *C_sp_MCSS17_006*||1 (560 aa), have enzymatic domains arranged in the N-terminus.

The C-terminal domains of *C_sp_PhB137*||1 (410 aa) and *C_sp_VKM_Ac-1796*||1 (411 aa) (Supplementary Figure S12d) have similarity with the carbohydrate-binding domain of endo-1,4-beta-xylanase C (PDB code 4XUP). Predicted structures of depolymerase *C_sp_MCSS17_015*||1 (707 aa) (Supplementary Figure S12e) and *C_sp_WW7*||3 (631 aa) feature a more complicated multidomain architecture, where the catalytic domain is located after the N-terminal domain and is attached to the β -barrel subdomain, which in turn is followed by another β -barrel domain. According to the results of an HHpred search, in the case of *C_sp_MCSS17_015*||1 (707 aa), the latter domain can play the role of the additional sugar-binding domain, as in a structurally similar sugar-binding protein (PDB code 4AVS) (Supplementary Figure S12e).

Interestingly, the N-terminal domains of Cluster 10 proteins vary in size and content. Some predicted structures, such as *C_sp_VKM_Ac-1376*||1 (545 aa) and *C_sp_MCSS17_015*||1 (707 aa), contain N-terminal parts composed of β -strands. Hypothetically, such domains could enhance substrate binding. In several models, such as *C_sp_VKM_Ac-1376*||2 (500 aa), the enzymatic domain is arranged between the upstream and downstream sequences, which in turn assemble a compact β -barrel structure, reminiscent of the topology of DPOs assigned to Cluster 8.

3. Discussion

Current information on bacteriophages infecting *Curtobacterium* sp. is very sparse. Numerous attempts to isolate lytic phages using traditional techniques [80,81] have resulted in the discovery of just three distinct phage types with prevailing ϕ 29-like podoviruses. The dominance, in the environment, of one, or few, genera of phages infective to a certain bacterial host species has been reported previously. Recent examples relevant to agriculture are the prevalence of *Limestoneviruses* among phages infecting potato pathogen *Dickeya solani* [82,83], and *Ficleduoviruses* among phages of aquaculture pathogen *Flavobacterium columnare* [84]. The accumulation of statistically robust data on available phage diversity either takes decades (as for *E. coli* or *Pseudomonas* sp.), or needs a concerted effort from numerous researchers (as for the SEA-PHAGE programme, studying phages of *Mycobacterium* sp.) [85,86].

A complementary approach is to assess the potential of temperate phages of the target bacteria, including inducible prophages encoded in host genomes. Generally, it is advised that temperate phages are avoided in phage therapy applications. However, when appropriate lytic phages are missing, or temperate phages have unique features, using the latter can be considered [87]. Besides selecting natural *vir* mutants with reduced lysogenic ability, it is possible to improve the behaviour of phages using gene editing approaches [88], or employing recombinant phage-derived enzymes with cytolytic properties [89,90].

Examination of the results of prophage prediction using genomes of different taxonomic groups indicated fewer predicted prophages in *Curtobacterium* bacteria than in most of the other analysed taxa, except for the genus *Clavibacter*, another member of the *Microbacteriaceae* family. This observation is interesting in light of the fact that relatively few *Curtobacterium* strains, (11 of 197 analysed genomes), contain CRISPR-Cas adaptive immune system regions in the search results, while previous research studies have estimated that 50% of sequenced bacterial genomes contain CRISPR [91,92]. It might be suggested that *Curtobacteria* have other effective antiphage defence mechanisms. An investigation of regions related to mobile elements could provide answers to these questions. It is noteworthy that a significant part of the PHASTER prediction results related to genomic regions containing the genes of cell-envelope-modifying enzymes. The cell walls of some *Curtobacterium* strains were shown to contain different glycopolymers, particularly rhamnan, and cell wall hydrolysates contained rhamnose, mannose and other saccharides [93]. Interestingly, the depolymerases assigned to Cluster 5 contained enzymatic domains similar to rhamnolacturonase and endo-xylogalacturonan.

An analysis of cell-envelope-degrading enzymes of prophage origin might provide insights into the phage resistance mechanisms of *Curtobacteria*. Several prophage depolymerases have been predicted to possess hyaluronidase enzymatic activity. Such proteins have been found in various Gram-positive bacteria, playing an important role in spread and growth [94]. Some phages infecting Gram-positive bacteria use hyaluronidase to break the hyaluronic acid capsule to penetrate the host cell [95,96].

Phages use bacterial receptors to adsorb to the host cell surface. Common cell receptors of Gram-positive bacteria used by phages include murein, cell wall teichoic acids and lipoteichoic acids [97,98]. Bacteria often use modification of the receptors to resist phage infection [99,100]. As part of the host–parasite ‘arms race’, phages evolve to counter the defensive mechanisms of bacteria. Several predicted and modelled phage depolymerases, including the phosphodiester α -N-acetylglucosaminidase (NAGPA) and α -L-fucosidase, can participate in the removal of cell receptors’ modifications, therefore preventing penetration into the host cell.

Most predicted depolymerase could be involved in peptidoglycan (PG) or polysaccharide deacetylation (Figure 11) [55,101–103]. O-acetylation of PG occurs at the C-6 hydroxyl of N-acetylmuramoyl residues and sterically blocks the activity of lysozymes [101]. O-acetylation of the capsular polysaccharide is important for bacteria and can achieve polysaccharide rigidity [104,105]. A wide variety of phage cell-envelope-degrading enzymes can indicate their importance as phage counter-defence mechanisms. This, together with the low number of prophages, enables the hypothesis that *Curtobacterium* high phage resistance may be associated with cell wall characteristics. This hypothesis needs further detailed study.

Phage endolysins encoded in predicted prophage-derived regions were represented by several groups showing different types of enzymatic activity, but most of the predicted lysins appeared to exhibit D,D-dipeptidase activity (Figure 11). Predicted structures of predicted endolysins were typical for Gram-positive bacteria, featuring a modular architecture that included at least two domains: catalytic domains and binding domains [106–108]. One endolysin, putative γ -D-glutamyl-L-diamino acid endopeptidases from *C_sp_MCSS17_007*||1, was predicted to have a four-domain architecture, while another endolysin, putative GH25 family muramidase from *C_sp_C1*||1, was modelled as a three-domain structure. Interestingly, regardless of the number of domains, the catalytic

domain was located in the N-terminal part of all proteins. A pronounced modular architecture of endolysins, together with a high level of accuracy of structure predictions using modern AI software, might be used for the design of chimeric proteins that are effective against *Curtobacterium* infections.

