

The e14 cryptic prophage belongs to a temperate phage lineage distributed across *Escherichia* and retains genes with potential anti-phage defence functions

Danna R. Gifford

Division of Evolution, Infection and Genomics, School of Biological Sciences, Faculty of Biology, Medicine and Health University of Manchester, Manchester, UK Email: danna.gifford@manchester.ac.uk

Abstract

The e14 region of *Escherichia coli* str. K-12 is a cryptic prophage that retains functional excision and integration machinery despite extensive loss of phage reproductive functions. However, its evolutionary relationship to other bacteriophages remains poorly resolved. Here, we identify related prophages across four *Escherichia* species, alongside related, independently isolated bacteriophages. The putative prophages were 34.7–43.5 kb, more than twice the size of e14 (15.2 kb), and retained much of the genomic architecture required for phage reproduction. Comparative genomic and phylogenetic analyses support e14 as a highly reduced member of a broader lineage of temperate bacteriophages. The retention of genes with established or putative anti-phage functions further raises the possibility that e14 represents a form of ‘defence flotsam’, potentially protecting its host against subsequent phage infections. More broadly, cryptic prophages may provide reservoirs of anti-phage defence functions long after losing their capacity for phage reproduction.

Keywords: *Escherichia coli*, cryptic prophage, comparative genomics, bacteriophage evolution, genome reduction, anti-phage defence

Introduction

Cryptic prophages are common components of bacterial genomes, but extensive gene loss can obscure their evolutionary origins. The *Escherichia coli* str. K-12 e14 element is an SOS-excisable, defective lambdoid prophage that retains several biologically active genes despite substantial loss of phage reproductive functions [1–3]. Previous comparisons identified similarities between e14 and other lambdoid prophages, including larger elements in *E. coli* O157, but its broader evolutionary context remains poorly resolved [3, 4]. Identifying more complete relatives could also help establish which functions have been lost during prophage degradation and which have been retained.

Here, comparative genomics across *Escherichia* was used to identify and characterise larger e14-like prophages in *E. coli*, *E. fergusonii*, *E. albertii* and *E. marmotae*. Notably, the recently sequenced *E. coli* JB2 genome [5] contains a 43.5-kb e14-like element with extensive conservation of gene content and synteny. Comparison between these elements and related isolated bacteriophages, including *Escherichia* phages vB_EcoM-689R6 and vB_EcoM-720R4 [6], allows reconstruction of the broader genomic architecture from which e14 may have been reduced. These comparisons reveal extensive loss of phage reproductive functions, consistent with minimising detrimental impacts on its host. Moreover, e14 retains several genes with known or potential anti-phage activities, raising the possibility that this defective prophage may continue to provide protection against phage infection.

Methods

To identify putative e14-like prophages, NCBI BLASTn was performed, querying the 15.2-kb e14 region from *E. coli* str. K-12 substr. MG1655 (RefSeq accession NC_000913.3) against non-*E. coli* strains of *Escherichia*. *E. coli* was excluded from this initial search as the vast majority of matches returned were from K-12 isolates, and were identical to the e14 region itself. Strong matches were found for *E. fergusonii*, *E. albertii* and *E. marmotae*. Within each species, the genome with the highest query coverage and nucleotide identity was retained as a representative for analysis. Manual inspection of the surrounding genome annotations indicated putative prophage genes. To identify putative complete e14-like prophages in *E. coli*, a subsequent NCBI BLASTn was performed, querying the putative prophage

from *E. fergusonii* XJ19MCE1 against *E. coli* genomes; this identified a strong match in the recently-sequenced *E. coli* JB2 genome, which was also retained for analysis (Table 1).

Table 1: Representative e14-like matches identified across *Escherichia* species using NCBI BLASTn. Query coverage and nucleotide identity are relative to the *E. coli* str. K-12 substr. MG1655 e14 region (RefSeq accession NC_000913.3:1196209-1211412) as a query sequence.

