

Lawrence Berkeley National Laboratory

LBL Publications

Title

Cross-family and phage-specific gene requirements for Klebsiella infection revealed by scalable RB-TnSeq genetic screens.

Permalink

<https://escholarship.org/uc/item/9294x99z>

Journal

PLOS Genetics, 22(7)

ISSN

1553-7390

Authors

Gittrich, Marissa R
Sanderson, Courtney M
Noel, Cara M
[et al.](#)

Publication Date

2026-07-01

DOI

10.1371/journal.pgen.1012233

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

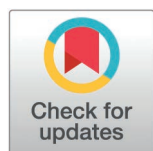
RESEARCH ARTICLE

Cross-family and phage-specific gene requirements for *Klebsiella* infection revealed by scalable RB-TnSeq genetic screens

Marissa R. Gittrich^{1,2}, Courtney M. Sanderson^{1,2}, Cara M. Noel¹, Erica Babusci³, Sumeyra C. Selbes⁴, Madeline Svab^{5,6,7}, Isabella Murray^{5,6,7}, Collis Bousliman^{5,8}, Ami Fofana¹, Aghiad Daboul¹, Jonathan Leopold¹, Alessandra Gonçalves de Melo⁹, Marion Urvoy¹, Sylvain Moineau⁹, Vivek K. Mutalik^{10*}, Matthew B. Sullivan^{1,2,10*}

1 Department of Microbiology, The Ohio State University, Columbus, Ohio, United States of America, **2** Center of Microbiome Science, The Ohio State University, Columbus, Ohio, United States of America, **3** School of the Environment and Natural Resources, The Ohio State University, Columbus, Ohio, United States of America, **4** Department of Psychology, The Ohio State University, Columbus, Ohio, United States of America, **5** Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, California, United States of America, **6** Molecular and Cell Biology Department, University of California Berkeley, Berkeley, California, United States of America, **7** Plant and Microbial Biology Department, University of California Berkeley, Berkeley, California, United States of America, **8** Department of Bioengineering, University of California Berkeley, Berkeley, California, United States of America, **9** Département de biochimie, de microbiologie et de bio-Informatique, Université Laval, Quebec, Canada, **10** Department of Civil, Environmental, and Geodetic Engineering, The Ohio State University, Columbus, Ohio, United States of America

* sullivan.948@osu.edu (MBS), vkmutalik@lbl.gov (VKM)



OPEN ACCESS

Citation: Gittrich MR, Sanderson CM, Noel CM, Babusci E, Selbes SC, Svab M, et al. (2026) Cross-family and phage-specific gene requirements for *Klebsiella* infection revealed by scalable RB-TnSeq genetic screens. PLoS Genet 22(7): e1012233. <https://doi.org/10.1371/journal.pgen.1012233>

Editor: Shumin Tan, Tufts University School of Medicine, UNITED STATES OF AMERICA

Received: November 10, 2025

Accepted: July 1, 2026

Published: July 20, 2026

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pgen.1012233>

Copyright: © 2026 Gittrich et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution,

Abstract

Bacteriophages are being cataloged at an accelerating pace and are recognized as key players in nutrient and energy cycling across ecosystems. Yet the bacterial genetic determinants that govern phage-host specificity and infection success remain poorly understood, particularly in clinically and ecologically important genera such as *Klebsiella* where prior receptor characterization has been almost entirely limited to capsulated strains. Here we used a randomly barcoded, genome-wide, loss-of-function transposon mutant library (RB-TnSeq) of *Klebsiella* sp. M5a1, a naturally acapsular, nitrogen-fixing rhizobacterium, to generate the first systematic, cross-family map of phage receptor gene dependencies in *Klebsiella*. Challenging the library against 25 double-stranded DNA phages spanning five families in 213 parallel assays, we identified 42 bacterial genes associated with phage infection, of which 15 had no prior association with phage infection in any bacterial system. Disruption of surface receptor biosynthesis genes conferred cross-resistance across multiple phage families, while intracellular gene disruptions had predominantly phage-specific effects. Clonal validation of eight genes confirmed LPS outer core biosynthesis genes as primary receptor determinants alongside additional host factors spanning outer membrane transport, cofactor biosynthesis, and two-component signaling. Comparative analysis across all 25 phages revealed that phage genus rather than family is the

and reproduction in any medium, provided the original author and source are credited.

Data availability statement: All 25 phage genome sequences are available from NCBI under accession numbers PP197198, PP197344, PP202993, PP444685, PP444686, PP444687, PP444688, PP505786, PP505787, PP526028, PP526029, PP554336, PP554337, PP554338, PP556864, PP556865, PP582754, PP582755, PP582756, PP582757, PP582758, PP597382, PP597383, PP597384, and PX412413. All RB-TnSeq fitness data underlying the figures and results presented in this manuscript are provided in S1 and S2 Tables. The processed data and analysis code used to generate the figures are publicly available on Figshare at 10.6084/m9.figshare.32928455. Phage stocks are available from the Félix d'Hérelle Reference Center for bacterial viruses at Université Laval.

Funding: Funding was provided by the U.S. Department of Energy (<https://www.energy.gov/>; award #DE-SC0020173 and #DE-SC0023307) to M.B.S. V.K.M. was also supported by the Biopreparedness Research Virtual Environment (BRaVE) Phage Foundry at Lawrence Berkeley National Laboratory by the U.S. Department of Energy, Office of Science, Office of Biological & Environmental Research under contract number DE-AC02-05CH11231. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

stronger predictor of host gene dependency profiles, a finding with direct implications for the functional annotation of uncharacterized phage isolates and rational phage cocktail design. Together, these findings provide a community resource for linking phage genomic diversity to functional host interaction space in this ecologically and clinically important genus.

Author summary

Bacteriophages, or phages, are viruses that infect bacteria and play central roles in shaping microbial communities across ecosystems. Despite their ecological importance, the bacterial genes that determine infection outcomes remain poorly understood. Using a scalable barcoded transposon sequencing (RB-TnSeq) approach, we mapped host gene requirements for infection by 25 diverse phages of *Klebsiella* sp. M5a1, a soil-associated plant-growth-promoting bacterium. We identified 42 bacterial genes associated with phage infection, spanning surface receptors, transcriptional regulators, and metabolic and protein-folding pathways. Infection strategies broadly clustered by phage genus and family, while phage-specific differences arose primarily in intracellular processes such as transcription and protein folding. These findings reveal both conserved and phage-specific host interactions that define infection strategies and establish a scalable framework for linking phage genomic diversity to function, with practical implications for predicting phage host range, designing phage cocktails for therapy, and understanding phage-driven dynamics in natural microbial communities.

Introduction

Phages, viruses that infect bacteria, have been increasingly recognized as key players in ecological and evolutionary processes across Earth's ecosystems [1–9]. In terrestrial environments, phages have been found to impact nutrient cycling [10, 11], shape bacterial community composition [12–14], and impact plant health [15, 16]. Despite over a century of phage research and growing recognition of their ecological importance, modeling phage-host interactions remains challenging due to limited studies of these interactions compared to the abundance and diversity of phages [17–19]. Even the basic question of “who infects whom” is often difficult to answer as (i) nearly genetically identical phages can exhibit drastically different host-range and infectivity at the bacterial strain level [20–25], and (ii) infectivity can change based on environmental conditions [25–27].

Mechanistically, some phage susceptibility differences can be explained by variations in bacterial receptors and in phage receptor binding proteins [20, 21, 28], presence or absence of host surface proteins [22], or phage defense elements like restriction-modification and CRISPR-Cas systems [23]. However, these factors do not fully explain the substantial variability in susceptibility observed across strains

and phages. In particular, the broader and strain-specific genetic determinants that modulate phage infection beyond well-characterized receptors and systems remain incompletely defined [24, 25]. In addition, there is a large discrepancy between the number of culturable phage-host model systems available for experimental testing (maybe hundreds to thousands) and the uncultivated genomic fragments known to exist (>15 million in IMG-VR) [29]. These genomic fragments represent only a small subset of global phage numbers thought to exist planet-wide ($\sim 10^{31}$) [30]. Critically, even the best *in silico* host prediction approaches are limited to predicting phage hosts to the bacterial genus level for uncultivated phage groups [31, 32], which represents a significant bottleneck in phage-driven microbiome design efforts [32]. These knowledge gaps in mechanistic understanding and understudied taxonomic variation in phage-host interactions make accurately predicting phage infectivity at the strain level challenging.

Recent advances in genome-wide functional genomics offer a scalable path toward resolving these gaps in phage-host interaction biology. Genome-wide transposon mutagenesis enables systematic interrogation of bacterial genes non-essential for growth under defined conditions and has been applied to characterize host factors involved in phage infection [33–36]. The combination of this approach with high-throughput randomly-barcoded transposon-site sequencing (RB-TnSeq), enables cost-effective, parallel fitness assessment of more than 100,000 transposon mutants in a single pooled experiment, quantifying the contribution of each disrupted gene to bacterial survival under a given selective pressure, including phage challenge [37–40]. Bacterial genes required for successful phage infection are identified through the selective enrichment of the mutated strains after phage exposure. This enrichment can also be quantified as a fitness advantage for the mutant attributed to that gene. Unlike traditional single-gene knockouts, RB-TnSeq provides genome-scale resolution of key host factors involved in phage infection pathways in a single sequencing run, substantially accelerating the systematic characterization of host factors governing phage infection.

RB-TnSeq has already revealed novel insights into phage biology, even in well-studied systems like *Escherichia coli* and *Salmonella enterica* [39, 40]. These studies have uncovered previously unrecognized host genes involved in phage susceptibility and highlighted genes conferring cross-resistance to diverse phages, such as genes involved in surface structures, transcriptional regulation, and stress responses. In one notable case, RB-TnSeq in *E. coli* revealed that cyclic di-GMP regulates secretion of the glycan receptor for phage N4, overturning a thirty-year-old model of phage adsorption and explaining the broad host range of N4-like phages [39, 41–43]. Most recently, while this manuscript was under review, Moriniere et al. (2026) [44] extended RB-TnSeq approach to an unprecedented scale, conducting genome-wide genetic screens across 255 taxonomically diverse *E. coli* phages and assigning receptor specificity to 193 phages across 19 receptor classes. By integrating these data with comparative genomics and machine learning, the authors demonstrated that phage receptor identity can be predicted from genome sequence alone, illustrating the predictive power that large-scale RB-TnSeq datasets can unlock.

Despite its clinical prominence as a leading cause of antibiotic-resistant infections, *Klebsiella* remains poorly characterized at the level of phage receptor biology [45]. Our understanding of *Klebsiella* phage-host interactions has been shaped almost entirely by studies of capsulated clinical strains, in which the capsular polysaccharide (CPS) serves as the dominant surface feature and principal phage receptor [33, 36, 46–48]. Recently, efforts have begun to systematically catalogue *Klebsiella* phage diversity, including the *Klebsiella* Phage Collection [49], as well as our prior work characterizing 24 *Klebsiella* phages that are publicly available through Université Laval [50], reflecting growing recognition of the need for organized phage resources in this genus. However, decades of research have uncovered only a handful of capsule-independent phage receptors [33, 36], and there has been no systematic study on experimentally mapping a diverse phage panel to different host factor determinants. This capsule-dominated view of *Klebsiella* phage receptor biology has failed to account for the genotype that emerges most frequently under phage selection pressure. Under phage predation, *K. pneumoniae* most frequently evolves resistance by shedding or modifying its capsule, with more than 97% of phage-resistant mutants carrying CPS biosynthesis mutations in controlled experiments [48]. Capsule loss can also arise reversibly without any genetic mutation [51]. These acapsular escape variants are no longer susceptible to capsule-targeting

phages and instead become susceptible to phages that use secondary receptor structures, namely LPS outer core and outer membrane porins. Characterizing receptors in acapsular strains therefore directly reveals the phage infection landscape relevant to environmental *Klebsiella* populations, where capsule-independent host-phage interactions are likely to predominate.

To address this gap, we selected *Klebsiella* sp. M5a1, a naturally non-capsulated, nitrogen-fixing soil bacterium relevant to nutrient cycling and plant root colonization [52–54], as the experimental host. We hypothesized that the absence of capsular structure would reveal the underlying LPS- and porin-dependent receptor landscape that is obscured by capsule signals in capsulated strain screens and can serve as a model system for systematically characterizing secondary or terminal receptors. As a demonstration, using this host, we assembled a collection of 25 phages spanning five families, capturing substantial genomic and morphological diversity. We challenged the *Klebsiella* sp. M5a1 RB-TnSeq library against all 25 phages in parallel, identified 42 bacterial genes associated with phage infection, and generated the first systematic, cross-family map of receptor gene dependencies across a diverse panel of *Klebsiella* phages.

Results and discussion

Challenging the *Klebsiella* sp. M5a1 RB-TnSeq library against 25 phages

We challenged the *Klebsiella* sp. M5a1 RB-TnSeq library against 25 phages [50] spanning five families, including the three most prevalent among *Klebsiella* phages, *Autographiviridae* (n=3), *Drexelvriidae* (n=2), *Straboviridae* (n=13), as well as two recently described novel families Ca. *Mavericviridae* (n=3) and Ca. *Rivulusviridae* (n=4) (Fig 1A). RB-TnSeq experiments were conducted as separate, single phage challenges at three phage-to-bacteria ratios (Methods, multiplicity of infection –of MOI 10, 1, and 0.1). Strain fitness was assessed from the log2 barcode abundance changes from the 4.5 hours post-infection relative to the initial library composition (“T0”) with individual strain fitness values averaged across all barcode abundances in a given gene to yield a gene-level fitness score (Fig 1B). To minimize false positives, we restricted analyses to genes meeting three criteria: (1) a fitness score of at least 4, representing the log₂ fold-change in mutant abundance relative to the no-phage control; (2) a t-like statistic of at least 5, reflecting consistency of the fitness estimate across independent barcodes; and (3) an estimated standard error no greater than 2, indicating high confidence in the calculated value. These thresholds were applied as previously described [39]. Loss-of-function mutants meeting these criteria and exhibiting increased fitness were interpreted as conferring partial or complete resistance to phage infection.

This experimental and analytical approach was then used to explore the bacterial genes required for phage infection across 213 genome-wide assays, identifying 42 bacterial genes, with each phage requiring between 1 and 15 (median=3) genes. The 42 genes span nine functional categories such as surface structures, inner membrane proteins, transcriptional regulators, carbon metabolism, energy cycling, amino acid biosynthesis and transport, protein transport, protein modification, and hypothetical proteins (Fig 2). Of these, 25 genes encode surface-associated factors and are likely involved in phage adsorption, while the remaining 17 encode intracellular functions. All 25 phages produced at least one high-scoring receptor hit, and eight had hits for more than one putative receptor (Fig 2).