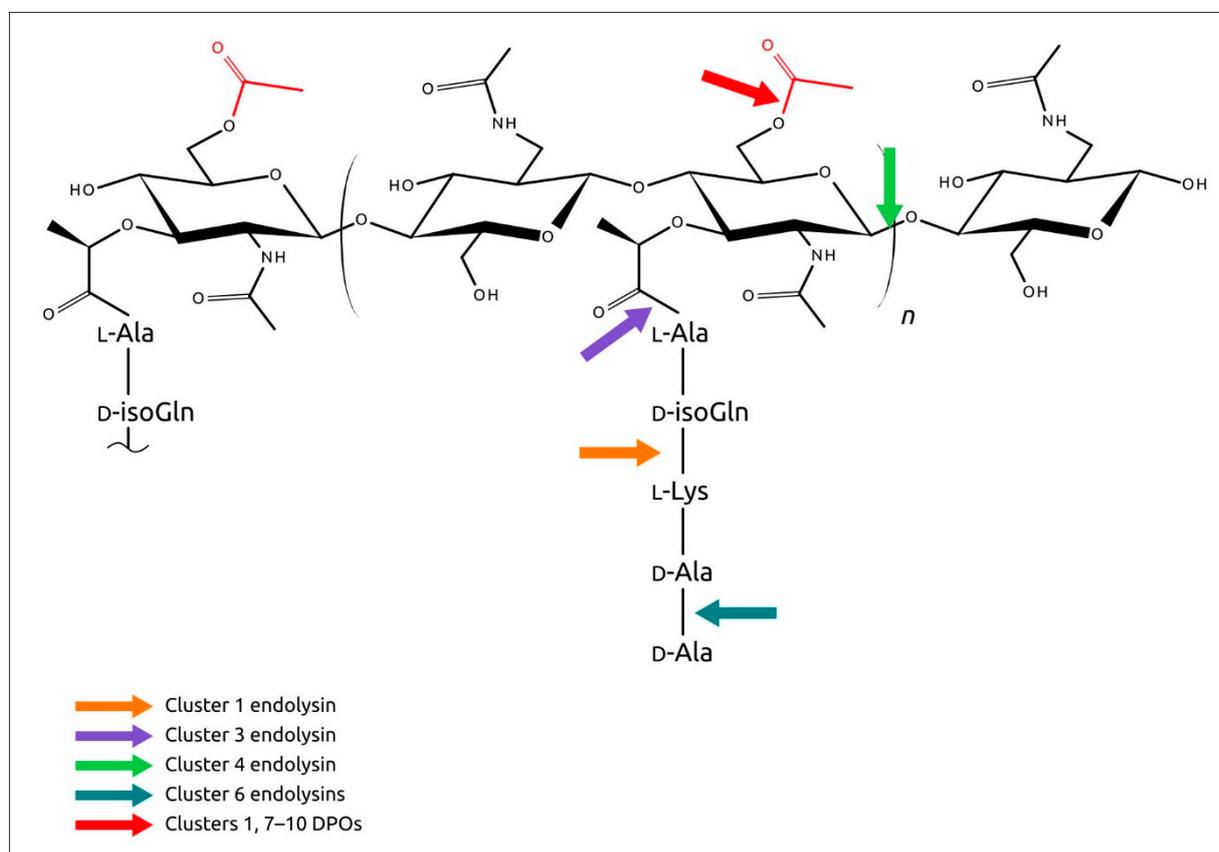


Figure 11. Possible sites for peptidoglycan cleavage, as predicted by cell-wall-degrading enzymes, indicated here with coloured arrows. The O-acetylation of the C-6 OH groups of MurNAc residues is indicated in red.

4. Materials and Methods

4.1. Search for Prophage-Derived Sequences

The *Curtobacterium* genomes were downloaded from the NCBI Genome Database [26] and re-annotated using the Prokka pipeline [109], with default settings. The search for possible prophage-derived regions in the genomes was conducted using the PhiSpy pipeline [28] and the PHASTER server [27]. The PhiSpy calculations and PHASTER searches were performed with default settings.

4.2. Prophage Annotation

Predicted prophage sequences were extracted using the Geneious Prime 2022.2.1 tools [110] and annotated using Prokka, the HHpred server [32], Phyre 2 [111] and a BLAST search of the NCBI non-redundant (nr/nt) database, as well as a BLAST search of custom databases using GenBank phage sequences. The Prokka settings included using a custom BLAST [31] database built with functionally annotated phage protein sequences extracted from annually annotated PDRs and NCBI GenBank RefSeq database [112] phage genome sequences. The HHpred search results were obtained using the PDB_mmCIF070, SCOPEe70, Pfam-A_v35 and UniProt-SwissProt-viral70 databases. Transmembrane regions were predicted using HHpred and TMHMM [113]. The genetic maps were visualised using Geneious Prime 2022.2.1. Comparisons between the predicted prophage regions were

performed and visualised using Easyfig [114], applying the TBLASTX [31] algorithm for the estimation of similarities among genomic loci.

4.3. Genomic and Phylogenetic Analysis

Average nucleotide identity was calculated using orthoANI, with default settings [56]. The pairwise nucleotide similarities among the predicted prophage-derived sequences and corresponding similarity matrix were computed using orthoANIu [56] and the Virus Intergenomic Distance Calculator (VIRIDIC) pipeline, with default settings being used [39]. The search for the CRISPR regions was conducted using the MinCED programme [115]. Protein sequence alignments were obtained using Clustal Omega [116] with [number of refinement iterations = 10, evaluate a full distance matrix for initial guide tree and full distance matrix = yes, cluster size for mBed guide trees = 100] settings. The terminase phylogenetic tree was constructed using the RAxML-NG [117] built-in raxmlGUI 2.0.9 graphical user interface [118], using the BLOSUM62+ F+I+G amino acid substitution model [119] and [-bs-metric tbe -tree rand{10} -bs-trees 1000] settings. The best amino acid substitution models were estimated using ModelTest-NG [120]. Robustness of the tree was assessed using a bootstrap analysis employing ten starting trees and 1000 bootstrap replicants, before calculating the transfer bootstrap expectation values. The resulting tree was visualised using the iTOL v6 server [121].

4.4. Prophage Induction Assay with Mitomycin C

The induction of prophages was performed as previously described [60], with modifications. Briefly, single colonies of *Curtobacterium* sp. strains VKM Ac-2098, VKM Ac-2884, VKM Ac-2861, VKM Ac-1796, VKM Ac-1376, VKM Ac-2889 and CFBP 3418 were picked from YD-agar (20 g dextrose, 10 g yeast extract, 20 g agar, distilled water up to 1 litre) plates, dropped in tubes containing 10 mL YD-broth (20 g dextrose, 10 g yeast extract, distilled water up to 1 litre) and left to grow overnight at 27 °C in personal bioreactor RTS-1C (Biosan, Riga, Latvia). Overnight bacterial cultures were diluted with 25 mL of fresh YD-broth to OD₆₀₀ of approximately 0,09 and then incubated at 27 °C, with shaking, at 300 rpm for 7 h to obtain a final OD₆₀₀ of 0.25. Then, several aliquots of these bacterial cultures were treated with different concentrations of mitomycin C (0.3 µg/mL, 0.5 µg/mL, 1 µg/mL, 3 µg/mL, 5 µg/mL and 7 µg/mL) or left without mitomycin C, as a control, and incubated under the same conditions for 22 h. After incubation, the samples were centrifuged at 7000 G for 20 min and then passed through 0.45-µm sterile membranes. The resulting filtrates were stored at 4 °C.

The presence of induced phages in the filtrates was tested against the same *Curtobacterium* sp. strains (VKM Ac-2098, VKM Ac-2884, VKM Ac-2861, VKM Ac-1796, VKM Ac-1376, VKM Ac-2889 and CFBP 3418), using the double-layer method [122]. For this, 200 µL of *Curtobacterium* bacterial cultures grown in YD-broth at 27 °C (~10⁸ CFU per mL) was mixed with 3 mL of soft agar (YD-broth supplemented with 0.7% agar). The mixtures were plated onto the YD-agar. Then, 10 µL of each filtrate was spotted on the soft agar lawns and incubated at 27 °C for 18 to 24 h.

