High-throughput mapping of the phage resistance landscape in *E. coli*

Vivek K. Mutalik*[1,2], Benjamin A. Adler[2,3], Harneet S. Rishi[4,5], Denish Piya[2,3], Crystal Zhong[1], Britt Koskella[6], Richard Calendar[7], Pavel Novichkov[1], Morgan N. Price[1], Adam M. Deutschbauer[1,2,8], Adam P. Arkin*[1,2,3]

[1]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA, 94720, USA;

[2]Innovative Genomics Institute, Berkeley, CA, 94704, USA;

[3]Department of Bioengineering, University of California - Berkeley, Berkeley, CA, 94720, USA;

[4] Biophysics Graduate Group, University of California - Berkeley, Berkeley, CA, 94720, USA;

[5] Designated Emphasis Program in Computational and Genomic Biology, University of California - Berkeley, Berkeley, CA, 94720, USA

[6]Department of Integrative Biology, University of California - Berkeley, Berkeley, CA, 94720, USA;

[7]Department of Molecular and Cell Biology, University of California - Berkeley, Berkeley, CA, 94720, USA;

[8]Department of Plant and Microbial Biology, University of California - Berkeley, Berkeley, CA, 94720, USA;

*Correspondence:

vkmutalik@lbl.gov and aparkin@lbl.gov

Abstract

Bacteriophages (phages) are critical players in the dynamics and function of microbial communities and drive processes as diverse as global biogeochemical cycles and human health. Phages tend to be predators finely tuned to attack specific hosts, even down to the strain level, which in turn defend themselves using an array of mechanisms. However, to date, efforts to rapidly and comprehensively identify bacterial host factors important in phage infection and resistance have yet to be fully realized. Here, we globally map the host genetic determinants involved in resistance to 14 phylogenetically diverse double-stranded DNA phages using two model *Escherichia coli* strains (K-12 and BL21) with known sequence divergence to demonstrate strain-specific differences. Using genome-wide loss-of-function and gain-of-function genetic technologies, we are able to confirm previously described phage receptors as well as uncover a number of previously unknown host factors that confer resistance to one or more of these phages. We uncover differences in resistance factors that strongly align with the susceptibility of K-12 and BL21 to specific phage. We also identify both phage specific mechanisms, such as the unexpected role of cyclic-di-GMP in host sensitivity to phage N4, and more generic defenses, such as the overproduction of colanic acid capsular polysaccharide that defends against a wide array of phages. Our results indicate that host responses to phages can occur via diverse cellular mechanisms. Our systematic and high-throughput genetic workflow to characterize phage-host interaction determinants can be extended to diverse bacteria to generate datasets that allow predictive models of how phage-mediated selection will shape bacterial phenotype and evolution. The results of this study and future efforts to map the phage resistance landscape will lead to new insights into the coevolution of hosts and their phage, which can ultimately be used to design better phage therapeutic treatments and tools for precision microbiome engineering.

Introduction:

Bacterial viruses, or bacteriophages (phages), are obligate parasites that infect specific bacterial strains. Phages represent the most abundant biological entities on earth, and are key ecological drivers of microbial community dynamics, activity and adaptation; thereby impacting environmental nutrient cycles, agricultural output, and human and animal health [1–8]. Despite nearly a century of pioneering molecular work, the mechanistic insights into phage specificity for a given host, infection pathways, and the breadth of bacterial responses to different phages have largely focused on a handful of individual bacterium-phage systems [9–13]. Bacterial sensitivity/resistance to phages is typically characterized using phenotypic methods such as cross-infection patterns against a panel of phages [14–27] or by whole-genome sequencing of phage-resistant mutants [28–32]. As such, our understanding of bacterial resistance mechanisms against phages remains limited, and the field is therefore in need of improved methods to characterize phage-host interactions, determine the generality and diversity of phage resistance mechanisms in nature, and identify the degree of specificity for each bacterial resistance mechanism across diverse phage types [13,25,26,33–47].

Unbiased and comprehensive genetic screens that are easily transferable and scalable to new phage-host combinations would be highly valuable for obtaining a deeper understanding of phage infection pathways and phage-resistance phenotypes [48–53]. Such genome-scale studies applied to different phage-host combinations have the unique potential to identify commonalities or differences in phage resistance patterns and mechanisms [18,25,28,54–56]. There have been few attempts to use genetic approaches for studying genome-wide host factors essential in phage infection. These loss-of-function (LOF) genetic screens broadly use bacterial saturation mutagenesis [48,53,57–60] or an arrayed library of single-gene deletion strains for studying phage-host interactions [49,50,52,61,62]. Consequently, these studies have generally involved laborious experiments on relatively few phages and their hosts, and scaling the approach to characterize hundreds of phages is challenging.