Connecting phage taxonomy to host gene dependency profiles

The scale of this genome-wide screen, encompassing 25 phages across five families, enabled us to ask a question that single phage-host studies cannot address; specifically, whether infection strategies are conserved within phage taxonomic groups or diverge even among closely related phages. To address this, we examined the extent to which RB-TnSeq fitness profiles cluster by phage taxonomy at the family and genus levels. Visual inspection revealed clear clustering by phage genus, with phage families also forming distinguishable groups (Fig 3A). Using the RB-TnSeq matrix of the high-scoring fitness genes, we generated Bray-Curtis dissimilarity matrices and visualized the data using Principal Coordinates Analysis (PCoA) and Non-metric Multidimensional Scaling (NMDS) ordination methods (Fig 3B and 3C). Pairwise

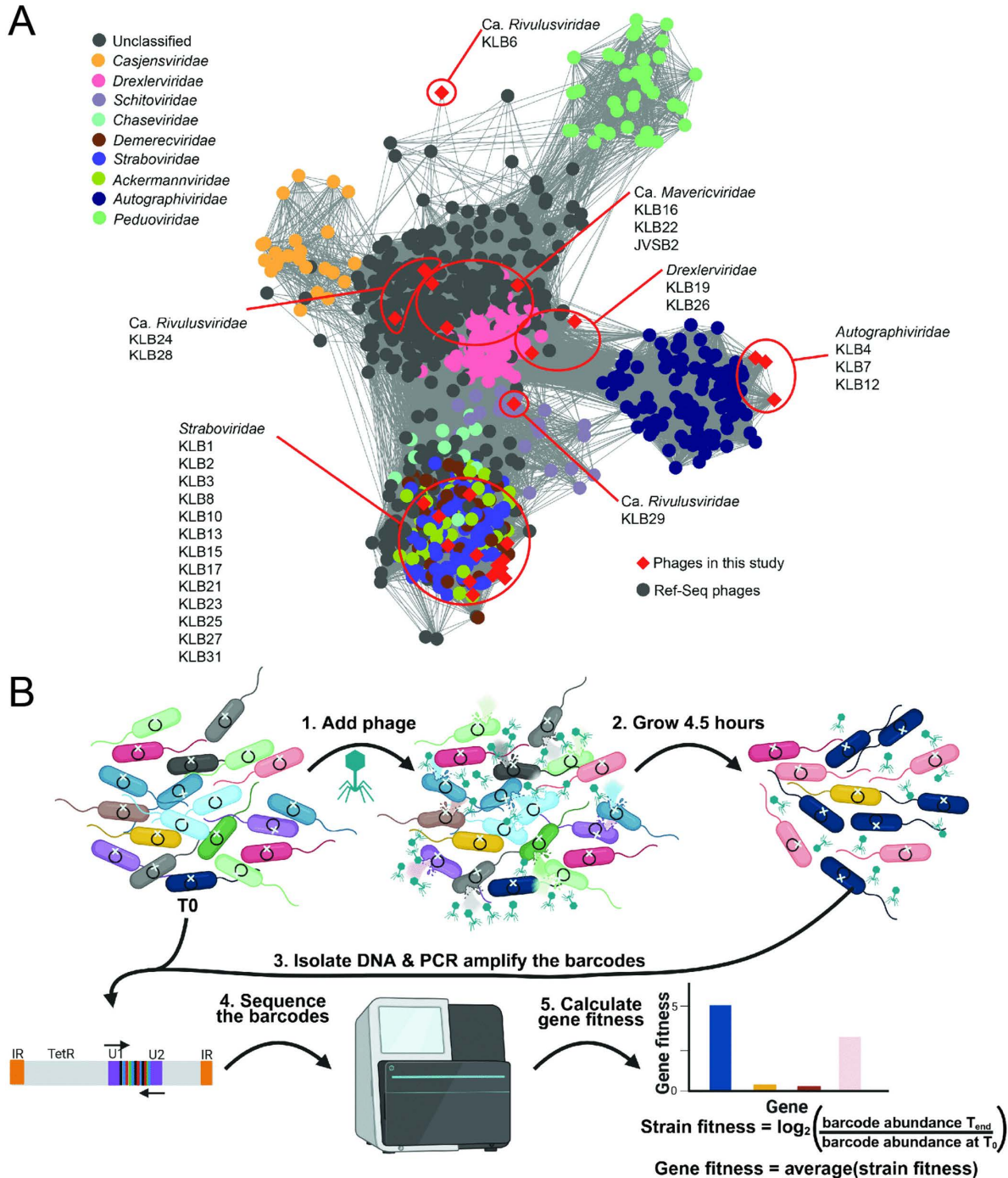


Fig 1. Overall study design. (A) Taxonomic diversity, determined through 'vConTACT3' gene sharing network analyses, for the 25 phages used in this study (red diamonds, and "KLB" named phages listed under each family name) compared to the 942 double-stranded DNA phages infecting *Klebsiella*, *Salmonella*, and *Escherichia* available from NCBI's RefSeq (version 220, 1/10/24). Lines represent detectable gene content similarity between the phages, nodes represent phage genomes, and the phages are color-coded according to family (see legend). (B) Overview of the RB-TnSeq experiment. RB-TnSeq library is initially barcode-sequenced to establish starting mutant frequencies, then challenged with one of the 25 phages, incubated in a plate reader, and the resultant mutant frequencies are estimated through barcode-sequencing. For each transposon knockout, the strain-level fitness is calculated by comparing the barcode abundance at the end of the infection to the initial abundance at T0 (see equation in the figure). At least three transposon knockouts or strains were required to establish each gene's average fitness. Created in BioRender. Gittrich, M. (2026) <https://BioRender.com/zh7126q>.

<https://doi.org/10.1371/journal.pgen.1012233.g001>

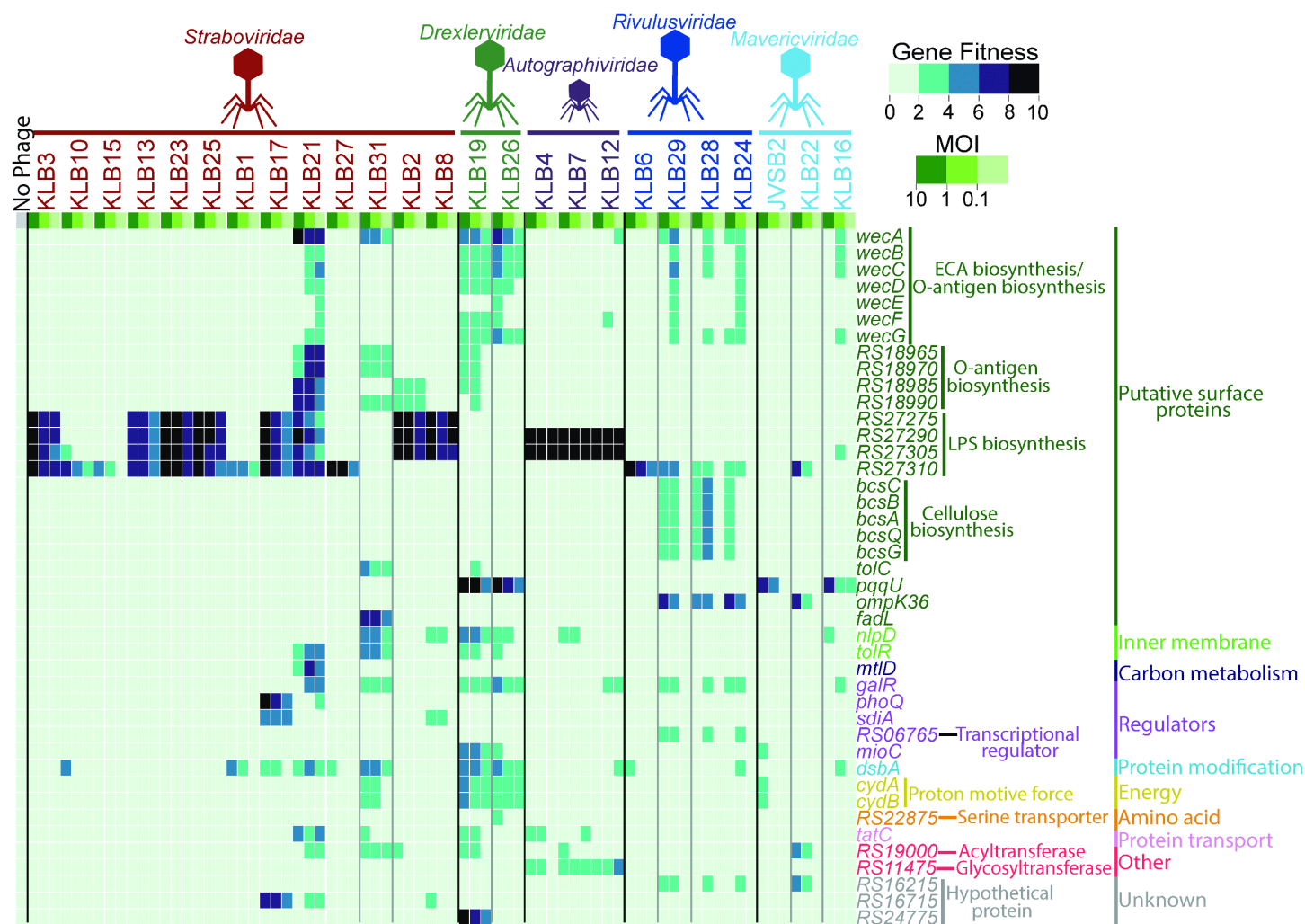


Fig 2. An RB-TnSeq mutant pool's gene fitness against diverse phage infections. For each assay, a subset of a *Klebsiella* sp. M5al RB-TnSeq mutant pool was infected by a single phage (25 phages in total) at a multiplicity of infection (MOI) of 0.1, 1, or 10 (Methods). We also included 'no-phage' culture where the SM buffer was mixed with the library in place of phage. Gene fitness was then calculated as described above (Fig 1B), showing high-confidence bacterial genes (a gene fitness score of ≥ 4). Each experiment was done in duplicate, and each gene had >3 unique gene knockouts. Vertical lines are used to separate phage families (black lines) or phage genera (gray lines). The underlying data for this figure can be found in [S1–S2 Tables](#). Created in BioRender. Gittrich, M. (2026) <https://BioRender.com/wpqy65>.

<https://doi.org/10.1371/journal.pgen.1012233.g002>

PERMANOVA analyses confirmed that phages clustered significantly according to their taxonomy at the family and genus levels (S3–S5 Tables) with one exception at the genus level that is likely attributable to the high genomic similarity between the two genera in question [47].

At the family level, all tested families exhibited significant clustering ($p_{\text{adjusted}} < 0.05$), indicating that phage families differ consistently in their host gene usage patterns. Where multiple genera were available for testing, genus-level comparisons consistently showed stronger statistical support and the highest R^2 values than the corresponding family-level comparisons (S3–S5 Tables), demonstrating that genus is a stronger predictor of host gene requirements for phage infection than family. Overall, our results suggest that phage genus membership carries quantifiable predictive information about infection strategy, and phylogenetic classification at genus resolution can inform host dependency modeling.

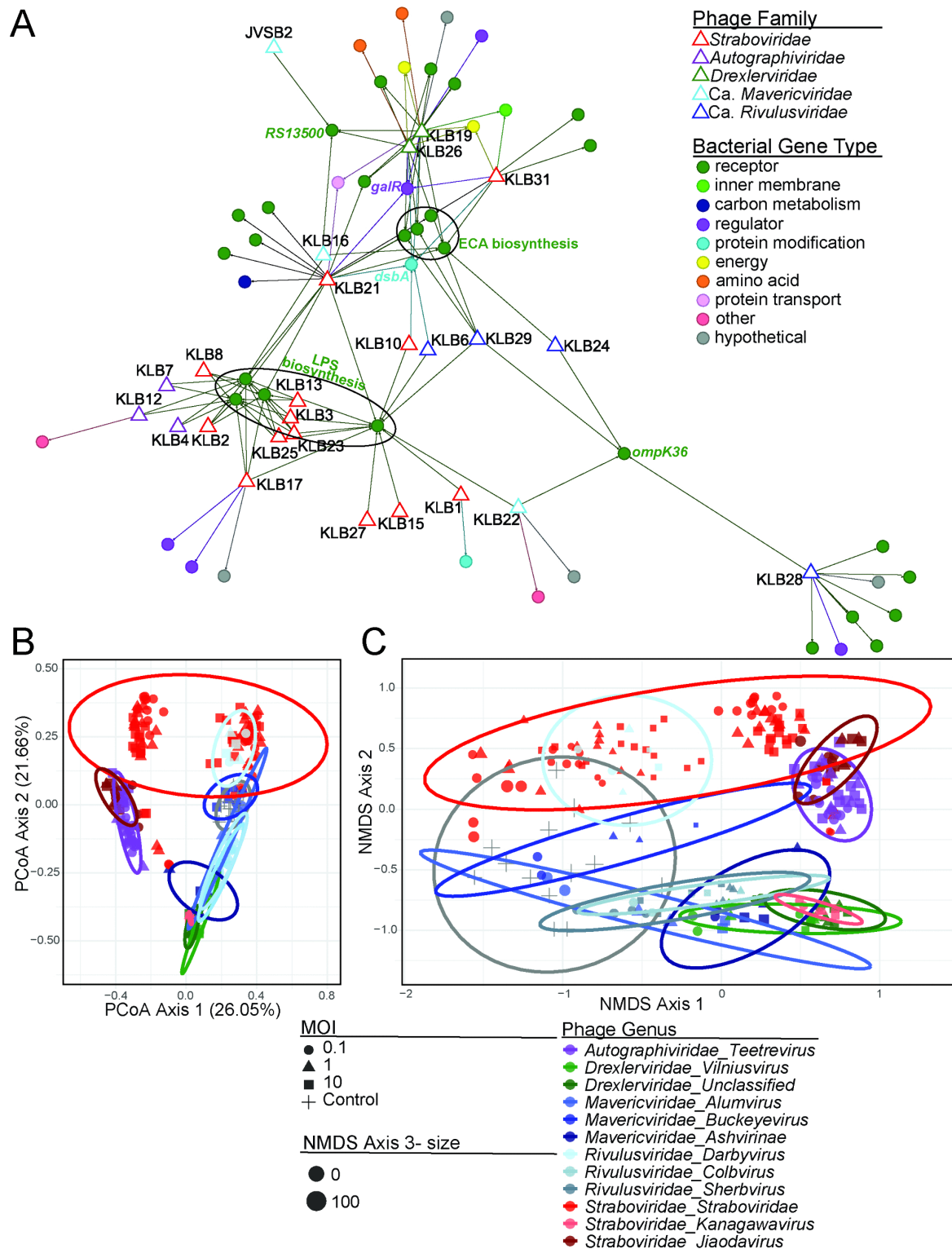


Fig 3. Patterns of phage infection. (A) Summary of bacterial gene requirement patterns between phages observed in our RB-TnSeq assays. Mixed-node network graph shows connections between phages (open triangles) and bacterial genes (solid circles). Gene nodes are colored by function (see

legend), and bacterial genes that provide broad cross-resistance (i.e., to four or more phages) are labeled. (B) Principal Coordinates Analysis (PCoA) and (C) Non-metric Multidimensional Scaling (NMDS) plots based on Bray-Curtis dissimilarity of fitness profiles across samples. Each point represents a sample, colored by phage genus and shaped according to the multiplicity of infection (MOI). Ellipses indicate 95% confidence intervals for each genus. In the NMDS plot, point size corresponds to variation along the third NMDS axis.