4.5. Electron Microscopy

To obtain preparations for microscopy, 100 mL of host culture was grown and prophage was induced, as described above. The resulting lysate was then concentrated and purified, according to the protocol described by Ackermann [123]. Centrifugation with ammonium acetate was carried out twice. Concentrated purified samples were placed on grids and stained with 1% aqueous uranyl acetate (pH 4.0). Prepared grids were examined using a JEM-2100 200 kV transmission electron microscope (JEOL, Tokyo, Japan).

4.6. PCR Analysis

PCR primers were constructed with Primer3 2.3.7 [124], using predicted sequences of major capsid proteins. 1-144F (CACCTTCAACGACATCCCCA) and 1-423R (GTAGTTGTC-

CCAGCCGTTGA) were selected to identify the phage **C_sp_VKM_Ac-2884||1** (280 bp product). Primers 2-119F (CGTCGCTGTCGTTCAACTTC) and 2-453R (GAAGTCGATCGTCGCCTTGA) were selected to identify the phage **C_sp_VKM_Ac-2884||2** (335 bp product). 5× ScreenMix (Evrogen, Russia) was used for PCR. Each 25 µL reaction contained 5 µL of ScreenMix, 0.3 µM of each primer and 25 ng of DNA, and the volume was adjusted using sterile Milli-Q water. Thermal cycling conditions were as follows: 94 °C for 3 min, followed by 34 cycles of melting at 94 °C for 30 s, annealing at 60 °C for 30 s, elongation at 30 °C for 30 s and finally incubation at 72 °C for 3 min. As a negative control, a reaction with the addition of an appropriate volume of water was used instead of DNA. PCR results were visualised on a 1.5% agarose gel containing ethidium bromide. For additional verification of the accuracy of determining the site in the genome, Sanger sequencing of the PCR product obtained was carried out.

4.7. Computational Modelling and Analysis of Protein Structure

Protein structures were modelled using AlphaFold 2.1, AlphaFold 2.2 [64] and RoseTTAFold, and visualised using Pymol 2.5 (Schrödinger Inc., New York, NY, USA) [125]. The models obtained were superimposed with the experimentally determined structures using Pymol. The robustness of structural alignments was assessed using root-mean-square deviation (RMSD), calculated using Pymol. Multiple protein structure alignment was carried out using mTM-align [126]. The phylogenetic tree was constructed based on the TM-score matrix [127], using the neighbour-joining method [128].

5. Conclusions

Due to the prospect of using phages and phage-derived antibacterials for therapy in the context of multi-drug-resistant bacterial infections, genomic studies of prophage-derived regions are of great interest. Studies of *Curtobacterium* genomes have indicated the presence of prophage-derived regions. The number of these regions appears to be smaller than in some other well-studied taxonomic groups, but the analysis and structural modelling of encoded proteins has highlighted the potential of cell-wall-degrading enzymes (CWDEs) for future use. The diversity of CWDEs may indicate the complex structure of the *Curtobacterium* cell envelope, and can facilitate an understanding of the mechanisms of *Curtobacterium* phage resistance.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms24021586/s1>.

Author Contributions: Conceptualisation, P.E. and K.M.; methodology, P.E. and A.P.; validation, P.E.; formal analysis, P.E., A.L., R.T., A.T., A.P., E.K., M.S. and K.M.; investigation, P.E., A.L., R.T., A.T. and E.K.; resources, K.M. and A.I.; data curation, P.E., A.L., A.L., R.T., E.K. and A.T.; writing—original draft preparation, P.E., K.M., A.P. and R.T.; writing—review and editing, P.E. and K.M.; visualisation, P.E., A.L. and R.T.; supervision, K.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Russian Science Foundation, grant #21-16-00047.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors thank Lubov Dorofeeva for providing samples of *Curtobacterium* spp. from the VKM collection.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Chase, A.B.; Arevalo, P.; Polz, M.F.; Berlemont, R.; Martiny, J.B.H. Evidence for Ecological Flexibility in the Cosmopolitan Genus *Curtobacterium*. *Front. Microbiol.* **2016**, *7*, 1874. [[CrossRef](#)] [[PubMed](#)]
2. Chase, A.B.; Gomez-Lunar, Z.; Lopez, A.E.; Li, J.; Allison, S.D.; Martiny, A.C.; Martiny, J.B.H. Emergence of Soil Bacterial Ecotypes along a Climate Gradient. *Environ. Microbiol.* **2018**, *20*, 4112–4126. [[CrossRef](#)]