Species/strain	Accession	Coordinates	Coverage (%)	Nucleotide identity (%)	Extracted fragment size (kb)
<i>E. coli</i> str. K-12 substr. MG1655	NC_000913.3	1196209-1211412	100	100	15.2
<i>E. coli</i> JB2	CP182203.1	3110039-3153542	91	97.07	43.5
<i>E. fergusonii</i> XJ19MCE1	CP125351.1	2656682-2699033	75	96.03	42.4
<i>E. albertii</i> RM9976	CP043262.1	2733991-2773208	70	98.18	39.2
<i>E. marmotae</i> F12YCO47	CP173213.1	2417684-2452364	49	95.63	34.7

Genomic regions corresponding to putative prophages were extracted from each genome. Homologous boundary sequences were identified by BLASTn in each candidate genomic context, and paired hits bracketing the candidate prophage were used to infer conservative *attL/attR*-associated extraction boundaries; where this was inconclusive, direct boundary homology or the original seed interval was retained (Supplementary File 1).

Extracted sequences and complete phages were annotated with Pharokka v1.10.1 [7] and Phold v1.2.6 [8]. Protein-level gene content and synteny were compared with Clinker v0.0.32 [9] using a minimum amino-acid identity of 30% (Supplementary File 2). NCBI SmartBLAST was used to identify homologous proteins and assess their taxonomic occurrence. Foldseek webserver (<https://search.foldseek.com/>) [10] was used to identify structural homologs and provide additional evidence for functional assignment, using the Big Fantastic Virus Database (BFVD, version 3) of predicted viral protein structures [11]. Structural matches were evaluated using the template modelling score (TM-score), which measures structural similarity on a scale from 0 to 1; root-mean-square deviation (RMSD, in units of angstroms i.e. Å), which measures the average positional difference between aligned protein structures; the expectation value (E-value), which estimates the number of matches of comparable statistical score expected by chance; and amino-acid sequence identity (%).

For phylogenetic analysis, focal prophages were searched against the INPHARED database [12] using Mash [13] ($k = 21$; sketch size = 1,000). Reference phages with Mash distances ≤ 0.15 were identified, retrieved from NCBI and ranked by their minimum distance to the focal prophages. Near-duplicate references (Mash distance < 0.002) were excluded (see Supplementary File 3 for final inclusion list). Proteins were predicted using PHANOTATE [14] and clustered using MMseqs2 [15], with a minimum sequence identity of 30% and bidirectional coverage of 50%. Six single-copy protein families containing e14, present in $\geq 25\%$ of retained genomes and with a median length ≥ 50 amino acids, were selected (given in Supplementary File 4). The six qualifying families were aligned using MAFFT [16], trimmed using ClipKIT ('smart-gap' mode) [17] and concatenated, with missing proteins represented by gaps. Taxa represented by fewer than three families were excluded, yielding an alignment of 21 sequences and 1,316 amino-acid positions. A maximum-likelihood phylogeny was inferred using IQ-TREE 2 [18], with automated model selection and partition merging ('MFP+MERGE'). Branch support was assessed using 1,000 ultrafast bootstrap and 1,000 SH-aLRT replicates. Two identified phages, vB_EcoM-689R6 (ON470584.1) and vB_EcoM-720R4 (ON470603.1) [6], were selected from the INPHARED search as complete phage comparators for the protein-level gene content and synteny analysis described previously.

GPT-6.1 Sol (Medium; OpenAI) was used to assist with initial coding of the analysis pipeline, and with debugging and consistency across pipeline steps. All code was reviewed and validated by the author. The pipeline is available in a public repository (<https://github.com/dannagifford/e14-like-prophages-escherichia>).

Results and discussion

Putative e14-like prophages were identified in *E. coli* JB2, *E. fergusonii*, *E. albertii* and *E. marmotae* (Figure 1), indicating that this phage lineage has circulated across multiple *Escherichia* species. Related complete *Escherichia* phages were also recovered from INPHARED, including vB_EcoM-689R6 and vB_EcoM-720R4, supporting the interpretation that the larger chromosomal elements represent substantially more complete members of the same phage lineage.