<https://doi.org/10.1371/journal.pgen.1012233.g003>

The divergence between genus-level and family-level predictive power was not uniform across all groups. Within the *Straboviridae*, *Slopekvirus* phages share the majority of their genome content divided into two distinct receptor dependency profiles, a pattern not resolvable at the family level. This suggests that within well-established families, genus-level resolution is necessary to capture meaningful variation in infection strategy.

Whether this relationship between taxonomic resolution and predictive power generalizes across broader *Klebsiella* phage diversity, and whether the novel families Ca. *Mavericviridae* and Ca. *Rivulusviridae* show distinct scaling behavior given their greater phylogenetic distance from the established families, remains to be determined as collections of sequenced isolates in these groups expand. For now, the data support genus as the important unit of functional prediction within the families represented here, while acknowledging that the optimal taxonomic resolution for predicting infection strategy may vary across the broader phage diversity infecting this host.

Receptor gene dependencies and intra-genus diversification

The majority of the 25 putative receptor hits encode genes involved in the biosynthesis of conserved surface structures, consistent with prior observations that phages preferentially target abundant, evolutionarily conserved surface features [55]. Among the surface-associated hits, the largest groups were enterobacterial common antigen biosynthesis (ECA) (n=7), cellulose biosynthesis (n=5), LPS biosynthesis (n=4), and genes involved in both LPS and ECA biosynthesis (n=4), alongside the individual loci *tolC*, *ompK36*, *fadL*, *pqqU* and two poorly characterized genes, RS16715 and RS24775 (Fig 2). RS16715 encodes a protein containing a DUF535 domain with homology to VirK, a family associated with outer membrane integrity and virulence in gram-negative bacteria, while RS24775 encodes a predicted inner membrane protein with homology to YqjF, a poorly characterized protein of unknown function. The roles of these two genes in phage infection have not been previously described and represent novel candidates for future characterization. Four of these groups, ECA, LPS, *ompK36*, and *pqqU* conferred cross-resistance to four or more phages across at least two phage families (Fig 2). This finding is consistent with previous studies showing that bacterial genes involved in phage adsorption such as those involved in LPS biosynthesis often provide cross-resistance across multiple phage families, and with some phages binding to more than one receptor [39, 40].

Enterobacterial common antigen (ECA) biosynthesis genes represented the largest single receptor-associated group in the screen, with seven genes required across multiple phages spanning at least two families. ECA is a conserved glycolipid present on the outer membrane of virtually all Enterobacteriaceae [56], and ECA biosynthesis genes have been identified as phage receptor determinants in genome-wide screens in *E. coli* and *Salmonella* [43], though its role in *Klebsiella* phage infection had not been previously described. The ECA-dependent phages in this dataset span more than one family, suggesting that exploitation of this surface structure has arisen independently across phage lineages, consistent with the general principle that abundant, conserved surface structures are convergently targeted by diverse phage groups [55]. Several bacterial mutants in both ECA and LPS biosynthesis pathways showed elevated fitness scores for the same phages, raising the possibility that these phages can engage either surface structure and that receptor redundancy may buffer against resistance arising from disruption of a single biosynthesis pathway. Whether these hits reflect direct ECA-dependent binding or dependence on a related glycan produced by the same biosynthetic pathway, as recently described in *E. coli* [43], remains to be determined in *Klebsiella*.

Five genes involved in cellulose biosynthesis also showed high fitness score, all associated with a subset of phages from the same family. Cellulose is an extracellular matrix component produced by many soil-dwelling bacteria, including

Klebsiella species, where it contributes to biofilm formation and surface attachment [57]. Its presence on the bacterial surface makes it a structurally accessible candidate receptor, and cellulose has been identified as a phage receptor in other gram-negative bacteria [58, 59]. Given that *Klebsiella* sp. M5al is a soil isolate in which cellulose production is likely ecologically relevant, these hits may reflect receptor interactions that are particularly important in environmental rather than clinical contexts, where cellulose-producing biofilm communities predominate.

To examine how receptor diversity arises within a single genus, we focused on the *Slopekvirus* genus, the largest group in our collection with ten phages that share 199 of an average of 277 genes per genome. Despite this high genomic similarity, the phages separated into two distinct receptor dependency profiles: all ten required *wabO*, but six additionally required *wabI*, *wabN*, and *wabH*, indicating differences in the specific LPS outer core regions recognized (Fig 4A and 4B).

Comparison of receptor-binding protein sequences offered a structural explanation for this divergence. In *Straboviridae* phages, the long tail fiber tip is responsible for receptor binding [60], and amino acid comparisons of the distal tail fiber

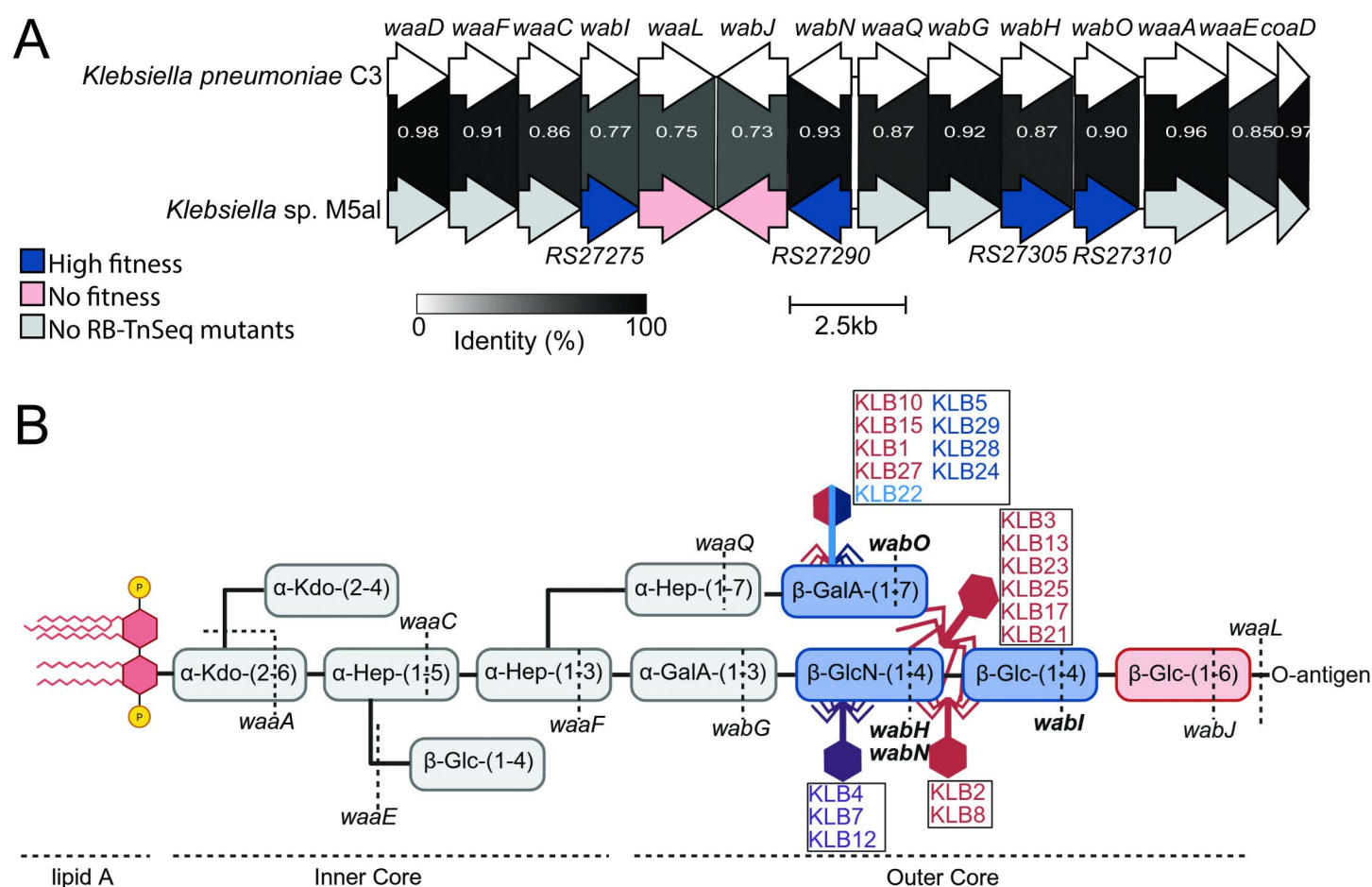


Fig 4. Impact of lipopolysaccharide (LPS) on phage infection. (A) Representation of the three operons involved in LPS biosynthesis. Bacterial genes are colored as follows: putative essential genes (colored gray), genes with high and significant fitness scores (fitness score >4, blue), and genes that did not have significant fitness scores (pink). The gene names RS27275, RS27290, RS27305, and RS27310 were inferred by comparison to the previously characterized strain *K. pneumoniae* C3. Operons were visualized using Clinker. Links between genes denote percent amino acid identity between deduced proteins. (B) Biochemical context showing the role of the LPS operon in the synthesis of the LPS core, and the RB-TnSeq-inferred interaction with phages. Phages are represented at the inferred binding region based on RB-TnSeq data, and color-coded by the phage family. Created in BioRender. Gittrich, M. (2026) <https://BioRender.com/e551yl3>.

<https://doi.org/10.1371/journal.pgen.1012233.g004>

subunit across all ten *Slopekvirus* phages revealed high conservation in the N-terminal approximately 160 residues but substantial divergence in the C-terminal receptor-binding region (Fig 5A and 5B). The two receptor-use profiles clustered by C-terminal sequence identity, and the receptor-binding domain of KLB25 showed greater similarity to the *wabl/wabN/wabH/wabO*-requiring group than to the *wabO*-only group, suggesting potential horizontal acquisition of this region consistent with its LPS binding profile. These data suggest that intra-genus divergence in receptor usage is likely driven by sequence variation in the distal tail fiber, and that subtle genetic variation among closely related phages likely produces functionally distinct host interactions.

Receptor usage and phage genus have emerged as key predictors of phage infectivity at the strain level [28, 61–63]. The present findings reinforce this view, further demonstrating that even phages sharing high genomic similarity within the same genus can rely on distinct receptor structures, with divergence in receptor-binding proteins as the primary driver. This observation aligns with experimental evolution studies showing that minor genetic changes can substantially alter host specificity, including single amino acid substitutions in receptor-binding proteins that redirect phage adsorption to different surface structures [20, 22, 44, 64]. Broad host-range screens consistently support this model across phage families [43, 51, 65–74].

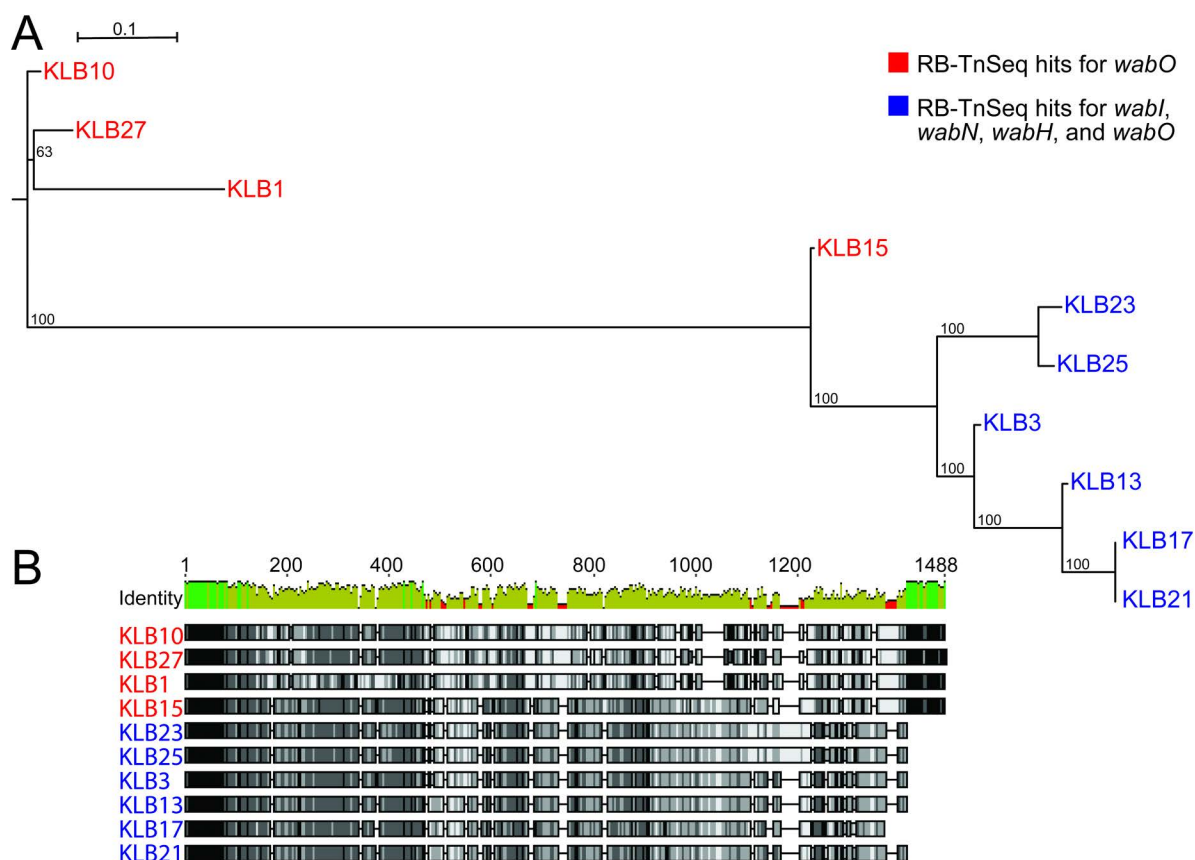


Fig 5. Genomic and structural differences in the long tail fiber of *Slopekvirus* phages correlate with LPS binding profiles. (A) Phylogenetic tree of the distal tail fiber subunit from *Slopekvirus* phages. Phages requiring only *wabO* for infection are labeled in red; phages requiring *wabl*, *wabN*, *wabH*, and *wabO* are labeled in blue. (B) Multiple sequence alignment of the distal long tail fiber subunit across all *Slopekvirus* phages, highlighting conserved and variable regions associated with receptor binding specificity.

<https://doi.org/10.1371/journal.pgen.1012233.g005>

Non-receptor host factors uncovered in RB-TnSeq screens

Beyond surface receptor genes, we identified 17 bacterial genes involved in intracellular functions, the majority of which were phage-specific, associated with only one to three phages (Fig 2). These genes generally produced lower fitness scores than the surface receptor hits, suggesting a more modest or indirect contribution to phage infection rather than an essential receptor role. These non-receptor dependencies span transcriptional regulation, inner membrane integrity, energy metabolism, protein folding, and amino acid transport. While none of these genes have been directly associated in the context of phage infection in *Klebsiella*, their known roles in cell envelope physiology and bacterial physiology provide a basis for hypothesizing how their disruption might influence phage infection.