3. Undabarrena, A.; Beltrametti, F.; Claverías, F.P.; González, M.; Moore, E.R.B.; Seeger, M.; Cámara, B. Exploring the Diversity and Antimicrobial Potential of Marine Actinobacteria from the Comau Fjord in Northern Patagonia, Chile. *Front. Microbiol.* **2016**, *7*, 1135. [[CrossRef](#)] [[PubMed](#)]
4. Ellington, A.J.; Bryan, N.C.; Christner, B.C.; Reisch, C.R. Draft Genome Sequences of Actinobacterial and Betaproteobacterial Strains Isolated from the Stratosphere. *Microbiol. Resour. Announc.* **2021**, *10*, e01009-21. [[CrossRef](#)] [[PubMed](#)]
5. Scales, N.C.; Chase, A.B.; Finks, S.S.; Malik, A.A.; Weihe, C.; Allison, S.D.; Martiny, A.C.; Martiny, J.B.H. Differential Response of Bacterial Microdiversity to Simulated Global Change. *Appl. Environ. Microbiol.* **2022**, *88*, e0242921. [[CrossRef](#)]
6. Liu, J.; Song, M.; Wei, X.; Zhang, H.; Bai, Z.; Zhuang, X. Responses of Phyllosphere Microbiome to Ozone Stress: Abundance, Community Compositions and Functions. *Microorganisms* **2022**, *10*, 680. [[CrossRef](#)]
7. Hone, H.; Mann, R.; Yang, G.; Kaur, J.; Tannenbaum, I.; Li, T.; Spangenberg, G.; Sawbridge, T. Profiling, Isolation and Characterisation of Beneficial Microbes from the Seed Microbiomes of Drought Tolerant Wheat. *Sci. Rep.* **2021**, *11*, 11916. [[CrossRef](#)]
8. Gerna, D.; Clara, D.; Allwardt, D.; Mitter, B.; Roach, T. Tailored Media Are Key to Unlocking the Diversity of Endophytic Bacteria in Distinct Compartments of Germinating Seeds. *Microbiol. Spectr.* **2022**, *10*, e0017222. [[CrossRef](#)]
9. Bziuk, N.; Maccario, L.; Straube, B.; Wehner, G.; Sørensen, S.J.; Schikora, A.; Smalla, K. The Treasure inside Barley Seeds: Microbial Diversity and Plant Beneficial Bacteria. *Environ. Microbiome* **2021**, *16*, 20. [[CrossRef](#)]
10. Schillaci, M.; Raio, A.; Sillo, F.; Zampieri, E.; Mahmood, S.; Anjum, M.; Khalid, A.; Centritto, M. Pseudomonas and Curtobacterium Strains from Olive Rhizosphere Characterized and Evaluated for Plant Growth Promoting Traits. *Plants* **2022**, *11*, 2245. [[CrossRef](#)]
11. Díez-Méndez, A.; Rivas, R. Improvement of Saffron Production Using Curtobacterium Herbarum as a Bioinoculant under Greenhouse Conditions. *AIMS Microbiol.* **2017**, *3*, 354–364. [[CrossRef](#)] [[PubMed](#)]
12. Dufresne, A.; Salanoubat, M.; Partensky, F.; Artiguenave, F.; Axmann, I.M.; Barbe, V.; Duprat, S.; Galperin, M.Y.; Koonin, E.V.; Le Gall, F.; et al. Genome Sequence of the Cyanobacterium Prochlorococcus Marinus SS120, a Nearly Minimal Oxyphototrophic Genome. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 10020–10025. [[CrossRef](#)]
13. Janisiewicz, W.J.; Buyer, J.S. Culturable Bacterial Microflora Associated with Nectarine Fruit and Their Potential for Control of Brown Rot. *Can. J. Microbiol.* **2010**, *56*, 480–486. [[CrossRef](#)] [[PubMed](#)]
14. Osdaghi, E.; Young, A.J.; Harveson, R.M. Bacterial Wilt of Dry Beans Caused by Curtobacterium Flaccumfaciens Pv. Flaccumfaciens: A New Threat from an Old Enemy. *Mol. Plant Pathol.* **2020**, *21*, 605–621. [[CrossRef](#)] [[PubMed](#)]
15. Mallick, P. A Rare Case of Curtobacterium Flaccumfaciens Infection in the Eye: A Case Report. *Trop. Med. Health* **2022**, *50*, 64. [[CrossRef](#)]
16. Francis, M.J.; Doherty, R.R.; Patel, M.; Hamblin, J.F.; Ojaimi, S.; Korman, T.M. Curtobacterium Flaccumfaciens Septic Arthritis Following Puncture with a Coxspur Hawthorn Thorn. *J. Clin. Microbiol.* **2011**, *49*, 2759–2760. [[CrossRef](#)]
17. Buonauro, R.; Moretti, C.; da Silva, D.P.; Cortese, C.; Ramos, C.; Venturi, V. The Olive Knot Disease as a Model to Study the Role of Interspecies Bacterial Communities in Plant Disease. *Front. Plant Sci.* **2015**, *6*, 434. [[CrossRef](#)]
18. Kamei, I.; Yoshida, T.; Enami, D.; Meguro, S. Coexisting Curtobacterium Bacterium Promotes Growth of White-Rot Fungus Stereum sp. *Curr. Microbiol.* **2012**, *64*, 173–178. [[CrossRef](#)]
19. Evseev, P.; Lukianova, A.; Tarakanov, R.; Tokmakova, A.; Shneider, M.; Ignatov, A.; Miroshnikov, K. Curtobacterium spp. and Curtobacterium flaccumfaciens: Phylogeny, Genomics-Based Taxonomy, Pathogenicity, and Diagnostics. *Curr. Issues Mol. Biol.* **2022**, *44*, 889–927. [[CrossRef](#)]
20. Casjens, S. Prophages and Bacterial Genomics: What Have We Learned so Far? *Mol. Microbiol.* **2003**, *49*, 277–300. [[CrossRef](#)]
21. Brüssow, H.; Canchaya, C.; Hardt, W.-D. Phages and the Evolution of Bacterial Pathogens: From Genomic Rearrangements to Lysogenic Conversion. *Microbiol. Mol. Biol. Rev.* **2004**, *68*, 560–602. [[CrossRef](#)] [[PubMed](#)]
22. Canchaya, C.; Fournous, G.; Brüssow, H. The Impact of Prophages on Bacterial Chromosomes. *Mol. Microbiol.* **2004**, *53*, 9–18. [[CrossRef](#)] [[PubMed](#)]
23. Argov, T.; Azulay, G.; Pasechnek, A.; Stadnyuk, O.; Ran-Sapir, S.; Borovok, I.; Sigal, N.; Herskovits, A.A. Temperate Bacteriophages as Regulators of Host Behavior. *Curr. Opin. Microbiol.* **2017**, *38*, 81–87. [[CrossRef](#)]
24. Davies, E.V.; Winstanley, C.; Fothergill, J.L.; James, C.E. The Role of Temperate Bacteriophages in Bacterial Infection. *FEMS Microbiol. Lett.* **2016**, *363*, fnw015. [[CrossRef](#)]
25. Schroven, K.; Aertsen, A.; Lavigne, R. Bacteriophages as Drivers of Bacterial Virulence and Their Potential for Biotechnological Exploitation. *FEMS Microbiol. Rev.* **2021**, *45*, fuaa041. [[CrossRef](#)]
26. Home-Genome-NCBI. Available online: <https://www.ncbi.nlm.nih.gov/genome> (accessed on 11 November 2021).
27. Arndt, D.; Marcu, A.; Liang, Y.; Wishart, D.S. PHAST, PHASTER and PHASTEST: Tools for Finding Prophage in Bacterial Genomes. *Brief. Bioinform.* **2017**, *20*, 1560–1567. [[CrossRef](#)]
28. Akhter, S.; Aziz, R.K.; Edwards, R.A. PhiSpy: A Novel Algorithm for Finding Prophages in Bacterial Genomes That Combines Similarity- and Composition-Based Strategies. *Nucleic Acids Res.* **2012**, *40*, e126. [[CrossRef](#)] [[PubMed](#)]
29. Sirén, K.; Millard, A.; Petersen, B.; Gilbert, M.T.P.; Clokie, M.R.J.; Sicheritz-Pontén, T. Rapid Discovery of Novel Prophages Using Biological Feature Engineering and Machine Learning. *NAR Genom. Bioinform.* **2021**, *3*, lqaa109. [[CrossRef](#)]
30. Roach, M.J.; McNair, K.; Michalczyk, M.; Giles, S.K.; Inglis, L.K.; Pargin, E.; Barylski, J.; Roux, S.; Decewicz, P.; Edwards, R.A. Philypmics 2021: Prophage Predictions Perplex Programs. *F1000 Res.* **2022**, *10*, 758. [[CrossRef](#)]
31. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic Local Alignment Search Tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [[CrossRef](#)]