The 43.5-kb region from *E. coli* JB2 showed the highest coverage of the e14 query among the larger prophages (91%, with 97.07% nucleotide identity), whereas *E. albertii* RM9976 had higher identity (98.18%) but lower coverage (70%). Consistent with previous studies showing that e14-like prophages preferentially integrate at the *icd* locus, e14-like prophages were found at this site in all but one species. The exception was *E. marmotae*, in which the e14-like prophage was located adjacent to gene *sbcB*, and was also lacking integrase and excisionase genes. Whether this reflects integration at an alternative site or subsequent genomic rearrangement remains unknown. Interestingly, the region immediately downstream of *icd* in *E. marmotae* contains several phage-related pseudogenes, including frameshifted genes encoding a terminase small subunit and a capsid protein, consistent with a cryptic prophage at this locus more degraded than e14.

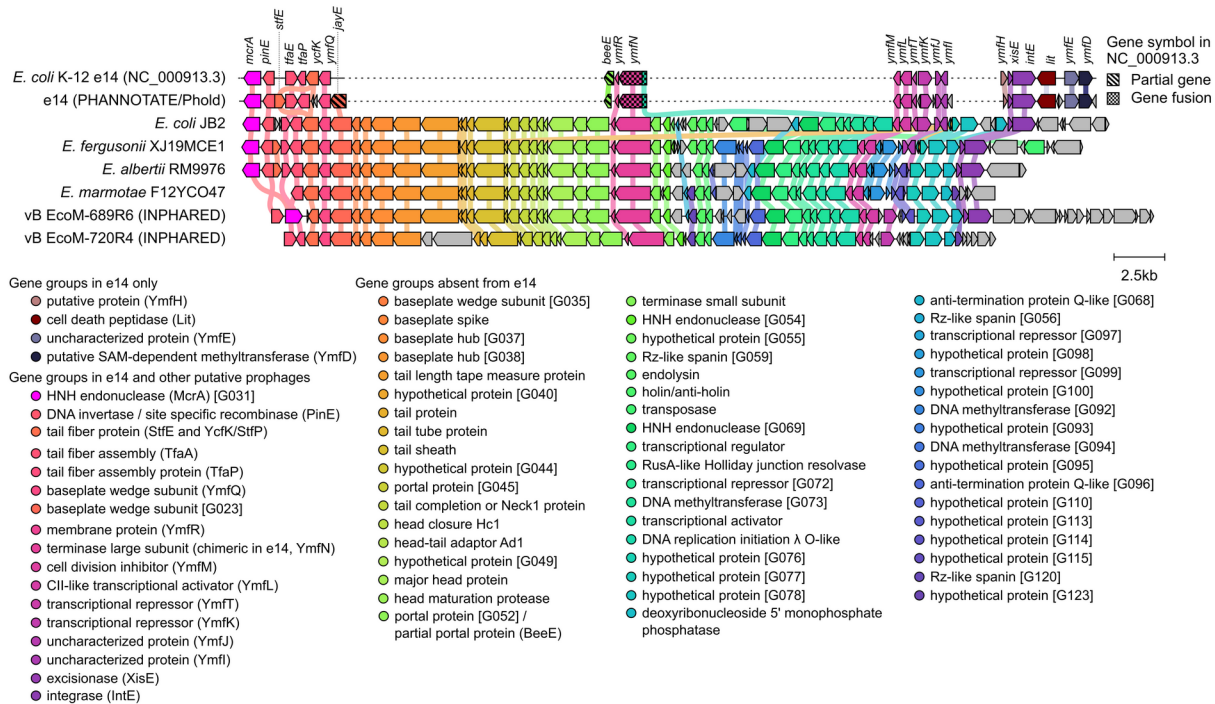


Figure 1: Protein-level comparison of *E. coli* str. K-12 e14 and related prophages from *Escherichia* reveals extensive loss of genes involved in genes required for productive phage replication.

Proteins sharing a Clinker gene homology group at a 30% amino-acid identity threshold are shown with the same colour. Where different homologous groups are identified by the same product name, the group number is also shown (see Supplementary File 2). The top row indicates gene symbols as labeled in the *E. coli* str. K-12 substr. MG1655 reference sequence (RefSeq accession NC_000913.3), with genes identified as pseudogenes indicated by brackets.