The cytochrome bd oxidase subunits *cydA* and *cydB* were associated with phages KLB19, KLB26, and KLB31, all of which also produced fitness hits for *dsbA*. Given that cytochrome bd oxidase contributes to the reducing power required by DsbA [75, 76], this co-occurrence suggests a potential functional link between these genes during infection. The transcriptional regulator SdiA was associated with phage KLB17; disruption of *sdiA* in *Klebsiella* leads to increased bio-film formation, impaired cell division, and filamentous morphology [77], any of which could alter phage adsorption. The inner membrane proteins NlpD and TolR, associated with phages KLB19, KLB21, and KLB31, are both implicated in cell envelope architecture. NlpD activates peptidoglycan amidases during cell division [77], and its disruption produces outer membrane defects that could affect phage binding or lysis. TolR is a component of the Tol-Pal system, which spans the inner membrane and periplasm and maintains outer membrane integrity; mutations in Tol-Pal genes in *E. coli* disrupt LPS structure [78], suggesting an indirect effect on phage receptor accessibility rather than a direct role in DNA translocation for the *Straboviridae* phages in this study. The core component of the twin-arginine translocation system *tatC* mutants showed fitness in presence of KLB19, KLB21, and KLB4. To assess whether the *tatC* association could reflect direct Tat-dependent export of a phage protein or screen hit, all phage-encoded proteins and high-fitness bacterial genes were systematically examined for canonical Tat signal motifs using SignalP-6.0. None were identified (S6 Table), ruling out the simplest model and implicating a more indirect role for the Tat pathway, potentially through export of host proteins that maintain surface accessibility for the relevant phages.

Finally, we note that because the phages used in this screen were isolated on *Klebsiella* sp. M5a1, they have by definition overcome or evaded the defense systems present in this strain. Disruption of those defense genes would confer no fitness advantage in the RB-TnSeq library, and this screen design is therefore not suited to evaluating the contribution of defense systems to host range in this strain.

Experimental validation of RB-TnSeq hits

To confirm that RB-TnSeq fitness scores reflect genuine phage resistance phenotypes, we generated individual, sequence-confirmed knockout strains for eight genes spanning receptor biosynthesis, outer membrane transport, and regulatory functions. These include the LPS outer core biosynthesis genes *wabI* (RS27275), *wabN* (RS27290), and *wabH* (RS27305); the outer membrane porin *ompK36*; the outer membrane long-chain fatty acid transporter *fadL*; the pyrrolo-quinoline quinone biosynthesis gene *pqqU*; the predicted inner membrane protein RS24775; and the periplasmic disulfide bond catalyst *dsbA*. Complementation was performed for knockouts to exclude polar effects on neighboring genes.

Validated surface receptor determinants

LPS outer core biosynthesis genes

The RB-TnSeq screen identified *wabI*, *wabN*, and *wabH* as conferring resistance to a subset of *Slopekvirus* phages (Figs 4A–B and 6A). Disruption of each gene conferred complete resistance to these phages, with the exception of partial resistance observed in *Autographiviridae* phages, consistent with potential alternative receptor usage (Fig 6A). Adsorption

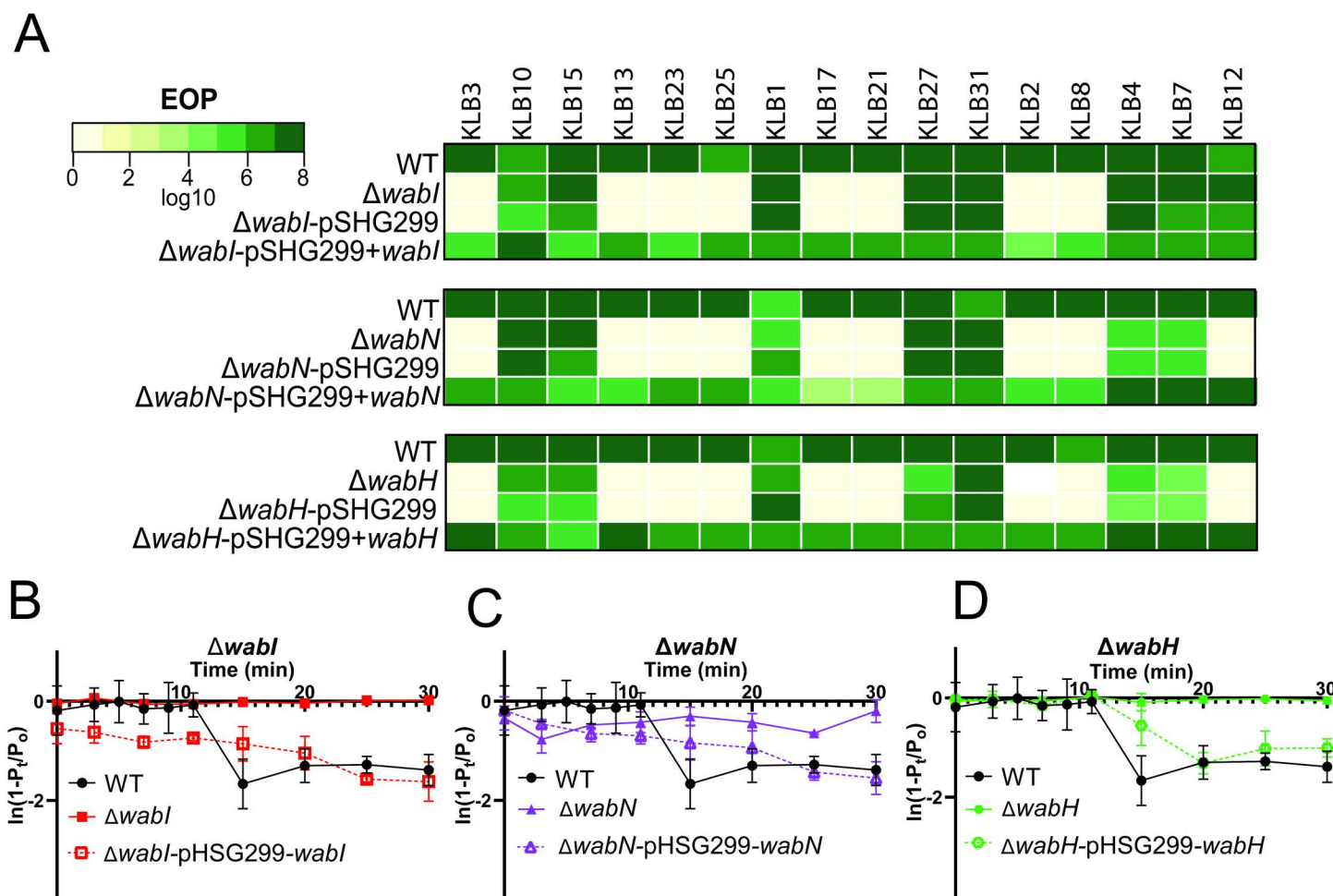


Fig 6. Impact of lipopolysaccharide (LPS) biosynthesis gene knockouts on phage infection. (A) Titers of the phages that required LPS to infect WT, Δ RS27275, Δ RS27290, and Δ RS27305, each of these mutants with the empty vector, and each mutant with a complemented gene. Black denotes phages not assayed as they did not require LPS based on the RB-TnSeq assay. Rate of adsorption of phage KLB2 for each of the three LPS mutants: (B) RS27275, (C) RS27290, and (D) RS27305. In adsorption assay panels (B to D), the solid black line represents the WT control, showing baseline phage adsorption.

<https://doi.org/10.1371/journal.pgen.1012233.g006>

assays confirmed that phage binding was abolished in all three knockout backgrounds, directly supporting the role of these LPS outer core biosynthesis genes as functional phage receptor determinants (Fig 6B, 6C and 6D). Complementation of each knockout restored wild-type infectivity, confirming that resistance was attributable to loss of the specific gene rather than polar effects on neighboring loci. These genes had not been characterized as phage receptors in any *Klebsiella* strain prior to this study.

Mutations in *wabH*, *wabN*, or *wabI* alter outer core structure in a manner that reduces phage adsorption or infection efficiency, while complete removal of O-antigen via *waaL* or *wabJ* disruption has little to no impact on phage binding. Thus, although upstream disruptions may indirectly affect O-antigen composition, their fitness effects are more parsimoniously explained by structural changes to the outer core rather than by loss of O-antigen per se. Repeated attempts to generate a clean deletion of *wabO* were unsuccessful; this hit is therefore retained as a high-confidence candidate for future validation.

OmpK36

The screen identified *ompK36* as conferring cross-resistance across *Ca. Rivulusviridae* family. A $\Delta ompK36$ knockout was generated and tested by efficiency of plating against KLB24, KLB28, and KLB29 phages with a corresponding high fitness score. Disruption of *ompK36* conferred strong resistance to all the three phages and complementation restored phage infectivity, confirming OmpK36 as a functional *Ca. Rivulusviridae* family receptor in *Klebsiella* sp. M5a1 (Fig 7A). OmpK36 was previously identified as a receptor in a capsulated *K. pneumoniae* strain, where it functioned as a secondary receptor accessed after capsule depolymerization [33]. Its validation here as a primary receptor in an acapsular host directly demonstrates that the receptor landscape exposed in the absence of a capsule is both distinct from and complementary to that characterized in capsulated strains.

FadL

The screen identified *fadL*, encoding an outer membrane long-chain fatty acid transporter, as a hit for phage KLB31, a member of the *Straboviridae*. A $\Delta fadL$ knockout was generated and infectivity assessed by efficiency of plating. Disruption of *fadL* conferred resistance to KLB31, and complementation restored wild-type infectivity, identifying FadL as a functional

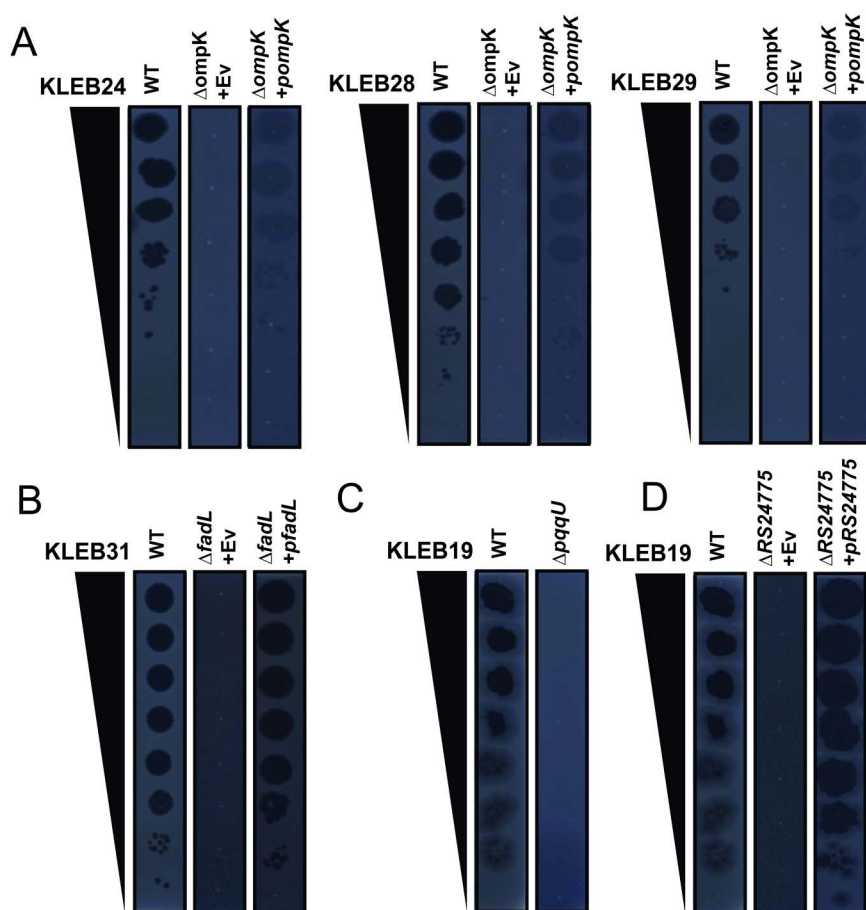


Fig 7. Experimental validations of top-scoring RB-TnSeq hits. Efficiency of plating for phages with RB-TnSeq hits in (A) $\Delta ompK36$, (B) $\Delta fadL$, (C) $\Delta RS24775$, (D) $\Delta pqqU$. WT is wild type host strain, Ev is an empty vector, p is for complementation plasmid.

<https://doi.org/10.1371/journal.pgen.1012233.g007>

receptor determinant for this phage (Fig 7B). FadL has been previously identified as a phage receptor in *E. coli* [39, 43], and this work represents the first characterization of *fadL* as a receptor in *Klebsiella*.

PqqU

The screen identified *pqqU*, involved in pyrroloquinoline quinone biosynthesis, as a hit for phages KLB19 and KLB26, both members of the *Drexleviridae*. A $\Delta pqqU$ knockout was generated and tested by the efficiency of plating against KLB19. Disruption of *pqqU* conferred resistance to KLB19 (Fig 7C). The role of PqqU in phage infection had not been previously described in *Klebsiella*, but has been found in other phage-host systems [43, 79]. Its identification here raises the possibility that PQQ-dependent surface modifications influence receptor accessibility for a subset of *Drexleviridae* phages. Notably, while *pqqU* was required for both KLB19 and KLB26, RS24775 was identified as an additional hit specific to KLB19 (see below) but not KLB26, suggesting that KLB19 may depend on a broader set of host factors for productive infection than KLB26, despite both phages sharing the *pqqU* requirement. Whether RS24775 and PqqU act in the same or independent pathways during KLB19 infection represents a question for future investigation.

Validated non-receptor host factors

RS24775

The screen identified RS24775, encoding a predicted inner membrane protein with homology to YqjF, as a hit for phage KLB19, a *Drexleviridae* phage. A $\Delta RS24775$ knockout was generated and infectivity assessed by efficiency of plating. Disruption conferred resistance to KLB19, and complementation restored wild-type infectivity (Fig 7D), indicating RS24775 role in KLB19 phage infection. The function of YqjF-family proteins is not well characterized in any organism, and this represents the first reported association of this protein family with phage susceptibility. Its inner membrane localization suggests a possible role in post-adsorption infection steps rather than direct receptor function, though this awaits experimental investigation.

DsbA

The screen identified *dsbA* as a hit for seven phages across the *Straboviridae* and *Drexleviridae* families. The fitness signal for *dsbA*, while reproducible across multiple independent transposon insertions, was lower in magnitude than that observed for the receptor genes described above, predicting partial rather than complete resistance. To test this prediction, plating efficiency and liquid growth inhibition were compared in wild-type versus $\Delta dsbA$ cells. The $\Delta dsbA$ mutant did not confer complete resistance, with efficiency of plating measurements revealing modest reductions for a subset of phages (S1 Fig), and liquid growth assays showing a significant decrease in infectivity for phage KLB1 in the $\Delta dsbA$ background while other phages showed no significant difference (S1 Fig). DsbA catalyses disulfide bond formation in periplasmic and outer membrane proteins, and its disruption is known to impair the folding and surface presentation of multiple outer membrane proteins in gram-negative bacteria. In *Klebsiella* sp. M5a1, where LPS outer core and outer membrane porins serve as primary receptor structures, DsbA-dependent folding of one or more surface components could plausibly influence their accessibility or structural integrity as phage receptor targets, accounting for the partial resistance phenotype observed across multiple phage families.