32. Gabler, F.; Nam, S.-Z.; Till, S.; Mirdita, M.; Steinegger, M.; Söding, J.; Lupas, A.N.; Alva, V. Protein Sequence Analysis Using the MPI Bioinformatics Toolkit. *Curr. Protoc. Bioinform.* **2020**, *72*, e108. [[CrossRef](#)] [[PubMed](#)]
33. Gokey, T.; Halavaty, A.S.; Minasov, G.; Anderson, W.F.; Kuhn, M.L. Structure of the Bacillus Anthracis DTDP-L-Rhamnose Biosynthetic Pathway Enzyme: DTDP- α -D-Glucose 4,6-Dehydratase, RfbB. *J. Struct. Biol.* **2018**, *202*, 175–181. [[CrossRef](#)] [[PubMed](#)]
34. Theodorou, I.; Courtin, P.; Palussière, S.; Kulakauskas, S.; Bidnenko, E.; Péchoux, C.; Fenaille, F.; Penno, C.; Mahony, J.; van Sinderen, D.; et al. A Dual-Chain Assembly Pathway Generates the High Structural Diversity of Cell-Wall Polysaccharides in *Lactococcus Lactis*. *J. Biol. Chem.* **2019**, *294*, 17612–17625. [[CrossRef](#)] [[PubMed](#)]
35. Balzaretto, S.; Taverniti, V.; Guglielmetti, S.; Fiore, W.; Minuzzo, M.; Ngo, H.N.; Ngere, J.B.; Sadiq, S.; Humphreys, P.N.; Laws, A.P. A Novel Rhamnose-Rich Hetero-Exopolysaccharide Isolated from *Lactobacillus Paracasei* DG Activates THP-1 Human Monocytic Cells. *Appl. Environ. Microbiol.* **2017**, *83*, e02702-16. [[CrossRef](#)] [[PubMed](#)]
36. Zatyka, M.; Thomas, C.M. Control of Genes for Conjugative Transfer of Plasmids and Other Mobile Elements. *FEMS Microbiol. Rev.* **1998**, *21*, 291–319. [[CrossRef](#)]
37. Zhou, Y.; Liang, Y.; Lynch, K.H.; Dennis, J.J.; Wishart, D.S. PHAST: A Fast Phage Search Tool. *Nucleic Acids Res.* **2011**, *39*, W347–W352. [[CrossRef](#)]
38. Czajkowski, R. May the Phage Be With You? Prophage-Like Elements in the Genomes of Soft Rot *Pectobacteriaceae*: *Pectobacterium* spp. and *Dickeya* spp. *Front. Microbiol.* **2019**, *10*, 138. [[CrossRef](#)]
39. Moraru, C.; Varsani, A.; Kropinski, A.M. VIRIDIC—A Novel Tool to Calculate the Intergenomic Similarities of Prokaryote-Infecting Viruses. *Viruses* **2020**, *12*, 1268. [[CrossRef](#)]
40. Juhala, R.J.; Ford, M.E.; Duda, R.L.; Youlton, A.; Hatfull, G.F.; Hendrix, R.W. Genomic Sequences of Bacteriophages HK97 and HK022: Pervasive Genetic Mosaicism in the Lambdoid Bacteriophages. *J. Mol. Biol.* **2000**, *299*, 27–51. [[CrossRef](#)]
41. Pedulla, M.L.; Ford, M.E.; Houtz, J.M.; Karthikeyan, T.; Wadsworth, C.; Lewis, J.A.; Jacobs-Sera, D.; Falbo, J.; Gross, J.; Pannunzio, N.R.; et al. Origins of Highly Mosaic Mycobacteriophage Genomes. *Cell* **2003**, *113*, 171–182. [[CrossRef](#)]
42. Evseev, P.; Sykilinda, N.; Gorshkova, A.; Kurochkina, L.; Ziganshin, R.; Drucker, V.; Miroshnikov, K. Pseudomonas Phage PaBG—A Jumbo Member of an Old Parasite Family. *Viruses* **2020**, *12*, 721. [[CrossRef](#)] [[PubMed](#)]
43. Fokine, A.; Rossmann, M.G. Molecular Architecture of Tailed Double-Stranded DNA Phages. *Bacteriophage* **2014**, *4*, e28281. [[CrossRef](#)] [[PubMed](#)]
44. Evseev, P.; Shneider, M.; Miroshnikov, K. Evolution of Phage Tail Sheath Protein. *Viruses* **2022**, *14*, 1148. [[CrossRef](#)]
45. Shao, Q.; Trinh, J.T.; Zeng, L. High-Resolution Studies of Lysis–Lysogeny Decision-Making in Bacteriophage Lambda. *J. Biol. Chem.* **2019**, *294*, 3343–3349. [[CrossRef](#)] [[PubMed](#)]
46. Bose, B.; Auchtung, J.M.; Lee, C.A.; Grossman, A.D. A Conserved Anti-Repressor Controls Horizontal Gene Transfer by Proteolysis. *Mol. Microbiol.* **2008**, *70*, 570–582. [[CrossRef](#)] [[PubMed](#)]
47. LeBowitz, J.H.; McMacken, R. The Bacteriophage Lambda O and P Protein Initiators Promote the Replication of Single-Stranded DNA. *Nucleic Acids Res.* **1984**, *12*, 3069–3088. [[CrossRef](#)]
48. Ravin, V.; Ravin, N.; Casjens, S.; Ford, M.E.; Hatfull, G.F.; Hendrix, R.W. Genomic Sequence and Analysis of the Atypical Temperate Bacteriophage N15. *J. Mol. Biol.* **2000**, *299*, 53–73. [[CrossRef](#)]
49. Krupovic, M.; Koonin, E.V. Multiple Origins of Viral Capsid Proteins from Cellular Ancestors. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E2401–E2410. [[CrossRef](#)]
50. Taylor, A.L. BACTERIOPHAGE-INDUCED MUTATION IN ESCHERICHIA COLI*. *Proc. Natl. Acad. Sci. USA* **1963**, *50*, 1043–1051. [[CrossRef](#)]
51. Harshey, R.M. The Mu Story: How a Maverick Phage Moved the Field Forward. *Mob. DNA* **2012**, *3*, 21. [[CrossRef](#)]
52. Young, R. Phage Lysis: Three Steps, Three Choices, One Outcome. *J. Microbiol.* **2014**, *52*, 243–258. [[CrossRef](#)] [[PubMed](#)]
53. Ramanculov, E.; Young, R. An Ancient Player Unmasked: T4 RI Encodes a t-Specific Antiholin. *Mol. Microbiol.* **2001**, *41*, 575–583. [[CrossRef](#)] [[PubMed](#)]
54. Guo, T.; Xin, Y.; Zhang, C.; Kong, J. A Cytoplasmic Antiholin Is Embedded In Frame with the Holin in a *Lactobacillus Fermentum* Bacteriophage. *Appl. Environ. Microbiol.* **2018**, *84*, e02518-17. [[CrossRef](#)] [[PubMed](#)]
55. Latka, A.; Maciejewska, B.; Majkowska-Skrobek, G.; Briers, Y.; Drulis-Kawa, Z. Bacteriophage-Encoded Virion-Associated Enzymes to Overcome the Carbohydrate Barriers during the Infection Process. *Appl. Microbiol. Biotechnol.* **2017**, *101*, 3103–3119. [[CrossRef](#)] [[PubMed](#)]
56. Lee, I.; Ouk Kim, Y.; Park, S.-C.; Chun, J. OrthoANI: An Improved Algorithm and Software for Calculating Average Nucleotide Identity. *Int. J. Syst. Evol. Microbiol.* **2016**, *66*, 1100–1103. [[CrossRef](#)] [[PubMed](#)]
57. Tse, V.Y.; Liu, H.; Kapinos, A.; Torres, C.; Cayabyab, B.C.S.; Fett, S.N.; Nakashima, L.G.; Rahman, M.; Vargas, A.S.; Reddi, K.; et al. Exploring Genomic Conservation in Actinobacteriophages with Small Genomes. *IEEE Access* **2022**, *7*, 87931–87942.
58. Taxonomy. Available online: <https://talk.ictvonline.org/taxonomy/> (accessed on 27 April 2022).
59. Turner, D.; Kropinski, A.M.; Adriaenssens, E.M. A Roadmap for Genome-Based Phage Taxonomy. *Viruses* **2021**, *13*, 506. [[CrossRef](#)]
60. Dion, M.B.; Plante, P.-L.; Zufferey, E.; Shah, S.A.; Corbeil, J.; Moineau, S. Streamlining CRISPR Spacer-Based Bacterial Host Predictions to Decipher the Viral Dark Matter. *Nucleic Acids Res.* **2021**, *49*, 3127–3138. [[CrossRef](#)]
61. Stern, A.; Mick, E.; Tirosh, I.; Sagy, O.; Sorek, R. CRISPR Targeting Reveals a Reservoir of Common Phages Associated with the Human Gut Microbiome. *Genome Res.* **2012**, *22*, 1985–1994. [[CrossRef](#)]