To overcome these technological limitations, we have developed three high-throughput genetic technologies that enable fast, economical and quantitative genome-wide screens for gene function, which are suitable for discovering host genes critical for phage infection and bacterial resistance. Random barcode transposon site sequencing (RB-TnSeq) allows genome-wide insertional mutagenesis leading to LOF mutations [63]; a pooled CRISPR interference (CRISPRi) approach, which allows partial inhibition of gene function via transcriptional inhibition [64]; and dual-barcoded shotgun expression library sequencing (Dub-seq) [65], which queries the effects of gene over-expression. All three technologies can be assayed across many conditions at low-cost as RB-TnSeq and Dub-seq use randomized DNA barcodes to assay strain abundance (Barseq [66]); while quantification of the pooled CRISPRi strains only requires deep sequencing of the guide sequences.

In RB-TnSeq, genome-wide transposon insertional mutant libraries labeled with unique DNA barcodes are generated, and next-generation sequencing methods are used to map the transposon insertions and DNA barcode at loci in genomes. Although RB-TnSeq can be applied on a large scale across multiple bacteria through barcode sequencing [67], it is limited to non-essential genes. Partial loss-of-function (LOF) assays such as CRISPRi use a catalytic null mutant of the Cas9 protein (dCas9) guided by chimeric single-guide RNA (sgRNA) to programmably knockdown gene expression, thereby allowing the probing of all genes (including essential genes) and more precise targeting of intergenic regions [64]. We have recently applied CRISPRi technology to systematically query the importance of ~13,000

genomic features of *E. coli* in different conditions (Rishi et al 2020 submitted). CRISPRi has been applied to different organisms to study essential genes [68–76], and has been recently applied to *E. coli* to uncover host factors involved in T4, 186 and λ phage infection [51]. Lastly, Dub-Seq uses shotgun cloning of randomly sheared DNA fragments of a host genome on a dual-barcoded replicable plasmid and next-generation sequencing to map the barcodes to the cloned genomic regions. The barcode association with genomic fragments and genes contained on those fragments enables a parallelized gain-of-function (GOF) screen, as demonstrated in our recent work [65]. In contrast to LOF genetic screens, GOF screens to study gene dosage effects on phage resistance have not been broadly reported, except for a recent work on T7 phage using an *E. coli* single gene overexpression library [50]. There are indications that enhanced gene dosage can be an effective way to search for dominant-negative mutants, antisense RNAs, or other regulatory genes that may block phage growth cycle [9,10,13,50,55,77]. Such GOF screens, when applied in a high-throughput format across diverse phages, may yield novel mechanisms of phage resistance that LOF screens may not uncover.

In this study, we employ these three technologies (RB-TnSeq, Dub-seq and CRISPRi) as a demonstration of “portable” and “scalable” platforms for rapidly probing phage-host interaction mechanisms. Using *E. coli* K-12 strain and 14 diverse double-stranded DNA phages, we show that our screens successfully identify known receptors and other host factors important in infection pathways, and they also yield additional novel loci that contribute to phage resistance. We validate some of these new findings by deleting or overexpressing individual genes and quantifying fitness in the presence of phage. Additionally, we used RB-TnSeq and Dub-seq to compare similarity and distinctiveness in phage resistance displayed by *E. coli* strains K-12 and BL21. The comparison of two historical lineages of *E. coli* allowed us to examine how strain-level divergence of genotype can lead to differential susceptibility in phage resistance. Finally, we discuss the implications and extensibility of our approaches and findings to other bacteria-phage combinations and how these datasets can provide a foundation for understanding phage ecology and engineering phage for therapeutic applications.

Results and Discussion:

Mapping genetic determinants of phage resistance using high-throughput LOF and GOF methods

To demonstrate the efficacy of our approaches to characterize phage resistance mechanisms, we first used *E. coli* K-12 strains (BW25113 for RB-TnSeq and Dub-seq screens, and MG1655 for CRISPRi screen) as hosts (Fig.1). As described later, for comparative purposes, we also performed high-throughput genetic assays in the *E. coli* BL21 strain background. Despite *E. coli* being a well-studied model organism [78,79], there are significant knowledge gaps regarding gene function [80] and phage interaction mechanisms [26,27,46,81,82]. Different serotypes of *E. coli* are also important pathogens with significant global threat and are crucial players in specific human-relevant ecologies [83–85], leading to the question of whether strain variation is also important in predicting the response to phage-mediated selection or whether the mechanisms are likely to be conserved.