Phylogenetic analysis of six conserved protein families further supports the evolutionary relationship between e14 and larger phage and prophage genomes (Figure 2). The *E. coli* str. K-12 e14 sequence clustered with the prophage from *E. fergusonii*, within a broader group containing the prophages from *E. albertii* and *E. marmotae*, phage vB_EcoM-720R4 and the experimentally characterised temperate phage

SfIV. SfIV is a serotype-converting phage of *Shigella flexneri*, illustrating that this group includes functional temperate phages capable of altering host phenotypes [19]. More distant references included vB_EcoM-689R6 and several other characterised temperate phages. Among these, SfV modifies the *S. flexneri* O-antigen to confer serotype V [20], while the inducible, infectious coliphage cdtI carries genes encoding cytolethal distending toxin and a type III secretion effector [21]. Ayreon, isolated from pond water using an *E. coli* K-12 derivative, shares substantial sequence similarity with phage cdtI but lacks its toxin gene cluster. Its temperate lifestyle has been experimentally demonstrated, but its attachment sequence suggests integration into *prfC*, indicating that related genomes need not share e14's specificity for the *icd* locus [22]. The set also included 2H10, a cultured temperate coliphage originating from the infant gut [23]. These relationships place e14 among relatives that retain functional temperate lifestyles and carry different forms of host-modifying cargo, consistent with its origin through extensive reduction of a more complete temperate phage genome. However, the small number of conserved proteins and the potentially mosaic nature of phage genomes mean that this phylogeny may not capture the evolutionary history of the complete elements [24].