The co-occurrence of *cydA* and *cydB* fitness signals across the same phages is consistent with this interpretation, as cytochrome bd oxidase contributes to the reducing power that supports DsbA activity [75]. The attenuated $\Delta dsbA$ phenotype may additionally reflect functional redundancy with the disulfide bond isomerase DsbC, which cooperates with DsbA in maintaining periplasmic protein folding and could partially compensate for DsbA loss in certain contexts [75]. A prior transposon screen in *Acinetobacter* also identified *dsbA* as a phage infection determinant, where its disruption truncated

the capsular polysaccharide and paradoxically increased phage adsorption [80]. Additionally, in phage lambda assays, *dsbA* disruption was found to have a moderate delay in lysis timing [81]. In *Klebsiella* sp. M5a1, which is naturally acapsular, this capsule-dependent mechanism cannot apply, and the opposing direction of the phenotype, namely partial resistance rather than increased susceptibility, suggests that DsbA influences phage infection through a distinct, capsule-independent pathway in this host. Whether a common surface-remodelling regulatory mechanism underlies DsbA effects across both systems represents an open question for future investigation.

Collectively, the 42 genes identified in this work reveal the breadth of host functions co-opted by diverse *Klebsiella* phages, from surface receptor biosynthesis to intracellular metabolism, and provide a resource for directing studies of phage-host interaction mechanisms (Figs 2 and 8, S7 Table).

Conclusion

Our study delivers the first genome-wide, cross-family map of phage infection dependencies in *Klebsiella*, built deliberately on an acapsular host that reveals the receptor landscape invisible to prior capsule-centric studies. Systematic

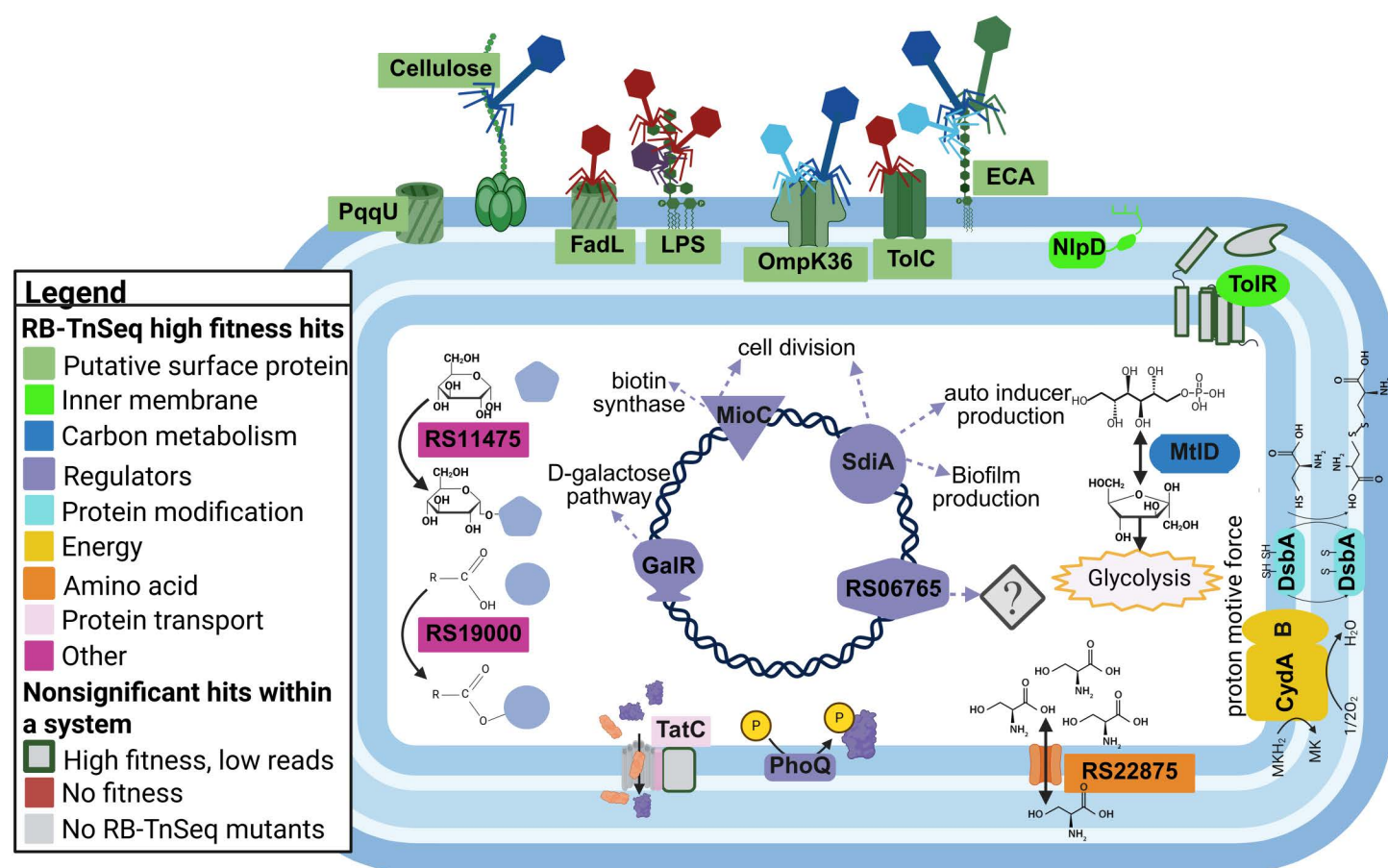


Fig 8. Overview of phage-bacteria interactions during phage infection. A total of 42 bacterial genes were identified in the RB-TnSeq screen. Bacterial genes are color-coded by their predicted function: surface receptor (green), inner membrane protein (light green), carbon metabolism (blue), regulators (purple), protein folding (light blue), energy production (yellow), protein transport (light pink), and other proteins (dark pink). For bacterial genes that are part of a complex, the complex is color-coded based on RB-TnSeq results. Green border: high fitness but read count less than 100; gray: no RB-TnSeq mutants available. Created in BioRender. Gittrich, M. (2026) <https://BioRender.com/rhi60kf>.

<https://doi.org/10.1371/journal.pgen.1012233.g008>

characterization of phage receptors in *Klebsiella* has been largely unexplored with decades of individual phage studies in capsulated strains that collectively produced knowledge of a handful of receptors for a small number of phages [28, 65–67, 82]. By examining phage infection in a non-capsulated host, we uncovered 42 high scoring host interaction determinants that influence phage infection and are likely obscured in capsule-centric studies. These results expand our understanding of phage-host interactions and highlight overlooked genetic factors that may govern resistance and host range.

To move from descriptive studies to predictive frameworks, we must identify conserved infection strategies across taxonomic boundaries. Our findings show that bacterial gene requirements for infection significantly cluster by phage genus and family, suggesting that infection strategies are often taxonomically structured. However, even within these groups, we observed divergences in host gene dependencies, particularly in receptor usage and intracellular processes. This variation underscores the value of pairing phage genomes with systematic host genetic screens like RB-TnSeq. This type of classification provides a functional layer of prediction that goes beyond sequence similarity and can help us understand the mechanisms of host-range.

In summary, this work advances the understanding of *Klebsiella* phage-host interactions by revealing the receptor architecture beneath the capsule, and provides a community resource for linking phage genomics to functional host interactions. These findings support the rational design of phage cocktails that account for capsule-loss escape variants, inform host-range prediction, and contribute to the broader goal of moving from descriptive phage genomics to mechanistically grounded, predictive models of phage infectivity.

Methods

Bacterial strains and growth conditions

Bacterial strains, phages, and their sources are listed in (S8 Table). *Klebsiella* sp. M5a1 was grown in LB broth at 30 °C, with shaking at 150 rpm. All *E. coli* strains used for plasmid maintenance and conjugation were grown in LB at 37°C and shaking at 150 rpm. All bacterial strains were stored long-term in -80 °C with 15% glycerol.

Bacteriophages and propagation

The *Klebsiella* sp. M5a1 phages were isolated, annotated, and 24 of the 25 phages were described by our group previously [50]. Phages were amplified using standard protocols. Briefly, *Klebsiella* sp. M5a1 was grown to mid-exponential (OD of 0.3) and five phage plaques were added to the culture. The infected culture was then incubated overnight, and the lysates were centrifuged and filtered using a 0.22 µm filter. Phage titer was determined by diluting lysates in SM buffer (100mM NaCl, 8mM MgSO₄, 50mM Tris, pH 7.5) and mixing 100 µL of diluted phages with 3 mL of 0.7% LB agar supplemented with 10 mM CaCl₂ and 10 mM MgSO₄, and 200 µL of overnight culture. Plates were incubated at 30 °C overnight.

Competitive growth experiments of the RB-TnSeq library

Klebsiella sp. M5a1 RB-TnSeq library was previously created by collaborators at the Lawrence Berkeley National Laboratory [33]. Briefly, the library was generated using the Tn5 transposase and consists of 134,723 unique barcodes. There were, on average, six knockouts per nonessential gene. A single aliquot (1 mL) of *Klebsiella* sp. M5a1 RB-TnSeq library stored at -80 °C was thawed and transferred to LB supplemented with 50 µg/mL of kanamycin. The culture was grown to mid-exponential at 30 °C with shaking at 150 rpm until the OD reached between 0.2–0.5, measured with an Epoch 2 plate reader. From this culture, three 1-mL aliquots were sampled and spun at 21,130 x g to pellet. These bacterial pellets were used as the reference for BarSeq (termed T0). The library was then diluted to an OD of 0.04 in LB medium prepared at 2 × concentration, containing 20 g tryptone, 10 g yeast extract, and 10 g NaCl per liter. In a 48-well plate, 350 µL of the diluted RB-TnSeq mixture was then mixed in a 1:1 ratio with 350 µL of phages diluted in SM buffer at three MOIs (0.1, 1, and 10). The plate was covered with a Breathe-Easy film, and the 48-well plate was then placed in an Epoch 2 plate

reader at 30 °C with double orbital shaking at 150 rpm and OD 600 nm readings every 15 minutes for 4.5 hours. At the end of the experiment, each well contents were transferred to a 1.5-mL microcentrifuge tube and spun at 21,130 x g to pellet the bacterial cells. The supernatant was removed, and the bacterial pellets were placed at -80 °C until DNA was isolated for BarSeq sequencing.

For each phage, there were at least two biological replicates for each time point, with additional replicates sometimes performed for clarity (S9 Table). In total, there were 213 phage-infected RB-TnSeq assays. Across these 213 assays, we identified 669 significant fitness hits (S10 Table) that dispersed across 63 bacterial genes. These hits were further curated based on read abundance (>100 reads for three strains) to reduce the chances of false positives (n = 16) and checked for potential polar effects (n = 5), resulting in 42 bacterial genes with significant fitness changes (Fig 2, S11 Table).

BarSeq of RB-TnSeq pooled fitness assay samples

We isolated genomic DNA from RB-TnSeq library samples using the DNeasy Blood and Tissue kit (Qiagen). We performed the 98°C BarSeq PCR protocol as described previously [38, 39] with the following modifications. Barcodes were PCR amplified using “BarSeq_V4” primers, as described previously [34]. These included a 10-bp index sequence for P7 primers and an internal 8-bp index (to identify Illumina index hopping), which allowed multiplexing up to 8 plates of 96-well BarSeq PCR samples (total 768 genome-wide assay samples). All experiments done on the same day and sequenced on the same lane were considered as a ‘set’. Equal volumes (5 µl) of the individual BarSeq PCRs were pooled, and 50 µl of the pooled PCR product were purified with the DNA Clean and Concentrator kit (Zymo Research). The final BarSeq library was eluted in 40 µl of sterile water.

Data processing and analysis of BarSeq reads

BarSeq reads were converted to counts per barcode using the MultiCodes.pl script in the feba code base (<https://bitbucket.org/berkeleylab/feba/>), with the -minQuality 0 option and -bs4 (BarSeq 4) for BarSeq primers. RB-TnSeq fitness data was analyzed as previously described [38, 39]. Briefly, the fitness value of each strain (an individual transposon mutant) was the normalized $\log_2(\text{strain barcode abundance at end of experiment} / \text{strain barcode abundance at the start of the experiment})$. The fitness value of each gene was the weighted average of the fitness of its strains. Because RB-TnSeq transposons were redundant, with on average six barcodes interrupting each nonessential gene and a minimum of three barcodes for each gene in our screen, the fitness score for each gene was determined by averaging the fitness levels of these strains. A positive fitness score indicated that the knocked-out host gene increased the relative fitness of the strain in the presence of phages, suggesting that the gene product was necessary for successful phage infection. Conversely, a negative fitness score suggested that the knocked-out strain was more susceptible to phage infection. We applied filters on experiments such that the mean reads per strain were ≥ 10 . In a typical experiment without stringent positive selection, a gene fitness score (fit) > 1 and an associated t-like statistic > 4 sufficed to give a low rate of false positives. Due to the stringent positive selection in phage assays, we required that $\text{fit} \geq 4$, $t \geq 5$, $\text{standard error} = \text{fit}/t \leq 2$, and $\text{fit} \geq \text{maxFit} - 8$, where maxFit is the highest fitness value of any gene in an experiment. The limit of maxFit - 8 was chosen based on our earlier work [39].

Annotation of bacterial genes identified in RB-TnSeq

To understand the functional significance of the 42 identified bacterial genes required for phage infection, we first classified them into nine functional categories: surface structures, inner membrane proteins, transcriptional regulators, carbon metabolism, energy cycling, amino acid biosynthesis and transport, protein transport, protein modification, other genes such as glycosyltransferase and acyltransferases that do not have specified pathways, as well as genes consisting of hypothetical proteins with no known function (Fig 2). The first eight functional categories were relatively straightforward to interpret. In contrast, the latter “other” category and hypothetical proteins were more difficult to analyze.

To refine our understanding of these latter two categories, the thirteen genes were more deeply assessed via homology-based searches, domain analyses, protein structure predictions, operon predictions, and constructed protein-protein interaction networks (S2 Fig; S12–S13 Tables) [83–87]. Together, these analyses provided complementary lines of evidence to generate putative functional annotations for otherwise uncharacterized proteins.

Five of these bacterial genes did not have enough evidence to associate with a specific gene or functional pathway. *RS27275*, *RS27305*, and *RS27310* were identified as family 1 glycosyltransferases, which mediate the transfer of sugar moieties to form glycosidic bonds. This glycosyltransferase family involves several pathways, including exopolysaccharide biosynthesis, colanic acid production, and lipopolysaccharide (LPS) biosynthesis. Notably, *RS27275*, *RS27290*, *RS27305*, and *RS27310* were clustered within eight genes of one another and were organized into three consecutive operons within the bacterial genome, all of which were associated with LPS biosynthesis. Multiple Sequence Alignment (MSA) confirmed that these genes were >80% similar to known homologs [88, 89], and they were renamed respectively *RS27275-wabI*, *RS27290-wabN*, *RS27305-wabH*, and *RS27310-wabO*. Additionally, *RS18965*, *RS18970*, *RS18985*, and *RS18990* were predicted to play a role in O-antigen and enterobacterial common antigen (ECA) biosynthesis and were placed in the surface structure category.

To see if some of the proteins had TAT domains, indicating these proteins are exported via the Tat system, we used SignalP-6.0 to identify proteins with the TAT domains.

Generation of individual *Klebsiella* sp. M5al knockouts

All individual *Klebsiella* sp. M5al single-gene deletions were created using one of two approaches depending on the gene target.