Here, we used a previously constructed *E. coli* K-12 BW25113 RB-TnSeq library [63] and defined *E. coli* K-12 MG1655 CRISPRi library (Rishi et al 2020 submitted). To study host gene dosage and overexpression effects on phage resistance, we used a prior reported GOF Dub-

seq library of *E. coli* K-12 BW25113 [65]. We sourced 14 diverse *E. coli* phages with double-stranded DNA genomes, belonging to Myoviridae, Podoviridae and Siphoviridae families (within the order Caudovirales) (Fig. 1). These phages include 11 canonical and well-studied coliphages each having overlapping but distinct mechanisms of host recognition, entry, replication and host lysis [86], and two recently reported phages, which are known to kill pathogenic shiga-toxin producing *E. coli* (STEC, O157:H7), and one novel coliphage. These 14 phages include T-series phages (T2, T3, T4, T5, T6, T7), N4, 186, λ*c*I857, P1*vir*, P2, and novel isolates of T-like phages (T6-like LZ4, STEC infecting T4-like CEV1 and T5-like CEV2).

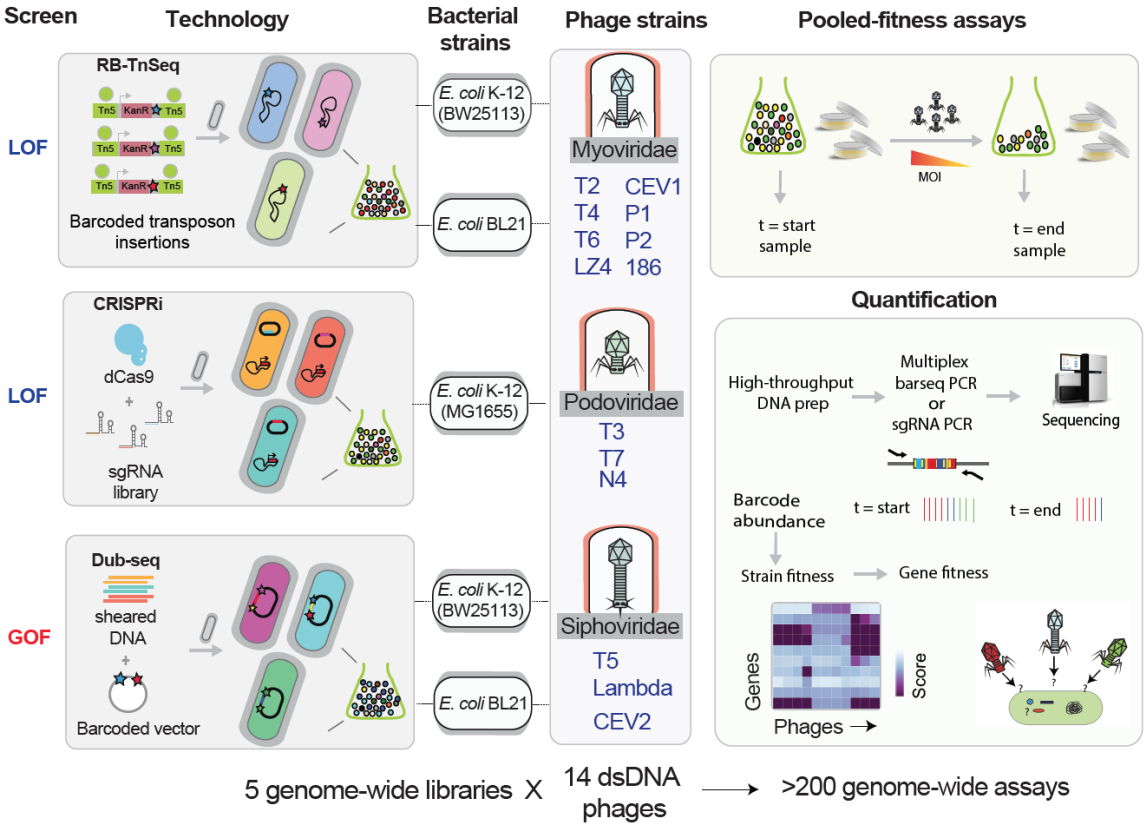


Figure 1. Overview of high-throughput genome-wide screens. We used barcoded loss-of-function (LOF) technologies (RB-TnSeq and CRISPRi) and a gain-of-function (GOF) technology (Dub-seq) in *E. coli* K-12 (BW25113 and MG1655) to screen for host factors important in phage infection and resistance. In *E. coli* BL21, we performed RB-TnSeq and Dub-seq (but not CRISPRi). We sourced 14 diverse *E. coli* phages with double-stranded DNA genomes, belonging to Myoviridae, Podoviridae and Siphoviridae families, and performed pooled fitness screens in both planktonic and solid agar formats. Disruption or overexpression of certain genes provide fitness to host in the presence of phages and we monitor these changes by quantifying the abundance of the DNA barcode or sgRNA associated with each strain. The individual strain abundances are then converted to gene fitness scores (normalized log2 change in the abundance of mutants in that gene).

The barcoded LOF or GOF libraries are grown competitively in a single-pot assay and challenged with phages. Disruption or overexpression of certain genes provides different fitness dynamics to the strains in the libraries. Due to the selection pressure of the phage, these strains either are eliminated from the population or enriched, and we monitor these changes by quantifying the abundance of the DNA barcode associated with each strain. The individual strain abundances are then converted to gene fitness scores (Methods), which we define as the normalized log2 change in the abundance of mutants in that gene. Positive gene

fitness scores ("positive hits") indicate that the gene(s) disrupted or overexpressed lead to an increase in relative fitness in the presence of a particular phage (i.e. these genetic changes lead to strain being effectively resistant to the phage). A positive fitness score for a gene when disrupted indicates that the gene product is needed for phage infection cycle, while a positive fitness score for a gene when overexpressed indicates that particular gene (or a fragment encoding that gene) prevents the phage infection cycle. Negative fitness values, which suggest reduced relative fitness, indicate that the gene(s) disruption or overexpression result in these strains being more sensitive to the phage than the typical strain in