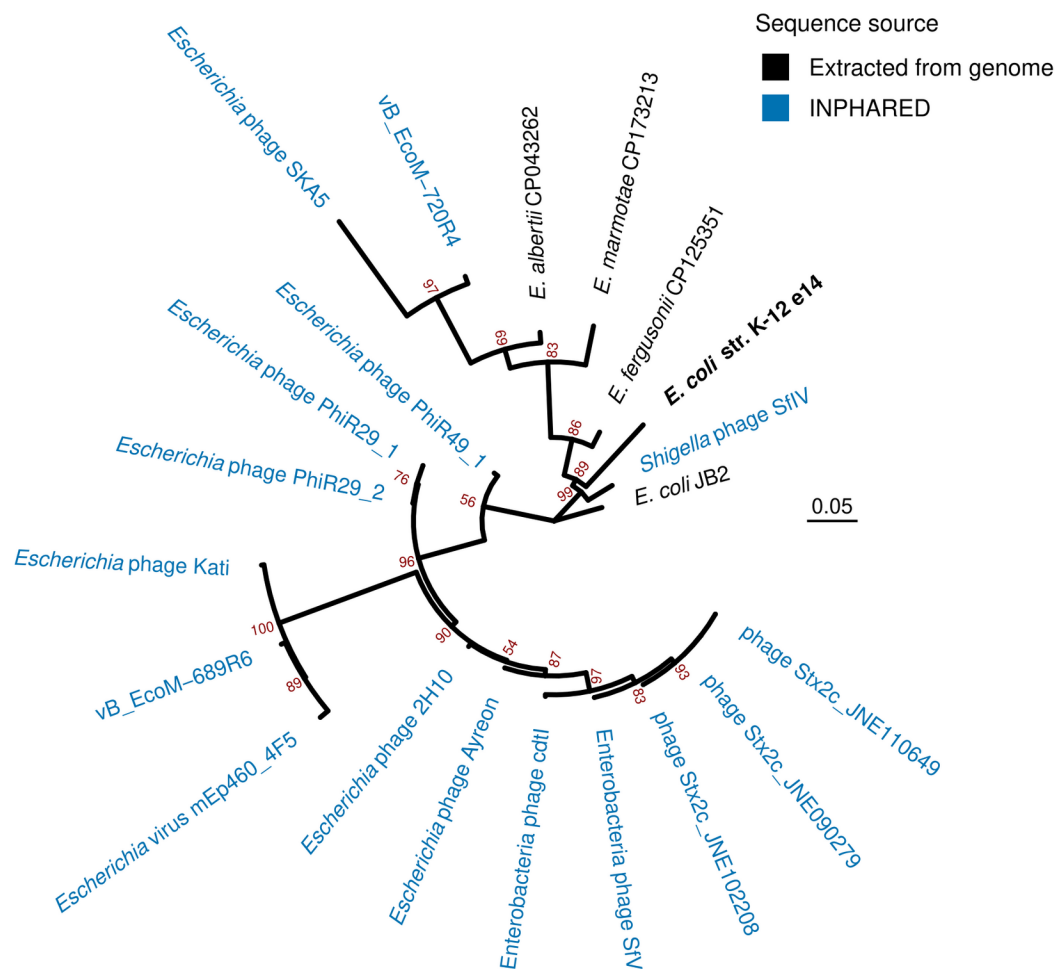


Figure 2: Maximum-likelihood phylogeny places *E. coli* str. K-12 e14 among a wider lineage of e14-like lambdoid phages across *Escherichia*. The phylogeny was reconstructed from a concatenated alignment of six conserved protein families shared with *E. coli* str. K-12 e14. Related prophages and reference bacteriophages were identified using Mash searches against INPHARED. The unrooted tree was inferred using IQ-TREE and subsequently rooted on the *E. coli* JB2 prophage for visualisation only. Node labels indicate ultrafast bootstrap support (%) from 1,000 replicates. The scale bar represents 0.05 substitutions per site.

Comparison with substantially more complete prophages reveals the extent and nature of gene loss by e14 (Figure 1). Whereas the larger elements retain extensive modules required for productive phage replication, including DNA replication and packaging, head and tail morphogenesis, and host lysis, most of these functions are absent from e14. Thus, the functions preferentially lost from e14 are largely those required to complete a lytic reproductive cycle and, ultimately, to kill the host. In contrast, e14 retains a smaller set of genes associated with integration and excision, regulation, and functions that can affect host fitness. This pattern resembles the general trajectory predicted for prophage degradation: genes required primarily for autonomous phage reproduction are preferentially lost, while genes whose functions can benefit the lysogen may persist [25, 26].

Why have the remaining components of e14 persisted after most phage reproductive functions were lost? e14 is unusual among *E. coli* str. K-12 cryptic prophages because it retains functional excision and integration machinery, with SOS induction excising a circular molecule that can subsequently reintegrate at its specific chromosomal attachment site [1, 2]. Complete loss of e14 is sometimes observed during experimental evolution [27]. The persistence of e14 despite its capacity for excision raises the possibility that some of its retained or subsequently acquired functions benefit the host. Cryptic prophage genes in K-12 can increase host fitness under particular environmental conditions, and deletion of e14 has been associated with altered stress resistance and biofilm phenotypes [28]. However, it is unclear how the remnant genes retained in e14 achieve this effect on bacterial phenotype.

One possible contribution to the persistence of e14 is that it may contribute to anti-phage defence, both against related lambdoid phages and unrelated bacteriophages. Several mechanisms could potentially protect against e14-like and other lambdoid phages. First, occupancy of the highly conserved *icd* attachment site could prevent integration by functional e14-like temperate phages that use the same site. Second, several e14 genes may interfere with subsequent lambdoid phage infections. The CI-like repressor YmfK might confer superinfection immunity against related e14-like phages, consistent with its proposed role as the e14 lysogenic repressor [3]. Foldseek analysis identified weak structural matches between the uncharacterised protein YmfE, detected here only in e14, and an uncharacterised protein from *Escherichia* phage vB_EcoM-705R2 (probability of homology = 0.69, E-value = 0.748, sequence identity = 11.6%), and a yet weaker match to the superinfection exclusion protein SieA from the P22-like phage UAB_Phi20 (probability of homology = 0.16, E-value = 8.63, TM-score = 0.331, RMSD = 19.03 Å) [29]. In P22, SieA mediates superinfection exclusion by preventing the injection of phage DNA into the host cell [30]. Although the structural match is weak, a similar role for YmfE remains an intriguing, if speculative, possibility. Another possibility involves YmfN, a chimeric protein comprising the N-terminus of a putative λ O-like DNA replication initiation protein fused to the C-terminal region of a large phage terminase, apparently resulting from a deletion event [3] (Figure 1). In some phages, the C-terminal region of the large terminase mediates interactions with the capsid portal protein [31, 32], while inhibition of terminase-mediated DNA packaging is an established mechanism of phage interference [33]. YmfN could therefore potentially interfere with DNA packaging by related e14-like phages, although this remains speculative.