The majority of gene deletions were created using the VcDART, a CRISPR-associated transposase DART system, to make targeted knockouts [90]. The gRNAs are 32 nucleotides long and recognize targets with a 5'-CC-3' type I-F protospacer adjacent motif (PAM). For each deletion, gRNAs were designed targeting the first half of the gene of interest to ensure functional knockouts after transposon insertion (S14 Tables) with additional nucleotides added at the 5'- and 3'-ends to generate sticky ends for insertion into plasmid pBFC0619.

Oligos were inserted into plasmid pBFC0619 using a previously reported method involving the BsaI-HFv2 Golden Gate assembly kit (NEB). Briefly, 5 μ L of the Golden Gate assembly reaction was mixed with 50 μ L of competent *E. coli* and subjected to heat shock at 42 °C for 60 seconds, followed by immediate placement on ice for 5 minutes. Then, 1 mL of SOC broth (Super Optimal broth with Catabolite repression) was added, and cells were incubated at 37 °C with shaking at 150 rpm for 60 minutes before plating on LB agar supplemented with 10 μ g/mL gentamicin. SOC broth consisted of 2% (w/v) tryptone, 0.5% (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄, and 20 mM glucose. gRNA insertion was confirmed via PCR and Sanger sequencing.

For gRNA plasmids that only had high efficiency of transformation in *pir*⁺, the plasmid was isolated using Qiagen plasmid mini extraction kit. A 10 mL culture of *E. coli* *pir*⁺ transformant was grown in LB supplemented with 10 μ g/mL of gentamicin at 37 °C with shaking at 150 rpm. The 10 mL of culture was then pelleted, and the plasmid was isolated using the Qiagen plasmid mini extraction protocol. The plasmid was then electroporated into electrocompetent *E. coli* WM3064 (25 μ F; 200 Ω , 2.5 kV), incubated for 2 hours at 37 °C with shaking at 150 rpm, and plated on 10 μ g/mL gentamicin [91].

To generate the targeted knockout, the pBFC0619 plasmid with the inserted gRNA was transferred to *Klebsiella* via conjugation with *E. coli* WM3064 using previously established protocols [90]. The donor and acceptor strains were mixed in 1:1 ratios on top of a 0.22 μ m filter on an LB plate supplemented with DAP (0.06 mg/mL) and IPTG (0.1 M). IPTG induced the CRISPR-associated transposase to insert the transposon into the gene of interest. The plates were incubated at 30 °C for 6–8 hours. The filter was resuspended in 2 mL of LB and plated on LB supplemented with 25 μ g/mL of gentamicin. Transposon insertion into the gene of interest was confirmed with Sanger sequencing.

For a subset of deletions (*pqqU*, *fadL*, *RS24775*), targeted knockouts were generated using the Lambda Red recombination system carried on plasmid pSim5 [92]. Electrocompetent *Klebsiella* sp. M5al cells were prepared as follows [93]. Cells were grown overnight to saturation, diluted 1:100 into fresh LB, and grown to an OD600 of approximately 0.1. The Lambda Red system was induced by shifting cultures to 42°C and growth was continued until an OD600 of approximately 0.5 was reached. Cultures were immediately placed on ice for 15–30 minutes, washed three times with ice-cold 10% glycerol, and concentrated 100-fold in ice-cold 10% glycerol. A deletion cassette consisting of a kanamycin resistance marker flanked by sequences homologous to the regions immediately upstream and downstream of the target gene was amplified by PCR. Fifty microlitres of electrocompetent cells were electroporated with 50 ng of the PCR product using 0.1 mm cuvettes at 1.8 kV, recovered in 1 mL of LB for 2 hours at 30°C, and plated on LB agar with kanamycin selection at 30°C for 24 hours. Successful integration of the kanamycin cassette at the target locus was confirmed by whole-genome sequencing. All *Klebsiella* sp. M5al mutants were stored at -80°C in LB supplemented with 15% glycerol.

Complementation of *Klebsiella* sp. M5al knockouts

Complementation experiments were performed using two approaches depending on the gene target.

For the majority of complementation constructs, the target bacterial gene was PCR-amplified using primers incorporating SbfI and SphI restriction sites (S15 Table). The amplified DNA was gel-purified and digested alongside the plasmid vector pSHG299 using SbfI and SphI restriction enzymes (New England Biolabs). For ligation, 66 ng of the digested PCR product were combined with 33 ng of linearized pSHG299 (2:1 molar ratio), T4 DNA ligase, and the supplied buffer, and incubated under standard ligation conditions. The resulting ligation mixture was transformed into *E. coli* pir+ via heat shock. Transformants were screened by colony PCR to verify the presence of the insert. Plasmid DNA was isolated from positive clones using the QIAprep Spin Miniprep Kit (Qiagen) and electroporated into electrocompetent *Klebsiella* sp. M5al cells. Following electroporation, cells were recovered for 2 hours at 37 °C in 1 mL of SOC broth, then plated on LB agar containing 50 µg/mL of kanamycin. Successful plasmid integration into the genome was confirmed by colony PCR and Sanger sequencing. For complementation of *RS24775* and *fadL*, a modified vector was constructed by replacing the kanamycin resistance marker in pSHG299 with a chloramphenicol resistance marker to enable selection in the (pSim5 enabled) kanamycin-marked deletion backgrounds. The target genes were PCR-amplified with primers incorporating overlapping sequences compatible with the modified vector. Amplified inserts were cloned into the linearized chloramphenicol-marked vector using the NEBioLabs Gibson Assembly mix according to the manufacturer's instructions. The assembled constructs were electroporated into the corresponding deletion strains using the electrocompetent cell preparation protocol described above. Following electroporation, cells were recovered for 2 hours at 30°C in 1 mL of LB broth and plated on LB agar containing 300 µg/mL chloramphenicol. Successful construct integration was confirmed by colony PCR and Sanger sequencing. All complementation strains were stored at -80°C in 15% glycerol until use.

Phage infection efficiency of *Klebsiella* sp. M5al knockouts

Phage resistance was assayed through efficiency of plating (EOP) experiments. Bacteria were grown overnight at 30 °C with shaking at 150 rpm and with 25 µg/mL of gentamicin. For *RS24775* and *fadL* deletion strains carrying empty or complementation constructs, overnight cultures were supplemented with both 50 µg/mL kanamycin and 300 µg/mL chloramphenicol to maintain both the deletion and complementation constructs simultaneously. An aliquot of 300 µL of the overnight culture was mixed with 5 mL of top agar (0.7% agar) and solidified at room temperature. Phages were diluted to a titer of 10⁸ PFU/mL and then ten-fold serially diluted in SM buffer. An aliquot of 5 µL of each phage dilution was then spotted onto the solidified bacterial lawn. The plates were incubated overnight at 30 °C. EOP was calculated as the ratio of the average effective phage titer on the *Klebsiella* sp. M5al knockout compared to the wild-type (WT) *Klebsiella* sp. M5al host. All plaquing experiments were performed with at least three biological replicates and with at least three overnight cultures for each bacterial strain.

Phage mediated growth inhibition assays

Phage-mediated killing of *Klebsiella* sp. M5al was quantified using a plate reader–based growth inhibition assay. Bacterial cultures were grown to mid-exponential phase ($OD_{600nm} = 0.3–0.6$) and diluted in $2 \times LB$ to an OD of 0.1. $MgSO_4$ and $CaCl_2$ were supplemented to a final concentration of 20 mM each. Phages were diluted in SM buffer to a final titer of 4.1×10^8 PFU/mL, corresponding to a multiplicity of infection (MOI) of 6 prior to mixing with bacteria (resulting in a final MOI of 3 after combining equal volumes of phage and host).

Phage suspensions (350 μL) were aliquoted into a 48-well flat-bottom microtiter plate according to a predefined layout, followed by the addition of 350 μL of the diluted *Klebsiella* suspension, for a final volume of 700 μL per well. Plates were sealed with Breathe-Easy film and incubated in an Epoch 2 microplate reader. OD was measured kinetically every 5 minutes over 12 hours, with shaking at 205 cpm using a double-orbital setting between reads.

Growth inhibition was quantified by calculating the area under the curve (AUC) using GraphPad Prism. To account for baseline growth differences, AUC values from noninfected controls were subtracted from those of the corresponding phage-infected cultures. Statistical significance comparing phage infection of WT bacteria to $\Delta dsbA$ was done using a two-tailed *t*-test in GraphPad Prism. *P* values are denoted as follows: $P < 0.05$ (*), $P < 0.01$ (**), and $P < 0.001$ (***).

Examining the patterns of bacterial resistance

RB-TnSeq network graphs were constructed using Gephi. The python scripts to generate the Gephi graphs were previously described [36]. In all graphs, edges were calculated based on datasets found in [S16 Table](#). Each mixed node graph in [Fig 3](#) was drawn with weight one between a phage node (fixed) and a gene node if that gene conferred resistance.

To assess patterns in bacterial gene dependencies across phages, we performed Principal Coordinates Analysis (PCoA) and Non-metric Multidimensional Scaling (NMDS) using Bray-Curtis dissimilarity matrices. All analyses were conducted in R (version 4.3.0). RB-TnSeq gene fitness data were imported and transformed into matrix format. Since for viral RB-TnSeq we only look at the positive RB-TnSeq values as the negative fitness values were indicative of phage infection, we set negative values to zero to avoid distortion in similarity calculations.

For both the PCoA and NMDS, we calculated Bray-Curtis dissimilarity matrices using the *vegdist* function (117). PCoA was performed using *cmdscale*, and the percentage of variation explained by each axis was calculated from the eigenvalues. For NMDS, we used the *metaMDS* function [94] with three dimensions ($k = 3$) and a maximum of 100 iterations (*trymax* = 100). The NMDS stress value was used to assess model fit, with values < 0.2 considered acceptable [95]. The third NMDS dimension was scaled by a factor of 100 to be visualized as point size in *ggplot2*.

PCoA and NMDS coordinates were merged with sample metadata and visualized using *ggplot2* [96], with colors corresponding to phage family, genus, or species, and ellipses representing 95% confidence intervals. To test for significant differences along PCoA axes, we performed Kruskal-Wallis tests followed by Dunn's multiple comparisons with Bonferroni correction using the *dunn.test* package [97, 98].

To statistically evaluate the clustering of samples by phage taxonomy, we used permutational multivariate analysis of variance (PERMANOVA) [99], implemented via *adonis2* in the *vegan* package, at the phage family, genus, and species levels. Post-hoc pairwise PERMANOVA tests were performed using the *pairwise.adonis* function from the *pairwiseAdonis* package [100]. Results of PERMANOVA are shown in [S3–S5 Tables](#).

Examination of differences in *Slopekvirus* long tail fibers

To understand the differences in the bacterial surface receptor genes for the *Slopekviruses* we examined the long-tail fiber protein that is required for phage adsorption [60]. To build an MSA phylogenetic tree to look at where the phage tail fibers clustered based on the bacterial surface gene requirements, we input the AA sequences into NGPhylogeny.fr PhyML/OneClick advanced workflow using the standard settings and adding bootstrap calculations. Tree used in [Fig 5A](#) was

visualized with iTOL [101]. The MSA alignment used in Fig 5B was also exported from the workflow and visualized using Geneious Prime 2020.05.

Supporting information

S1 Fig. STRING protein-protein interaction network of *Klebsiella* sp. M5al genes with high fitness scores from RB-TnSeq scores. Nodes are colored according to STRING local network clusters: LPS biosynthesis (red), enterobacterial common antigen biosynthesis (green), O-antigen biosynthesis (yellow), and cellulose biosynthesis (blue). White nodes represent genes that do not belong to a local network cluster or KEGG pathway. Edge thickness indicates the confidence score of the predicted protein-protein interactions based on STRING's scoring system with a minimum confidence score of 0.7.
(TIF)

S2 Fig. Minimal impact of the *Klebsiella* $\Delta dsbA$ mutation on phage infectivity. (A) Efficiency of plating (EOP) for six phages with RB-TnSeq hits in $\Delta dsbA$, measured on wild-type (WT, blue) and $\Delta dsbA$ (red) strains. (B) Difference in area under the curve (AUC) of bacterial growth (OD) between uninfected and phage-infected cultures in WT and $\Delta dsbA$ backgrounds. AUC was calculated in GraphPad, and to normalize for inherent growth differences between WT and $\Delta dsbA$, the AUC of non-infected controls was subtracted from that of each phage-infected culture. Statistical significance comparing phage infection of WT bacteria to $\Delta dsbA$ was done using a two-tailed t-test in GraphPad Prism. P values are denoted as follows: $P < 0.05$ (*), $P < 0.01$ (**), and $P < 0.001$ (***).
(TIF)

S1 Table. RB-TnSeq data used to generate Fig 2.
(XLSX)

S2 Table. Raw RB-TnSeq data.
(XLSX)

S3 Table. Permanova of RB-TnSeq data at family level.
(XLSX)

S4 Table. Permanova of RB-TnSeq data at genus level.
(XLSX)

S5 Table. Permanova of RB-TnSeq data at species level.
(XLSX)

S6 Table. Proteins predicted to have TAT signaling domains in *Klebsiella* sp. M5al and phages.
(XLSX)

S7 Table. Literature on bacterial genes previously found in phage infection. Functions and references to bacterial genes that have been found in previous phage infection across different hosts and phages.
(XLSX)

S8 Table. Phages and bacteria used in the study.
(XLSX)

S9 Table. Reasonings behind additional replicates for RB-TnSeq assays.
(XLSX)

S10 Table. List of RB-TnSeq hits with significant, high fitness scores.

(XLSX)

S11 Table. List of high fitness genes removed from RB-TnSeq assay and reasonings for removal.

(XLSX)

S12 Table. List of genes identified in RB-TnSeq with hypothetical or unknown pathway functions. We ran each of these genes through sequence similarity, hidden markov model profiles, domain searches, structural predictions, and literature search.

(XLSX)

S13 Table. Operon prediction for the *Klebsiella* sp. M5al genome.

(XLSX)

S14 Table. gRNAs used for generating single gene Kos.

(XLSX)

S15 Table. primers used to insert the KO gene back into *Klebsiella* sp. M5al.

(XLSX)

S16 Table. data used to generate the gephi graph examining the phage-bacterial gene interactions.

(XLSX)

Author contributions

Conceptualization: Marissa R. Gittrich, Courtney M. Sanderson, Vivek K. Mutalik, Matthew B. Sullivan.

Data curation: Marissa R. Gittrich, Cara M. Noel.

Formal analysis: Marissa R. Gittrich, Vivek K. Mutalik.

Funding acquisition: Sylvain Moineau, Vivek K. Mutalik, Matthew B. Sullivan.

Investigation: Marissa R. Gittrich, Courtney M. Sanderson, Cara M. Noel, Erica Babusci, Sumeyra C. Selbes, Madeline Svab, Isabella Murray, Collis Bousliman, Ami Fofana, Aghiad Daboul, Jonathan Leopold, Alessandra Gonçalves de Melo, Vivek K. Mutalik, Matthew B. Sullivan.

Methodology: Marissa R. Gittrich, Courtney M. Sanderson, Erica Babusci, Sumeyra C. Selbes, Ami Fofana, Aghiad Daboul, Jonathan Leopold, Alessandra Gonçalves de Melo, Vivek K. Mutalik.