the library. Lastly, fitness scores near zero indicate no fitness change for the mutated or overexpressed gene under the assayed condition. Due to the strong selection of phage infection, we anticipated (and indeed observed) that genetically modified strains in our libraries with resistance to a phage would lead to very strong positive fitness values. While these very strong positive phenotypes are readily interpretable, one consequence is that strains with relatively poor or neutral fitness scores will be swept from the population. Thus intermediate resistance factors to phage infection will have similar low or negative fitness values as a neutral mutant. As such, most of our focus in this study is on the strong positive fitness scores (Methods). Challenges in identifying intermediate phage-resistant mutants in the presence of highly resistant phage receptor mutations are well appreciated in the field [87].

RB-TnSeq identifies known receptors and host factors for all 14 phages

To perform genome-wide transposon-based LOF assays, we recovered a frozen aliquot of the *E. coli* K-12 RB-TnSeq library in Lysogeny Broth (commonly known as LB [88]) to mid-log phase, collected a cell pellet for the "start", and used the remaining cells to inoculate an LB culture supplemented with different dilutions of a phage in SM buffer. After 8 hrs of phage infection in planktonic cultures, we collected the surviving phage-resistant strains or "end" samples (Methods). We also repeated these fitness assays on solid media by plating the library post-phage adsorption, incubating the plates overnight and collecting all surviving phage-resistant colonies. We hypothesized that, given the spatial structure and possibility of phage refuges, fitness experiments on solid media might provide a less stringent selection environment than in liquid pooled assays, such that less fit survivors could potentially be detected. For all assays, recovered genomic DNA from surviving strains was used as templates to PCR amplify the barcodes for sequencing (Methods). The strain fitness and gene fitness scores were then calculated as previously described [63].

In total, we performed 68 RB-TnSeq assays across 14 phages at varying multiplicity of infection (MOI) and 9 no-phage control assays (Methods). For planktonic assays, the gene fitness scores were reproducible across different phage MOIs (Fig. 2A), with a median pairwise correlation of 0.90. Because of stronger positive selection in the presence of phages (relative to our typical RB-TnSeq fitness assays with stress compounds and in defined growth media [63,67]), to determine reliable effects across experiments we used stringent filters (gene fitness score ≥ 5, t-like statistic ≥ 5 and estimated standard error ≤ 2) (Methods). Across all replicate experiments, we identified 354 high-scoring hits, which represent 52 unique genes and 133 unique gene x phage combinations (some genes were linked with more than phage)(Supplementary Table S1). In all phage experiments there was at least one gene with a high-confidence effect. In the presence of some phages (for example, T5 and T6), we observed enrichment of strains with disruption in only one gene (encoding the phage receptor),

while other phages (such as 186 and λ cI857 phages) showed enrichment with disruptions in multiple genes (Fig. 2A).