Two other e14 genes have established roles in defence against unrelated bacteriophages, particularly the T-even phages. *mcrA* encodes a modification-dependent restriction endonuclease that restricts phages containing certain modified cytosines in their DNA [34, 35]. However, *mcrA* was not specific to e14, being present in most e14-like elements examined, including the complete phage vB_EcoM-689R6, but absent from *E. marmotae* and vB_EcoM-720R4. *lit* is well characterised as an inhibitor of bacteriophage T4 infection. It encodes a normally inactive protease that is activated by the Gol peptide, derived from the T4 major capsid protein gp23, during infection. Activated Lit cleaves the essential host translation factor EF-Tu, causing global inhibition of protein synthesis, preventing phage reproduction through host death [36, 37]. In contrast to *mcrA*, *lit* was detected only in e14 and was notably absent from *E. coli* JB2, despite it containing the most closely related e14-like prophage identified in this study. The origins of *lit* are somewhat obscure. Interestingly, *lit* and the neighbouring *ymfD* and *ymfE* genes lie between the attachment site and integrase, and have an unusually low G+C content. This region was previously proposed to have been acquired either by the ancestral phage before integration or by e14 subsequently [3]. Its absence from the other e14-like prophages examined here is consistent with a distinct evolutionary origin. Among *Escherichia*, *lit* was only detected by SmartBLAST to occur within *E. coli*, though homologs occur in other Enterobacteriales. Foldseek identified structural similarity between Lit and a putative

lysogenic conversion protein from Enterobacteria phage P2-EC46 (homology probability = 1, TM-score = 0.561, RMSD = 11.56 Å, E-value = 1.29×10^{-5}) consistent with a shared protein fold, despite low amino-acid sequence identity (14.1%). Further evidence for an anti-phage role comes from the finding that PinQ, a recombinase encoded by the Qin cryptic prophage, inverts a segment of e14 to generate the chimeric protein StfE2, which inhibits T2 adsorption [38].

Together, these findings support an origin of e14 as a degraded temperate bacteriophage, with substantially more complete relatives distributed across multiple *Escherichia* species. Its reduced genome reflects extensive loss of genes required for phage reproduction, alongside the retention of genes with potential benefits to the bacterial host. Given that anti-phage defence systems frequently cluster within mobile genetic regions [39, 40], e14 may represent a form of anti-phage defence “flotsam”: a degraded prophage that retains several potential defensive functions and has become a defence-island-like locus. Whether these functions confer protection against related or unrelated phages, and whether they have been selectively retained or subsequently acquired, remains to be established. More broadly, e14 raises the possibility that cryptic prophages are not just evolutionary relics, but may provide anti-phage defence functions. Such remnants could therefore represent an intermediate stage in the evolution of defence islands, in which the remains of past phage infections contribute to protection against future ones.

Data summary

All genome sequences analysed in this study are publicly available from NCBI under the accessions given in the Methods. Scripts used to perform the analyses are provided on GitHub (<https://github.com/dannagifford/e14-like-prophages-escherichia>).

Conflicts of interest

The author declares no conflict of interest.

Acknowledgements

The author thanks Rohan Maddamsetti and Jeffrey Barrick for an online discussion involving e14 that inspired this analysis. This work was supported by the Academy of Medical Sciences (Springboard Grant SBF007/100096), and UKRI-BBSRC (BB/X007979/1).