Project administration: Sylvain Moineau, Vivek K. Mutalik, Matthew B. Sullivan.

Resources: Sylvain Moineau, Vivek K. Mutalik.

Supervision: Sylvain Moineau, Vivek K. Mutalik, Matthew B. Sullivan.

Validation: Madeline Svab, Isabella Murray, Collis Bousliman, Vivek K. Mutalik.

Visualization: Marissa R. Gittrich, Courtney M. Sanderson, Cara M. Noel, Madeline Svab, Isabella Murray, Collis Bousliman, Marion Urvoy, Vivek K. Mutalik, Matthew B. Sullivan.

Writing – original draft: Marissa R. Gittrich, Courtney M. Sanderson, Vivek K. Mutalik, Matthew B. Sullivan.

Writing – review & editing: Marissa R. Gittrich, Courtney M. Sanderson, Cara M. Noel, Erica Babusci, Sumeyra C. Selbes, Madeline Svab, Isabella Murray, Collis Bousliman, Ami Fofana, Aghiad Daboul, Jonathan Leopold, Alessandra Gonçalves de Melo, Marion Urvoy, Sylvain Moineau, Vivek K. Mutalik, Matthew B. Sullivan.

References

- Batinovic S, Wassef F, Knowler SA, Rice DTF, Stanton CR, Rose J. Bacteriophages in Natural and Artificial Environments. *Pathogens*. 2019;8(3):3. <https://doi.org/10.3390/pathogens8030100>
- Łusiak-Szelachowska M, Weber-Dąbrowska B, Żaczek M, Borysowski J, Górski A. The Presence of Bacteriophages in the Human Body: Good, Bad or Neutral? *Microorganisms*. 2020;8(12):2012. <https://doi.org/10.3390/microorganisms8122012>
- Townsend EM, Kelly L, Muscatt G, Box JD, Hargraves N, Lilley D, et al. The Human Gut Phageome: Origins and Roles in the Human Gut Microbiome. *Front Cell Infect Microbiol*. 2021;11:643214. <https://doi.org/10.3389/fcimb.2021.643214> PMID: 34150671
- Emerson JB, Roux S, Brum JR, Bolduc B, Woodcroft BJ, Jang HB, et al. Host-linked soil viral ecology along a permafrost thaw gradient. *Nat Microbiol*. 2018;3(8):870–80. <https://doi.org/10.1038/s41564-018-0190-y> PMID: 30013236
- Trubl G, Jang HB, Roux S, Emerson JB, Solonenko N, Vik DR, et al. Soil Viruses Are Underexplored Players in Ecosystem Carbon Processing. *mSystems*. 2018;3(5):e00076-18. <https://doi.org/10.1128/mSystems.00076-18> PMID: 30320215
- Daly RA, Roux S, Borton MA, Morgan DM, Johnston MD, Booker AE, et al. Viruses control dominant bacteria colonizing the terrestrial deep biosphere after hydraulic fracturing. *Nat Microbiol*. 2019;4(2):352–61. <https://doi.org/10.1038/s41564-018-0312-6> PMID: 30510171
- Amundson KK, Borton MA, Daly RA, Hoyt DW, Wong A, Eder E, et al. Microbial colonization and persistence in deep fractured shales is guided by metabolic exchanges and viral predation. *Microbiome*. 2022;10(1):5. <https://doi.org/10.1186/s40168-021-01194-8> PMID: 35034639
- Zhong ZP, Tian F, Roux S, Gazitúa MC, Solonenko NE, Li YF. Glacier ice archives nearly 15,000-year-old microbes and phages. *Microbiome*. 2021;9(1):160. <https://doi.org/10.1186/s40168-021-01106-w>
- Zhong ZP, Rapp JZ, Wainaina JM, Solonenko NE, Maughan H, Carpenter SD. Viral ecogenomics of Arctic cryopeg brine and sea ice. *mSystems*. 2020;5(3). <https://doi.org/10.1128/msystems.00246-20>
- Kuzyakov Y, Mason-Jones K. Viruses in soil: Nano-scale undead drivers of microbial life, biogeochemical turnover and ecosystem functions. *Soil Biology and Biochemistry*. 2018;127:305–17. <https://doi.org/10.1016/j.soilbio.2018.09.032>
- Braga LPP, Spor A, Kot W, Breuil MC, Hansen LH, Setubal JC, et al. Impact of phages on soil bacterial communities and nitrogen availability under different assembly scenarios. *Microbiome*. 2020;8(1):1–14. <https://doi.org/10.1186/S40168-020-00822-Z> PMID: 32252805
- Wang Y, Liu Y, Wu Y, Wu N, Liu W, Wang X. Heterogeneity of soil bacterial and bacteriophage communities in three rice agroecosystems and potential impacts of bacteriophage on nutrient cycling. *Environ Microbiome*. 2022;17(1):17. <https://doi.org/10.1186/s40793-022-00410-8> PMID: 35387674
- Morella NM, Gomez AL, Wang G, Leung MS, Koskella B. The impact of bacteriophages on phyllosphere bacterial abundance and composition. *Mol Ecol*. 2018;27(8):2025–38. <https://doi.org/10.1111/mec.14542> PMID: 29457297
- Santos-Medellin C, Zinke LA, Ter Horst AM, Gelardi DL, Parikh SJ, Emerson JB. Viromes outperform total metagenomes in revealing the spatio-temporal patterns of agricultural soil viral communities. *ISME J*. 2021;15(7):1956–70. <https://doi.org/10.1038/s41396-021-00897-y> PMID: 33612831
- Yang K, Wang X, Hou R, Lu C, Fan Z, Li J, et al. Rhizosphere phage communities drive soil suppressiveness to bacterial wilt disease. *Microbiome*. 2023;11(1):16. <https://doi.org/10.1186/s40168-023-01463-8> PMID: 36721270
- Wang X, Tang Y, Yue X, Wang S, Yang K, Xu Y, et al. The role of rhizosphere phages in soil health. *FEMS Microbiol Ecol*. 2024;100(5):fae052. <https://doi.org/10.1093/femsec/fae052> PMID: 38678007
- Graham EB, Camargo AP, Wu R, Neches RY, Nolan M, Paez-Espino D, et al. A global atlas of soil viruses reveals unexplored biodiversity and potential biogeochemical impacts. *Nat Microbiol*. 2024;9(7):1873–83. <https://doi.org/10.1038/s41564-024-01686-x> PMID: 38902374
- Gregory AC, Zayed AA, Conceição-Neto N, Temperton B, Bolduc B, Alberti A. Marine DNA viral macro- and microdiversity from pole to pole. *Cell*. 2019;177(5):1109–1123.e14. <https://doi.org/10.1016/j.cell.2019.03.040> PMID: 31031001
- Ma B, Wang Y, Zhao K, Stirling E, Lv X, Yu Y, et al. Biogeographic patterns and drivers of soil viromes. *Nat Ecol Evol*. 2024;8(4):717–28. <https://doi.org/10.1038/s41559-024-02347-2> PMID: 38383853
- Boon M, Holtappels D, Lood C, van Noort V, Lavigne R. Host Range Expansion of Pseudomonas Virus LUZ7 Is Driven by a Conserved Tail Fiber Mutation. *Phage (New Rochelle)*. 2020;1(2):87–90. <https://doi.org/10.1089/phage.2020.0006> PMID: 36147895
- Le S, He X, Tan Y, Huang G, Zhang L, Lux R, et al. Mapping the tail fiber as the receptor binding protein responsible for differential host specificity of Pseudomonas aeruginosa bacteriophages PaP1 and JG004. *PLoS One*. 2013;8(7):e68562. <https://doi.org/10.1371/journal.pone.0068562> PMID: 23874674
- Habusha M, Tzipilevich E, Fiyaksel O, Ben-Yehuda S. A mutant bacteriophage evolved to infect resistant bacteria gained a broader host range. *Mol Microbiol*. 2019;111(6):1463–75. <https://doi.org/10.1111/mmi.14231> PMID: 30811056
- Hussain FA, Dubert J, Elsherbini J, Murphy M, VanInsberghe D, Arevalo P. Rapid evolutionary turnover of mobile genetic elements drives bacterial resistance to phages. *Science*. 2021;374(6566):488–92. <https://doi.org/10.1126/science.abb1083>
- Ceyssens PJ, Glonti T, Kropinski AM, Lavigne R, Chanishvili N, Kulakov L. Phenotypic and genotypic variations within a single bacteriophage species. *Virology Journal*. 2011;8(1):134. <https://doi.org/10.1186/1743-422X-8-134>
- Howard-Varona C, Hargreaves KR, Solonenko NE, Markillie LM, White RA 3rd, Brewer HM, et al. Multiple mechanisms drive phage infection efficiency in nearly identical hosts. *ISME J*. 2018;12(6):1605–18. <https://doi.org/10.1038/s41396-018-0099-8> PMID: 29568113

26. Howard-Varona C, Lindback MM, Fudyma JD, Krongauz A, Solonenko NE, Zayed AA, et al. Environment-specific virocell metabolic reprogramming. *ISME J*. 2024;18(1):wrae055. <https://doi.org/10.1093/ismejo/wrae055> PMID: [38552150](#)
27. Howard-Varona C, Lindback MM, Bastien GE, Solonenko N, Zayed AA, Jang HB, et al. Phage-specific metabolic reprogramming of virocells. *ISME J*. 2020;14(4):881–95. <https://doi.org/10.1038/s41396-019-0580-z>
28. Boeckaerts D, Stock M, Criel B, Gerstmanns H, De Baets B, Briers Y. Predicting bacteriophage hosts based on sequences of annotated receptor-binding proteins. *Sci Rep*. 2021;11(1):1467. <https://doi.org/10.1038/s41598-021-81063-4> PMID: [33446856](#)
29. Camargo AP, Nayfach S, Chen IMA, Palaniappan K, Ratner A, Chu K. IMG/VR v4: an expanded database of uncultivated virus genomes within a framework of extensive functional, taxonomic, and ecological metadata. *Nucleic Acids Research*. 2023;51(D1):D733–43. <https://doi.org/10.1093/nar/gkac1037>
30. Comeau AM, Hatfull GF, Krisch HM, Lindell D, Mann NH, Prangishvili D. Exploring the prokaryotic virosphere. *Res Microbiol*. 2008;159(5):306–13. <https://doi.org/10.1016/j.resmic.2008.05.001> PMID: [18639443](#)
31. Roux S, Camargo AP, Coutinho FH, Dabdoub SM, Dutilh BE, Nayfach S, et al. iPhoP: An integrated machine learning framework to maximize host prediction for metagenome-derived viruses of archaea and bacteria. *PLoS Biol*. 2023;21(4):e3002083. <https://doi.org/10.1371/journal.pbio.3002083> PMID: [37083735](#)
32. Nie W, Qiu T, Wei Y, Ding H, Guo Z, Qiu J. Advances in phage–host interaction prediction: in silico method enhances the development of phage therapies. *Brief Bioinform*. 2024;25(3):bbae117. <https://doi.org/10.1093/bib/bbae117>
33. Dunstan RA, Bamert RS, Tan KS, Imbulgoda U, Barlow CK, Taiaroa G, et al. Epitopes in the capsular polysaccharide and the porin OmpK36 receptors are required for bacteriophage infection of *Klebsiella pneumoniae*. *Cell Rep*. 2023;42(6):112551. <https://doi.org/10.1016/j.celrep.2023.112551> PMID: [37224021](#)
34. Kortright KE, Chan BK, Turner PE. High-throughput discovery of phage receptors using transposon insertion sequencing of bacteria. *Proceedings of the National Academy of Sciences*. 2020. <https://doi.org/10.1073/pnas.2001888117>
35. Christen M, Beusch C, Bösch Y, Cerletti D, Flores-Tinoco CE, Del Medico L, et al. Quantitative Selection Analysis of Bacteriophage ϕ CbK Susceptibility in *Caulobacter crescentus*. *J Mol Biol*. 2016;428(2 Pt B):419–30. <https://doi.org/10.1016/j.jmb.2015.11.018> PMID: [26593064](#)
36. Rotman E, McClure S, Glazier J, Fuerte-Stone J, Foldi J, Erani A, et al. Rapid design of bacteriophage cocktails to suppress the burden and virulence of gut-resident carbapenem-resistant *Klebsiella pneumoniae*. *Cell Host Microbe*. 2024;32(11):1988–2003.e8. <https://doi.org/10.1016/j.chom.2024.09.004> PMID: [39368473](#)
37. Price MN, Wetmore KM, Waters RJ, Callaghan M, Ray J, Liu H, et al. Mutant phenotypes for thousands of bacterial genes of unknown function. *Nature*. 2018;557(7706):503–9. <https://doi.org/10.1038/s41586-018-0124-0> PMID: [29769716](#)
38. Wetmore KM, Price MN, Waters RJ, Lamson JS, He J, Hoover CA. Rapid quantification of mutant fitness in diverse bacteria by sequencing randomly bar-coded transposons. *mBio*. 2015;6(3). <https://doi.org/10.1128/mBio.00306-15> PMID: [25968644](#)
39. Mutalik VK, Adler BA, Rishi HS, Piya D, Zhong C, Koskella B, et al. High-throughput mapping of the phage resistance landscape in *E. coli*. *PLoS Biol*. 2020;18(10):e3000877. <https://doi.org/10.1371/journal.pbio.3000877> PMID: [33048924](#)
40. Adler BA, Kazakov AE, Zhong C, Liu H, Kutter E, Lui LM, et al. The genetic basis of phage susceptibility, cross-resistance and host-range in *Salmonella*. *Microbiology*. 2021;167(12). <https://doi.org/10.1099/mic.0.001126>
41. Sellner B, Prakapaitė R, van Berkum M, Heinemann M, Harms A, Jenal U. A new sugar for an old phage: a c-di-GMP-dependent polysaccharide pathway sensitizes *Escherichia coli* for bacteriophage infection. *mBio*. 2021;12(6). <https://doi.org/10.1128/MBIO.03246-21> PMID: [34903045](#)
42. Junkermeier EH, Hengge R. A novel locally c-di-GMP-controlled exopolysaccharide synthase required for bacteriophage N4 infection of *Escherichia coli*. *mBio*. 2021;12(6):e03249–21. <https://doi.org/10.1128/mbio.03249-21> PMID: [34903052](#)
43. Maffei E, Shaidullina A, Burkholder M, Heyer Y, Estermann F, Druelle V, et al. Systematic exploration of *Escherichia coli* phage-host interactions with the BASEL phage collection. *PLoS Biol*. 2021;19(11):e3001424. <https://doi.org/10.1371/journal.pbio.3001424> PMID: [34784345](#)
44. Moriniere L, Noonan AJC, Kazakov A, Pena M, Svab M, Rivera-Lopez EO. Enabling the prediction of phage receptor specificity from genome data. *bioRxiv*. 2026. <https://doi.org/10.64898/2026.04.02.716166>
45. Le Bris J, Chen N, Supandy A, Rendueles O, Van Tyne D. Phage therapy for *Klebsiella pneumoniae*: Understanding bacteria-phage interactions for therapeutic innovations. *PLoS Pathog*. 2025;21(4):e1012971. <https://doi.org/10.1371/journal.ppat.1012971> PMID: [40198880](#)
46. Latka A, Lemire S, Grimon D, Dams D, Maciejewska B, Lu T, et al. Engineering the modular receptor-binding proteins of *Klebsiella* phages switches their capsule serotype specificity. *mBio*. 2021;12(3):e00455–21. <https://doi.org/10.1128/mBio.00455-21> PMID: [33947754](#)
47. Haudiquet M, Le Bris J, Nucci A, Bonnin RA, Domingo-Calap P, Rocha EPC, et al. Capsules and their traits shape phage susceptibility and plasmid conjugation efficiency. *Nat Commun*. 2024;15(1):2032. <https://doi.org/10.1038/s41467-024-46147-5> PMID: [38448399](#)
48. Fu Y, Yin M, Cao L, Lu Y, Li Y, Zhang L. Capsule mutations serve as a key strategy of phage resistance evolution of K54 hypervirulent *Klebsiella pneumoniae*. *Commun Biol*. 2025;8(1):257. <https://doi.org/10.1038/s42003-025-07687-8> PMID: [39966630](#)
49. Rothschild-Rodriguez D, Lambon KS, Kushwaha SK, Garushyants SK, Ertelt M, Latka A, et al. KlebPhaCol: a community-driven resource for *Klebsiella* research identified a novel phage family. *Nucleic Acids Res*. 2025;53(21):gkaf1122. <https://doi.org/10.1093/nar/gkaf1122> PMID: [41261852](#)