Phage infection initiates through an interaction of the phage with receptors on the outer surface of the cell envelope, which could be either lipopolysaccharide (LPS) moieties and/or proteinaceous components. Any changes in levels or structure of these receptors can compromise efficient phage infection, thereby leading to an improvement in host fitness in the presence of phages. Consequently, receptor mutants are the most commonly found candidates that display phage resistance [48,49,52,57,87,89]. To confirm the validity of our approach, we looked for receptors recognized by many of the canonical phages used in this study for which there is published data available [49,50,90–96]. Indeed, we found high fitness scores (fitness score > 10, corresponding to >1,000-fold enrichment of transposon mutants) for multiple known phage receptors (Fig. 2, Supplementary Table S1-S3). These included genes coding for protein receptors such as *fadL* (long-chain fatty acid transporter) for T2, *ompC* (outer membrane porin C) for T4, *fhuA* (ferrichrome outer membrane transporter) for T5, *tsx* (nucleoside-specific transporter) for T6, *nfrA* and *nfrB* (unknown function) for N4 and *lamB* (maltose outer membrane transporter) for λ. *fhuA* showed high fitness scores in presence of CEV2 phage, indicating CEV2 has similar receptor requirement as T5 [97]. We find *tsx* as the top scorer in the presence of novel LZ4 phage, and thus appears to have similar receptor requirement as T6 phage [98]. In addition to protein receptors, we also identified a few high-scoring genes that are known to interfere or regulate the expression of receptors, thereby impacting phage infection. For example, deletion of the EnvZ/OmpR two-component system involved in the regulation of *ompC* and gene products involved in regulation of *lamB* expression (*cyaA*, *malI*, *malT*) all show high fitness scores in presence of T4 and λ phages, respectively (Fig. 2A).

For phages utilizing LPS as their receptors, we found top scores for gene mutations within the *waa* gene cluster, which codes for enzymes involved in LPS core biosynthesis (Fig. 2). For example, *waaC*, *waaD*, *waaE* and *lpcA/gmhA* were the top scorers for T3 and T7 phages, while *waaC*, *waaD*, *waaE*, *waaF*, *waaJ*, *waaO*, *waaQ*, and *lpcA/gmhA* showed high fitness in presence of P1, P2 and 186 phages (Fig. 2A and Fig. 2B). Even though P1 and P2 phages have been studied for decades, the host factors important in their infection cycle are not fully characterized [91,99]. Our results show that all LPS core components are essential for an efficient P1, P2 and 186 phage infection. CEV1 phage seems to require both outer membrane porin OmpF and LPS core for efficient infection. In addition to *ompF* and its regulator the *envZ/ompR* two-component signaling system, a number of genes involved in the LPS core biosynthesis pathway (*waaC*, *waaD*, *waaE*, *waaF*, *waaG*, *waaP, galU* and *lpcA/gmhA*) and a regulator of genes involved in biosynthesis, assembly, and export of LPS core (*rfaH*) all showed high fitness scores (>10) in the presence of CEV1 phage. Among other top-scoring hits, genes encoding a putative L-valine exporter subunit (*ygaH*) and a diguanylate cyclase (*dgcJ*) showed stronger fitness in presence of N4 phage. Both YgaH and DgcJ were not previously known to be involved in N4 phage resistance. Finally, *igaA/yrfF,* which encodes a negative regulator of the Rcs phosphorelay pathway, shows strong fitness scores against 8 phages, indicating its importance to general phage resistance. Though *igaA* is an essential gene in *E. coli* [100], our RB-TnSeq library contained 9 disruptions in *igaA*'s cytoplasmic domain and these strains seem to tolerate the disruption (Supplementary Fig. S1). It is known that the Rcs signal transduction pathway functions as an envelope stress response system that monitors cell surface composition and regulates a large number of genes involved in diverse functions including colanic acid synthesis and biofilm formation [101].