References

1. Greener A, Hill CW. Identification of a novel genetic element in *Escherichia coli* K-12. *Journal of Bacteriology* 1980;144:312–321.
2. Brody H, Greener A, Hill CW. Excision and reintegration of the *Escherichia coli* K-12 chromosomal element e14. *Journal of Bacteriology* 1985;161:1112–1117.
3. Mehta P, Casjens S, Krishnaswamy S. Analysis of the lambdoid prophage element e14 in the *E. Coli* K-12 genome. *BMC Microbiology* 2004;4:4.
4. Casjens S. Prophages and bacterial genomics: What have we learned so far? *Molecular Microbiology* 2003;49:277–300.
5. Hu Z, Chen D, Shi T, Huang Y, Cui X, *et al.* Antimicrobial resistance of Enterobacteriaceae in rabbit farms: An underestimated reservoir harboring *mcr-1.1*. *Frontiers in Cellular and Infection Microbiology* 2025;15:1663852.
6. López-Díaz M, Bleriot I, Pacios O, Fernández-García L, Blasco L, *et al.* Molecular characteristics of phages located in carbapenemase-producing *Escherichia coli* clinical isolates: New phage-like plasmids. *bioRxiv*. Epub ahead of print 2022. DOI: 10.1101/2022.06.29.498070.
7. Bouras G, Nepal R, Houtak G, Psaltis AJ, Wormald P-J, *et al.* Pharokka: A fast scalable bacteriophage annotation tool. *Bioinformatics* 2023;39:btac776.
8. Bouras G, Grigson SR, Mirdita M, Heinzinger M, Papudeshi B, *et al.* Protein structure-informed bacteriophage genome annotation with Phold. *Nucleic Acids Research* 2026;54:gkaf1448.
9. Gilchrist CLM, Chooi Y-H. Clinker & clustermap.js: Automatic generation of gene cluster comparison figures. *Bioinformatics* 2021;37:2473–2475.

10. Kempen M Van, Kim SS, Tumescheit C, Mirdita M, Lee J, *et al.* Fast and accurate protein structure search with Foldseek. *Nature Biotechnology* 2024;42:243–246.
11. Kim RS, Pimenova O, Levy Karin E, Mirdita M, Steinegger M. BFVD v3–UniProt-complete, improved viral protein structure predictions. *bioRxiv*. Epub ahead of print 2026. DOI: 10.64898/2026.09.16.752260.
12. Cook R, Brown N, Redgwell T, Rihtman B, Barnes M, *et al.* INfrastructure for a PHAge REference Database: Identification of large-scale biases in the current collection of cultured phage genomes. *PHAGE* 2021;2:214–223.
13. Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, *et al.* Mash: Fast genome and metagenome distance estimation using MinHash. *Genome Biology* 2016;17:132.
14. McNair K, Zhou C, Dinsdale EA, Souza B, Edwards RA. PHANOTATE: A novel approach to gene identification in phage genomes. *Bioinformatics* 2019;35:4537–4542.
15. Steinegger M, Söding J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature Biotechnology* 2017;35:1026–1028.
16. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution* 2013;30:772–780.
17. Steenwyk JL, Buida TJ, Li Y, Shen X-X, Rokas A. ClipKIT: A multiple sequence alignment trimming software for accurate phylogenomic inference. *PLOS Biology* 2020;18:e3001007.
18. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, *et al.* IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution* 2020;37:1530–1534.
19. Jakhetia R, Talukder KA, Verma NK. Isolation, characterization and comparative genomics of bacteriophage SfIV: A novel serotype converting phage from *Shigella flexneri*. *BMC Genomics* 2013;14:677.
20. Allison GE, Angeles D, Tran-Dinh N, Verma NK. Complete genomic sequence of SfV, a serotype-converting temperate bacteriophage of *Shigella flexneri*. *Journal of Bacteriology* 2002;184:1974–1987.
21. Asakura M, Hinenoya A, Alam MS, Shima K, Zahid SH, *et al.* An inducible lambdoid prophage encoding cytolethal distending toxin (Cdt-I) and a type III effector protein in enteropathogenic *Escherichia coli*. *Proceedings of the National Academy of Sciences* 2007;104:14483–14488.
22. Vlot M, Nobrega FL, Wong CFA, Liu Y, Brouns SJJ. Complete genome sequence of the *Escherichia coli* phage Ayreon. *Genome Announcements* 2018;6:e01354–17.
23. Mathieu A, Dion M, Deng L, Tremblay D, Moncaut E, *et al.* Virulent coliphages in 1-year-old children fecal samples are fewer, but more infectious than temperate coliphages. *Nature Communications* 2020;11:378.
24. Hendrix RW, Smith MCM, Burns RN, Ford ME, Hatfull GF. Evolutionary relationships among diverse bacteriophages and prophages: All the world's a phage. *Proceedings of the National Academy of Sciences* 1999;96:2192–2197.
25. Bobay L-M, Touchon M, Rocha EPC. Pervasive domestication of defective prophages by bacteria. *Proceedings of the National Academy of Sciences* 2014;111:12127–12132.
26. Khan A, Burmeister AR, Wahl LM. Evolution along the parasitism-mutualism continuum determines the genetic repertoire of prophages. *PLOS Computational Biology* 2020;16:e1008482.
27. Maddamsetti R, Orman M. Parallel evolution of chimeric genes in microbial evolution experiments. *bioRxiv*. Epub ahead of print 2026. DOI: 10.64898/2026.06.14.732165.
28. Wang X, Kim Y, Ma Q, Hong SH, Pokusaeva K, *et al.* Cryptic prophages help bacteria cope with adverse environments. *Nature Communications* 2010;1:147.
29. Bardina C, Colom J, Spricigo DA, Otero J, Sánchez-Osuna M, *et al.* Genomics of three new bacteriophages useful in the biocontrol of *Salmonella*. *Frontiers in Microbiology* 2016;7:545.