62. Jones, P.; Binns, D.; Chang, H.-Y.; Fraser, M.; Li, W.; McAnulla, C.; McWilliam, H.; Maslen, J.; Mitchell, A.; Nuka, G.; et al. InterProScan 5: Genome-Scale Protein Function Classification. *Bioinformatics* **2014**, *30*, 1236–1240. [[CrossRef](#)]
63. Levy-Assaraf, M.; Voronov-Goldman, M.; Rozman Grinberg, I.; Weiserman, G.; Shimon, L.J.W.; Jindou, S.; Borovok, I.; White, B.A.; Bayer, E.A.; Lamed, R.; et al. Crystal Structure of an Uncommon Cellulosome-Related Protein Module from *Ruminococcus Flavefaciens* That Resembles Papain-Like Cysteine Peptidases. *PLoS ONE* **2013**, *8*, e56138. [[CrossRef](#)] [[PubMed](#)]
64. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. [[CrossRef](#)] [[PubMed](#)]
65. Baek, M.; DiMaio, F.; Anishchenko, I.; Dauparas, J.; Ovchinnikov, S.; Lee, G.R.; Wang, J.; Cong, Q.; Kinch, L.N.; Schaeffer, R.D.; et al. Accurate Prediction of Protein Structures and Interactions Using a Three-Track Neural Network. *Science* **2021**, *373*, 871–876. [[CrossRef](#)] [[PubMed](#)]
66. Kerff, F.; Petrella, S.; Mercier, F.; Sauvage, E.; Herman, R.; Pennartz, A.; Zervosen, A.; Luxen, A.; Frère, J.-M.; Joris, B.; et al. Specific Structural Features of the N-Acetylmuramoyl-L-Alanine Amidase AmiD from *Escherichia Coli* and Mechanistic Implications for Enzymes of This Family. *J. Mol. Biol.* **2010**, *397*, 249–259. [[CrossRef](#)] [[PubMed](#)]
67. Holm, L. Dali Server: Structural Unification of Protein Families. *Nucleic Acids Res.* **2022**, *50*, W210–W215. [[CrossRef](#)]
68. Xu, X.; Zhang, D.; Zhou, B.; Zhen, X.; Ouyang, S. Structural and Biochemical Analyses of the Tetrameric Cell Binding Domain of Lys170 from Enterococcal Phage F170/08. *Eur. Biophys. J.* **2021**, *50*, 721–729. [[CrossRef](#)] [[PubMed](#)]
69. Lemak, S.; Tchigvintsev, A.; Petit, P.; Flick, R.; Singer, A.U.; Brown, G.; Evdokimova, E.; Egorova, O.; Gonzalez, C.F.; Chernikova, T.N.; et al. Structure and Activity of the Cold-Active and Anion-Activated Carboxyl Esterase OLEI01171 from the Oil-Degrading Marine Bacterium *Oleispira Antarctica*. *Biochem. J.* **2012**, *445*, 193–203. [[CrossRef](#)]
70. Gorelik, A.; Illes, K.; Nagar, B. Crystal Structure of the Mannose-6-Phosphate Uncovering Enzyme. *Structure* **2020**, *28*, 426–436.e3. [[CrossRef](#)]
71. Ji, S.; Dix, S.R.; Aziz, A.A.; Sedelnikova, S.E.; Baker, P.J.; Rafferty, J.B.; Bullough, P.A.; Tzokov, S.B.; Agirre, J.; Li, F.-L.; et al. The Molecular Basis of Endolytic Activity of a Multidomain Alginate Lyase from *Defluviitalea Phaphyphila*, a Representative of a New Lyase Family, PL39. *J. Biol. Chem.* **2019**, *294*, 18077–18091. [[CrossRef](#)]
72. Leiman, P.G.; Kanamaru, S.; Mesyanzhinov, V.V.; Arisaka, F.; Rossmann, M.G. Structure and Morphogenesis of Bacteriophage T4. *Cell Mol. Life Sci.* **2003**, *60*, 2356–2370. [[CrossRef](#)]
73. Garcia-Doval, C.; van Raaij, M.J. Structure of the Receptor-Binding Carboxy-Terminal Domain of Bacteriophage T7 Tail Fibers. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 9390–9395. [[CrossRef](#)] [[PubMed](#)]
74. Drobiazko, A.Y.; Kasimova, A.A.; Evseev, P.V.; Shneider, M.M.; Klimuk, E.I.; Shashkov, A.S.; Dmitrenok, A.S.; Chizhov, A.O.; Slukin, P.V.; Skryabin, Y.P.; et al. Capsule-Targeting Depolymerases Derived from *Acinetobacter Baumannii* Prophage Regions. *Int. J. Mol. Sci.* **2022**, *23*, 4971. [[CrossRef](#)] [[PubMed](#)]
75. Smith, N.L.; Taylor, E.J.; Lindsay, A.-M.; Charnock, S.J.; Turkenburg, J.P.; Dodson, E.J.; Davies, G.J.; Black, G.W. Structure of a Group A Streptococcal Phage-Encoded Virulence Factor Reveals a Catalytically Active Triple-Stranded β -Helix. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 17652–17657. [[CrossRef](#)] [[PubMed](#)]
76. Greenfield, J.; Shang, X.; Luo, H.; Zhou, Y.; Heselpoth, R.D.; Nelson, D.C.; Herzberg, O. Structure and Tailspike Glycosidase Machinery of ORF212 from *E. Coli* O157:H7 Phage CBA120 (TSP3). *Sci. Rep.* **2019**, *9*, 7349. [[CrossRef](#)] [[PubMed](#)]
77. Petersen, T.N.; Kauppinen, S.; Larsen, S. The Crystal Structure of Rhamnogalacturonase A from *Aspergillus Aculeatus*: A Right-Handed Parallel Beta Helix. *Structure* **1997**, *5*, 533–544. [[CrossRef](#)]
78. Kovalová, T.; Koval, T.; Benešová, E.; Vodicková, P.; Spiwok, V.; Lipovová, P.; Dohnálek, J. Active Site Complementation and Hexameric Arrangement in the GH Family 29; a Structure-Function Study of α -l-Fucosidase Isoenzyme 1 from *Paenibacillus Thiaminolyticus*. *Glycobiology* **2019**, *29*, 59–73. [[CrossRef](#)]
79. Zhou, S.; Ko, T.-P.; Huang, J.-W.; Liu, W.; Zheng, Y.; Wu, S.; Wang, Q.; Xie, Z.; Liu, Z.; Chen, C.-C.; et al. Structure of a Gut Microbial Diltiazem-Metabolizing Enzyme Suggests Possible Substrate Binding Mode. *Biochem. Biophys. Res. Commun.* **2020**, *527*, 799–804. [[CrossRef](#)]
80. Tarakanov, R.I.; Lukianova, A.A.; Evseev, P.V.; Pilik, R.I.; Tokmakova, A.D.; Kulikov, E.E.; Toshchakov, S.V.; Ignatov, A.N.; Dzhaliylov, F.S.-U.; Miroshnikov, K.A. Ayka, a Novel Curtobacterium Bacteriophage, Provides Protection against Soybean Bacterial Wilt and Tan Spot. *Int. J. Mol. Sci.* **2022**, *23*, 10913. [[CrossRef](#)]
81. Four Novel Curtobacterium Phages Isolated from Environmental Samples. Available online: <https://www.researchsquare.com> (accessed on 7 November 2022).
82. Carstens, A.B.; Djurhuus, A.M.; Kot, W.; Jacobs-Sera, D.; Hatfull, G.F.; Hansen, L.H. Unlocking the Potential of 46 New Bacteriophages for Biocontrol of *Dickeya Solani*. *Viruses* **2018**, *10*, 621. [[CrossRef](#)]
83. Miroshnikov, K.A.; Evseev, P.V.; Lukianova, A.A.; Ignatov, A.N. Tailed Lytic Bacteriophages of Soft Rot Pectobacteriaceae. *Microorganisms* **2021**, *9*, 1819. [[CrossRef](#)]
84. Runtuvuori-Salmela, A.; Kunttu, H.M.T.; Laanto, E.; Almeida, G.M.F.; Mäkelä, K.; Middelboe, M.; Sundberg, L.-R. Prevalence of Genetically Similar Flavobacterium Columnare Phages across Aquaculture Environments Reveals a Strong Potential for Pathogen Control. *Environ. Microbiol.* **2022**, *24*, 2404–2420. [[CrossRef](#)] [[PubMed](#)]