50. Gittrich MR, Sanderson CM, Wainaina JM, Noel CM, Leopold JE, Babusci E, et al. Isolation and characterization of 24 phages infecting the plant growth-promoting rhizobacterium *Klebsiella* sp. M5al. *PLoS One*. 2025;20(2):e0313947. <https://doi.org/10.1371/journal.pone.0313947> PMID: 39982899
51. Mora-Quilis L, Sanjuán R, Domingo-Calap P. Reversible phenotypic resistance to phage infection via capsule downregulation in *Klebsiella pneumoniae*. *iScience*. 2025;28(12):114107. <https://doi.org/10.1016/j.isci.2025.114107> PMID: 41438077
52. Yu Z, Li S, Li Y, Jiang Z, Zhou J, An Q. Complete genome sequence of N2-fixing model strain *Klebsiella* sp. nov. M5al, which produces plant cell wall-degrading enzymes and siderophores. *Biotechnol Rep (Amst)*. 2017;17:6–9. <https://doi.org/10.1016/j.btre.2017.11.006> PMID: 29234606
53. Bao G, Zhang Y, Du C, Chen Z, Li Y, Cao Z, et al. Genome Sequence of *Klebsiella oxytoca* M5al, a Promising Strain for Nitrogen Fixation and Chemical Production. *Genome Announc*. 2013;1(1):e00074–12. <https://doi.org/10.1128/genomeA.00074-12> PMID: 23405358
54. Pengra RM, Wilson PW. Physiology of nitrogen fixation by *Aerobacter aerogenes*. *J Bacteriol*. 1958;75(1):21–5. <https://doi.org/10.1128/jb.75.1.21-25.1958> PMID: 13513555
55. Bertozzi Silva J, Storms Z, Sauvageau D. Host receptors for bacteriophage adsorption. *FEMS Microbiol Lett*. 2016;363(4). <https://doi.org/10.1093/femsle/fnw002>
56. Rai AK, Mitchell AM. Enterobacterial Common Antigen: Synthesis and Function of an Enigmatic Molecule. *mBio*. 2020;11(4):e01914–20. <https://doi.org/10.1128/mBio.01914-20> PMID: 32788387
57. Xue Y, Shi F, Zhou B, Shi Y, Luo W, Zhu J. Biofilm formation, antibiotic resistance, and virulence analysis of human and avian origin *Klebsiella pneumoniae* from Jiangsu, China. *Vet Sci*. 2025;12(7):628. <https://doi.org/10.3390/vetsci12070628>
58. Knecht LE, Heinrich N, Born Y, Felder K, Pelludat C, Loessner MJ, et al. Bacteriophage S6 requires bacterial cellulose for *Erwinia amylovora* infection. *Environ Microbiol*. 2022;24(8):3436–50. <https://doi.org/10.1111/1462-2920.15973> PMID: 35289468
59. Römmling U. The power of unbiased phenotypic screens - cellulose as a first receptor for the Schitoviridae phage S6 of *Erwinia amylovora*. *Environ Microbiol*. 2022;24(8):3316–21. <https://doi.org/10.1111/1462-2920.16010> PMID: 35415924
60. Hyman P, van Raaij M. Bacteriophage T4 long tail fiber domains. *Biophys Rev*. 2017;10(2):463–71. <https://doi.org/10.1007/s12551-017-0348-5> PMID: 29204885
61. Gencay YE, Gambino M, Prüssing TF, Brøndsted L. The genera of bacteriophages and their receptors are the major determinants of host range. *Environ Microbiol*. 2019;21(6):2095–111. <https://doi.org/10.1111/1462-2920.14597> PMID: 30888719
62. Gaborieau B, Vaysset H, Tesson F, Charachon I, Dib N, Bernier J, et al. Prediction of strain level phage-host interactions across the *Escherichia* genus using only genomic information. *Nat Microbiol*. 2024;9(11):2847–61. <https://doi.org/10.1038/s41564-024-01832-5> PMID: 39482383
63. Piel D, Bruto M, Labreuche Y, Blanquart F, Goudenège D, Barcia-Cruz R, et al. Phage-host coevolution in natural populations. *Nat Microbiol*. 2022;7(7):1075–86. <https://doi.org/10.1038/s41564-022-01157-1> PMID: 35760840
64. Meyer JR, Dobias DT, Weitz JS, Barrick JE, Quick RT, Lenski RE. Repeatability and contingency in the evolution of a key innovation in phage lambda. *Science*. 2012;335(6067):428–32. <https://doi.org/10.1126/science.1214449>
65. Beamud B, García-González N, Gómez-Ortega M, González-Candelas F, Domingo-Calap P, Sanjuan R. Genetic determinants of host tropism in *Klebsiella* phages. *Cell Rep*. 2023;42(2):112048. <https://doi.org/10.1016/j.celrep.2023.112048> PMID: 36753420
66. Wang C, Wang S, Jing S, Zeng Y, Yang L, Mu Y, et al. Data-Driven Engineering of Phages with Tunable Capsule Tropism for *Klebsiella pneumoniae*. *Adv Sci (Weinh)*. 2024;11(33):e2309972. <https://doi.org/10.1002/advs.202309972> PMID: 38937990
67. Townsend EM, Kelly L, Gannon L, Muscatt G, Dunstan R, Michniewski S, et al. Isolation and Characterization of *Klebsiella* Phages for Phage Therapy. *Phage (New Rochelle)*. 2021;2(1):26–42. <https://doi.org/10.1089/phage.2020.0046> PMID: 33796863
68. Nicolas M, Trotereau A, Culot A, Moodley A, Atterbury R, Wagemans J, et al. Isolation and Characterization of a Novel Phage Collection against Avian-Pathogenic *Escherichia coli*. *Microbiol Spectr*. 2023;11(3):e0429622. <https://doi.org/10.1128/spectrum.04296-22> PMID: 37140373
69. Moller AG, Lindsay JA, Read TD. Determinants of Phage Host Range in *Staphylococcus* Species. *Appl Environ Microbiol*. 2019;85(11):e00209–19. <https://doi.org/10.1128/AEM.00209-19> PMID: 30902858
70. Cazares D, Cazares A, Figueroa W, Guarneros G, Edwards RA, Vinuesa P. A novel group of promiscuous podophages infecting diverse gammaproteobacteria from river communities exhibits dynamic intergenus host adaptation. *mSystems*. 2021;6(1). <https://doi.org/10.1128/MSYSTEMS.00773-20>
71. Holmfeldt K, Howard-Varona C, Solonenko N, Sullivan MB. Contrasting genomic patterns and infection strategies of two co-existing Bacteroidetes podovirus genera. *Environ Microbiol*. 2014;16(8):2501–13. <https://doi.org/10.1111/1462-2920.12391> PMID: 24428166
72. Nilsson E, Bayfield OW, Lundin D, Antson AA, Holmfeldt K. Diversity and Host Interactions Among Virulent and Temperate Baltic Sea Flavobacterium Phages. *Viruses*. 2020;12(2):158. <https://doi.org/10.3390/v12020158> PMID: 32019073
73. Lourenço M, Osbelt L, Passet V, Gravey F, Megrian D, Strowig T. Phages against noncapsulated *Klebsiella pneumoniae*: broader host range, slower resistance. *Microbiol Spectr*. 2023:e04812–22. <https://doi.org/10.1128/spectrum.04812-22>
74. Loney RE, Delesalle VA, Chaudry BE, Czerpak M, Guffey AA, Goubet-McCall L, et al. A Novel Subcluster of Closely Related Bacillus Phages with Distinct Tail Fiber/Lysin Gene Combinations. *Viruses*. 2023;15(11):2267. <https://doi.org/10.3390/v15112267> PMID: 38005943

75. Cotter PA, Chepuri V, Gennis RB, Gunsalus RP. Cytochrome o (cyoABCDE) and d (cydAB) oxidase gene expression in *Escherichia coli* is regulated by oxygen, pH, and the *fnr* gene product. *J Bacteriol.* 1990;172(11):6333–8. <https://doi.org/10.1128/jb.172.11.6333-6338.1990> PMID: 2172211
76. Hernández Villamizar S, Chica Cárdenas LA, Morales Mancera LT, Vives Florez MJ. Anaerobiosis, a neglected factor in phage-bacteria interactions. *Applied and Environmental Microbiology.* 2023;89(12):e01491–23. <https://doi.org/10.1128/aem.01491-23> PMID: 37966212
77. Pacheco T, Gomes AÉI, Siqueira NMG, Assoni L, Darrieux M, Venter H, et al. SdiA, a quorum-sensing regulator, suppresses fimbriae expression, biofilm formation, and quorum-sensing signaling molecules production in *Klebsiella pneumoniae*. *Front Microbiol.* 2021;12:597735. <https://doi.org/10.3389/fmicb.2021.597735> PMID: 34234747
78. Szczepaniak J, Press C, Kleanthous C. The multifarious roles of Tol-Pal in Gram-negative bacteria. *FEMS Microbiol Rev.* 2020;44(4):490–506. <https://doi.org/10.1093/femsre/fuaa018> PMID: 32472934
79. Munder F, Voutsinos M, Hantke K, Venugopal H, Grinter R. High-affinity PQQ import is widespread in Gram-negative bacteria. *Sci Adv.* 2025;11(22):eadr2753. <https://doi.org/10.1126/sciadv.adr2753> PMID: 40446051
80. Bai J, Raustad N, Denoncourt J, van Opijnen T, Geisinger E. Genome-wide phage susceptibility analysis in *Acinetobacter baumannii* reveals capsule modulation strategies that determine phage infectivity. *PLoS Pathog.* 2023;19(6):e1010928. <https://doi.org/10.1371/journal.ppat.1010928> PMID: 37289824
81. Berry JD, Rajaure M, Young R. Spanin function requires subunit homodimerization through intermolecular disulfide bonds. *Mol Microbiol.* 2013;88(1):35–47. <https://doi.org/10.1111/mmi.12167> PMID: 23387988
82. Vieira MF, Duarte J, Domingues R, Oliveira H, Dias O. PhageDPO: A machine-learning based computational framework for identifying phage depolymerases. *Comput Biol Med.* 2025;188:109836. <https://doi.org/10.1016/j.compbimed.2025.109836> PMID: 39951981
83. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol.* 1990;215(3):403–10. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2) PMID: 2231712
84. Söding J, Biegert A, Lupas AN. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* 2005;33(Web Server issue):W244–8. <https://doi.org/10.1093/nar/gki408> PMID: 15980461
85. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc.* 2015;10(6):845–58. <https://doi.org/10.1038/nprot.2015.053> PMID: 25950237
86. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature.* 2021;596(7873):583–9. <https://doi.org/10.1038/s41586-021-03819-2> PMID: 34265844
87. Price M. PaperBLAST: Text Mining Papers for Information about Homologs. *mSystems.* 2017;2(4). <https://doi.org/10.1128/msystems.00039-17>
88. Regué M, Izquierdo L, Fresno S, Jiménez N, Piqué N, Corsaro MM. The incorporation of glucosamine into enterobacterial core lipopolysaccharide: two enzymatic steps are required. *J Biol Chem.* 2005;280(44):36648–56. <https://doi.org/10.1074/jbc.M506278200>
89. Fresno S, Jiménez N, Canals R, Merino S, Corsaro MM, Lanzetta R, et al. A second galacturonic acid transferase is required for core lipopolysaccharide biosynthesis and complete capsule association with the cell surface in *Klebsiella pneumoniae*. *J Bacteriol.* 2007;189(3):1128–37. <https://doi.org/10.1128/JB.01489-06> PMID: 17142396
90. Rubin BE, Diamond S, Cress BF, Crits-Christoph A, Lou YC, Borges AL, et al. Species- and site-specific genome editing in complex bacterial communities. *Nat Microbiol.* 2022;7(1):34–47. <https://doi.org/10.1038/s41564-021-01014-7> PMID: 34873292
91. Electrocompetent Cells. <https://barricklab.org/twiki/bin/view/Lab/ProtocolsElectrocompetentCells> Accessed 2025 August 6.
92. Datta S, Costantino N, Court DL. A set of recombineering plasmids for gram-negative bacteria. *Gene.* 2006;379:109–15. <https://doi.org/10.1016/j.gene.2006.04.018> PMID: 16750601
93. Vo PLH, Ronda C, Klompe SE, Chen EE, Acree C, Wang HH, et al. CRISPR RNA-guided integrases for high-efficiency, multiplexed bacterial genome engineering. *Nat Biotechnol.* 2021;39(4):480–9. <https://doi.org/10.1038/s41587-020-00745-y> PMID: 33230293
94. Vegan Tutorial. <https://peat-clark.github.io/BIO381/veganTutorial.html> Accessed 2025 July 15.
95. Kruskal JB. Nonmetric Multidimensional Scaling: A Numerical Method. *Psychometrika.* 1964;29(2):115–29. <https://doi.org/10.1007/bf02289694>
96. Wickham H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. 2016.
97. Dinno A. dunn.test: Dunn's Test of Multiple Comparisons Using Rank Sums. <https://cran.r-project.org/web/packages/dunn.test/index.html> 2024. Accessed 2025 July 15.
98. Dunn OJ. Multiple Comparisons Using Rank Sums. *Technometrics.* 1964;6(3):241–52. <https://doi.org/10.1080/00401706.1964.10490181>
99. Anderson MJ. Permutational Multivariate Analysis of Variance (PERMANOVA). Wiley StatsRef: Statistics Reference Online. John Wiley & Sons, Ltd. 2017:1–15. <https://doi.org/10.1002/9781118445112.stat07841>
100. Martínez Arbizu P. Pairwise multilevel comparison using adonis. <https://github.com/pmartinezarbizu/pairwiseAdonis> Accessed 2025 July 15.
101. Letunic I, Bork P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res.* 2024;52(W1):W78–82. <https://doi.org/10.1093/nar/gkae268> PMID: 38613393