30. Leavitt JC, Woodbury BM, Gilcrease EB, Bridges CM, Teschke CM, *et al.* Bacteriophage P22 SieA-mediated superinfection exclusion. *mBio* 2024;15:e02169–23.
31. Yeo A, Feiss M. Specific interaction of terminase, the DNA packaging enzyme of bacteriophage lambda, with the portal protein of the prohead. *Journal of Molecular Biology* 1995;245:141–150.
32. Lokareddy RK, Sankhala RS, Roy A, Afonine PV, Motwani T, *et al.* Portal protein functions akin to a DNA-sensor that couples genome-packaging to icosahedral capsid maturation. *Nature Communications* 2017;8:14310.
33. Boyd CM, Seed KD. A phage satellite manipulates the viral DNA packaging motor to inhibit phage and promote satellite spread. *Nucleic Acids Research* 2024;52:10431–10446.
34. Raleigh EA, Trimarchi R, Revel H. Genetic and physical mapping of the *mcrA* (*rgIA*) and *mcrB* (*rgIB*) loci of *Escherichia coli* K-12. *Genetics* 1989;122:279–296.
35. Hiom K, Sedgwick SG. Cloning and structural characterization of the *mcrA* locus of *Escherichia coli*. *Journal of Bacteriology* 1991;173:7368–7373.
36. Kao C, Snyder L. The *lit* gene product which blocks bacteriophage T4 late gene expression is a membrane protein encoded by a cryptic DNA element, e14. *Journal of Bacteriology* 1988;170:2056–2062.
37. Yu YT, Snyder L. Translation elongation factor Tu cleaved by a phage-exclusion system. *Proceedings of the National Academy of Sciences* 1994;91:802–806.
38. Kirigo J, Huelgas-Méndez D, Tomás M, Benedik MJ, García-Contreras R, *et al.* Adsorption of phage T2 is inhibited due to inversion of cryptic prophage DNA by the serine recombinase PinQ. *Nucleic Acids Research* 2025;53:gkaf1041.
39. Makarova KS, Wolf YI, Snir S, Koonin EV. Defense islands in bacterial and archaeal genomes and prediction of novel defense systems. *Journal of Bacteriology* 2011;193:6039–6056.
40. Hochhauser D, Millman A, Sorek R. The defense island repertoire of the *Escherichia coli* pan-genome. *PLOS Genetics* 2023;19:e1010694.