85. Jordan, T.C.; Burnett, S.H.; Carson, S.; Caruso, S.M.; Clase, K.; DeJong, R.J.; Dennehy, J.J.; Denver, D.R.; Dunbar, D.; Elgin, S.C.R.; et al. A Broadly Implementable Research Course in Phage Discovery and Genomics for First-Year Undergraduate Students. *mBio* **2014**, *5*, e01051-13. [[CrossRef](#)]
86. Hatfull, G.F. Mycobacteriophages. *Microbiol. Spectr.* **2018**, *6*, 1029–1055. [[CrossRef](#)] [[PubMed](#)]
87. Monteiro, R.; Pires, D.P.; Costa, A.R.; Azeredo, J. Phage Therapy: Going Temperate? *Trends Microbiol.* **2019**, *27*, 368–378. [[CrossRef](#)]
88. Cobb, L.H.; Park, J.; Swanson, E.A.; Beard, M.C.; McCabe, E.M.; Rourke, A.S.; Seo, K.S.; Olivier, A.K.; Priddy, L.B. CRISPR-Cas9 Modified Bacteriophage for Treatment of Staphylococcus Aureus Induced Osteomyelitis and Soft Tissue Infection. *PLoS ONE* **2019**, *14*, e0220421. [[CrossRef](#)] [[PubMed](#)]
89. Oliveira, H.; São-José, C.; Azeredo, J. Phage-Derived Peptidoglycan Degrading Enzymes: Challenges and Future Prospects for In Vivo Therapy. *Viruses* **2018**, *10*, 292. [[CrossRef](#)]
90. Drulis-Kawa, Z.; Majkowska-Skrobek, G.; Maciejewska, B. Bacteriophages and Phage-Derived Proteins—Application Approaches. *Curr. Med. Chem.* **2015**, *22*, 1757–1773. [[CrossRef](#)]
91. Grissa, I.; Vergnaud, G.; Pourcel, C. The CRISPRdb Database and Tools to Display CRISPRs and to Generate Dictionaries of Spacers and Repeats. *BMC Bioinform.* **2007**, *8*, 172. [[CrossRef](#)]
92. Makarova, K.S.; Wolf, Y.I.; Alkhnbashi, O.S.; Costa, F.; Shah, S.A.; Saunders, S.J.; Barrangou, R.; Brouns, S.J.J.; Charpentier, E.; Haft, D.H.; et al. An Updated Evolutionary Classification of CRISPR–Cas Systems. *Nat. Rev. Microbiol.* **2015**, *13*, 722–736. [[CrossRef](#)]
93. Zaychikov, V.A.; Potekhina, N.V.; Dmitrenok, A.S.; Fan, D.; Tul'skaya, E.M.; Dorofeeva, L.V.; Evtushenko, L.I. Cell Wall Rhamnan in Actinobacteria of the Genus *Curtobacterium*. *Microbiology* **2021**, *90*, 343–348. [[CrossRef](#)]
94. Rivera Starr, C.; Engleberg, N.C. Role of Hyaluronidase in Subcutaneous Spread and Growth of Group A Streptococcus. *Infect. Immun.* **2006**, *74*, 40–48. [[CrossRef](#)]
95. Di Lallo, G.; Falconi, M.; Iacovelli, F.; Frezza, D.; D'Addabbo, P. Analysis of Four New Enterococcus Faecalis Phages and Modeling of a Hyaluronidase Catalytic Domain from Saphexavirus. *PHAGE* **2021**, *2*, 131–141. [[CrossRef](#)] [[PubMed](#)]
96. Hynes, W.L.; Walton, S.L. Hyaluronidases of Gram-Positive Bacteria. *FEMS Microbiol. Lett.* **2000**, *183*, 201–207. [[CrossRef](#)]
97. Dowah, A.S.A.; Clokie, M.R.J. Review of the Nature, Diversity and Structure of Bacteriophage Receptor Binding Proteins That Target Gram-Positive Bacteria. *Biophys. Rev.* **2018**, *10*, 535–542. [[CrossRef](#)]
98. Bertozzi Silva, J.; Storms, Z.; Sauvageau, D. Host Receptors for Bacteriophage Adsorption. *FEMS Microbiol. Lett.* **2016**, *363*, fnw002. [[CrossRef](#)]
99. Rostøl, J.T.; Marraffini, L. (Ph)Ighting Phages—How Bacteria Resist Their Parasites. *Cell Host Microbe* **2019**, *25*, 184–194. [[CrossRef](#)]
100. Brockhurst, M.A.; Koskella, B.; Zhang, Q.-G. Bacteria-Phage Antagonistic Coevolution and the Implications for Phage Therapy. In *Bacteriophages: Biology, Technology, Therapy*; Harper, D., Abedon, S., Burrowes, B., McConville, M., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 1–21. ISBN 978-3-319-40598-8.
101. Sychantha, D.; Brott, A.S.; Jones, C.S.; Clarke, A.J. Mechanistic Pathways for Peptidoglycan O-Acetylation and De-O-Acetylation. *Front. Microbiol.* **2018**, *9*, 2332. [[CrossRef](#)] [[PubMed](#)]
102. Russo, T.A.; Luke, N.R.; Beanan, J.M.; Olson, R.; Sauberan, S.L.; MacDonald, U.; Schultz, L.W.; Umland, T.C.; Campagnari, A.A. The K1 Capsular Polysaccharide of *Acinetobacter Baumannii* Strain 307-0294 Is a Major Virulence Factor. *Infect. Immun.* **2010**, *78*, 3993–4000. [[CrossRef](#)] [[PubMed](#)]
103. Knecht, L.E.; Veljkovic, M.; Fieseler, L. Diversity and Function of Phage Encoded Depolymerases. *Front. Microbiol.* **2020**, *10*, 2949. [[CrossRef](#)]
104. Hitri, K.; Kuttel, M.M.; De Benedetto, G.; Lockyer, K.; Gao, F.; Hansal, P.; Rudd, T.R.; Beamish, E.; Rijpkema, S.; Ravenscroft, N.; et al. O-Acetylation of Typhoid Capsular Polysaccharide Confers Polysaccharide Rigidity and Immunodominance by Masking Additional Epitopes. *Vaccine* **2019**, *37*, 3866–3875. [[CrossRef](#)]
105. Fiebig, T.; Cramer, J.T.; Bethe, A.; Baruch, P.; Curth, U.; Führung, J.I.; Buettner, F.F.R.; Vogel, U.; Schubert, M.; Fedorov, R.; et al. Structural and Mechanistic Basis of Capsule O-Acetylation in *Neisseria Meningitidis* Serogroup A. *Nat. Commun.* **2020**, *11*, 4723. [[CrossRef](#)] [[PubMed](#)]
106. Yuan, Y.; Peng, Q.; Gao, M. Characteristics of a Broad Lytic Spectrum Endolysin from Phage BtCS33 of *Bacillus Thuringiensis*. *BMC Microbiol.* **2012**, *12*, 297. [[CrossRef](#)] [[PubMed](#)]
107. Oechslin, F.; Menzi, C.; Moreillon, P.; Resch, G. The Multidomain Architecture of a Bacteriophage Endolysin Enables Intramolecular Synergism and Regulation of Bacterial Lysis. *J. Biol. Chem.* **2021**, *296*, 100639. [[CrossRef](#)]
108. Santos, S.B.; Oliveira, A.; Melo, L.D.R.; Azeredo, J. Identification of the First Endolysin Cell Binding Domain (CBD) Targeting *Paenibacillus Larvae*. *Sci. Rep.* **2019**, *9*, 2568. [[CrossRef](#)]
109. Seemann, T. Prokka: Rapid Prokaryotic Genome Annotation. *Bioinformatics* **2014**, *30*, 2068–2069. [[CrossRef](#)] [[PubMed](#)]
110. Geneious Bioinformatics Software for Sequence Data Analysis. Available online: <https://www.geneious.com/> (accessed on 11 November 2022).
111. Kelley, L.A.; Mezulis, S.; Yates, C.M.; Wass, M.N.; Sternberg, M.J.E. The Phyre2 Web Portal for Protein Modeling, Prediction and Analysis. *Nat. Protoc.* **2015**, *10*, 845–858. [[CrossRef](#)] [[PubMed](#)]
112. O'Leary, N.A.; Wright, M.W.; Brister, J.R.; Ciufu, S.; Haddad, D.; McVeigh, R.; Rajput, B.; Robbertse, B.; Smith-White, B.; Ako-Adjei, D.; et al. Reference Sequence (RefSeq) Database at NCBI: Current Status, Taxonomic Expansion, and Functional Annotation. *Nucleic Acids Res.* **2016**, *44*, D733–D745. [[CrossRef](#)]

113. Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E.L. Predicting Transmembrane Protein Topology with a Hidden Markov Model: Application to Complete Genomes. *J. Mol. Biol.* **2001**, *305*, 567–580. [[CrossRef](#)]
114. Sullivan, M.J.; Petty, N.K.; Beatson, S.A. Easyfig: A Genome Comparison Visualizer. *Bioinformatics* **2011**, *27*, 1009–1010. [[CrossRef](#)]
115. MinCED-Mining CRISPRs in Environmental Datasets 2022. Available online: <https://github.com/ctSkennerton/minced> (accessed on 11 November 2022).
116. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T.J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; et al. Fast, Scalable Generation of High-Quality Protein Multiple Sequence Alignments Using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 539. [[CrossRef](#)]
117. Kozlov, A.M.; Darriba, D.; Flouri, T.; Morel, B.; Stamatakis, A. RAxML-NG: A Fast, Scalable and User-Friendly Tool for Maximum Likelihood Phylogenetic Inference. *Bioinformatics* **2019**, *35*, 4453–4455. [[CrossRef](#)] [[PubMed](#)]
118. Edler, D.; Klein, J.; Antonelli, A.; Silvestro, D. RaxmlGUI 2.0: A Graphical Interface and Toolkit for Phylogenetic Analyses Using RAxML. *Methods Ecol. Evol.* **2021**, *12*, 373–377. [[CrossRef](#)]
119. Veerassamy, S.; Smith, A.; Tillier, E.R.M. A Transition Probability Model for Amino Acid Substitutions from Blocks. *J. Comput. Biol.* **2003**, *10*, 997–1010. [[CrossRef](#)]
120. Darriba, D.; Posada, D.; Kozlov, A.M.; Stamatakis, A.; Morel, B.; Flouri, T. ModelTest-NG: A New and Scalable Tool for the Selection of DNA and Protein Evolutionary Models. *Mol. Biol. Evol.* **2020**, *37*, 291–294. [[CrossRef](#)]
121. Letunic, I.; Bork, P. Interactive Tree Of Life (ITOL) v5: An Online Tool for Phylogenetic Tree Display and Annotation. *Nucleic Acids Res.* **2021**, *49*, W293–W296. [[CrossRef](#)] [[PubMed](#)]
122. Adams, M.H. *Bacteriophages*; Interscience Publishers: New York, NY, USA, 1959.
123. Ackermann, H.-W. Basic Phage Electron Microscopy. In *Bacteriophages: Methods and Protocols, Volume 1: Isolation, Characterization, and Interactions*; Clokie, M.R.J., Kropinski, A.M., Eds.; Methods in Molecular Biology™; Humana Press: Totowa, NJ, USA, 2009; pp. 113–126. ISBN 978-1-60327-164-6.
124. Untergasser, A.; Cutcutache, I.; Koressaar, T.; Ye, J.; Faircloth, B.C.; Remm, M.; Rozen, S.G. Primer3—New Capabilities and Interfaces. *Nucleic Acids Res.* **2012**, *40*, e115. [[CrossRef](#)] [[PubMed](#)]
125. PyMOL Pymol.Org. Available online: <https://pymol.org/2/> (accessed on 11 November 2021).
126. Dong, R.; Peng, Z.; Zhang, Y.; Yang, J. MTM-Align: An Algorithm for Fast and Accurate Multiple Protein Structure Alignment. *Bioinformatics* **2018**, *34*, 1719–1725. [[CrossRef](#)]
127. Zhang, Y.; Skolnick, J. Scoring Function for Automated Assessment of Protein Structure Template Quality. *Proteins* **2004**, *57*, 702–710. [[CrossRef](#)]
128. Saitou, N.; Nei, M. The Neighbor-Joining Method: A New Method for Reconstructing Phylogenetic Trees. *Mol. Biol. Evol.* **1987**, *4*, 406–425. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.