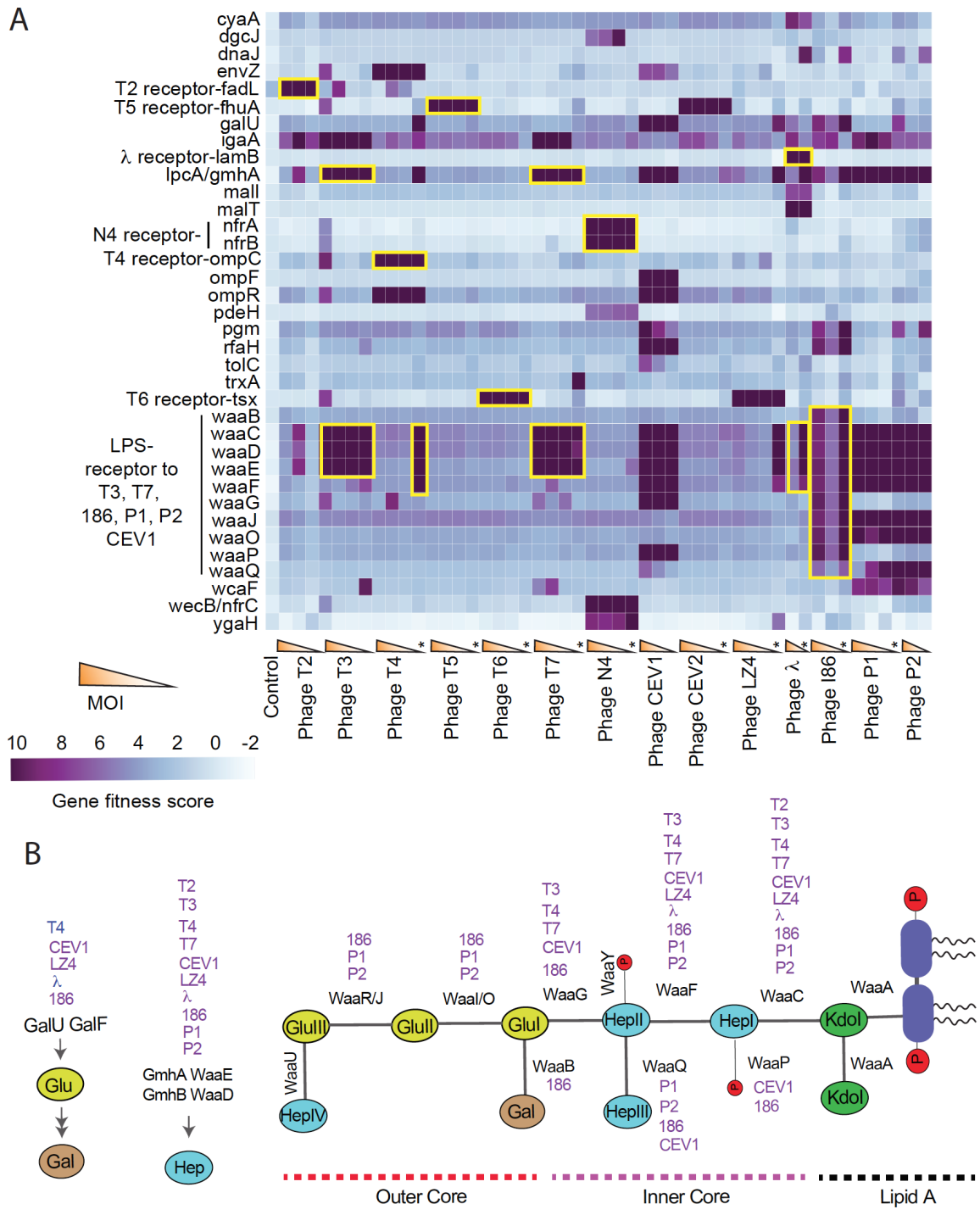


Figure 2. Heatmap of *E. coli* K-12 RB-TnSeq data for 14 dsDNA phages at different MOI. (A) Top 36 genes with high-confidence effects and a gene fitness score of >=6.5 in at least one phage assay are shown. The pooled fitness assays performed on solid agar plates are shown as stars. Yellow boxes highlight genes that encode known receptors for the marked phages. (B) Schematic of *E. coli* K-12 LPS structure with associated enzymes involved in LPS core biosynthesis. Top scoring candidates in the presence of a particular phage (at any MOI) are highlighted by associating each enzymatic step with phages.

Compared to assays performed in planktonic cultures, a few gene mutants showed stronger fitness effects on plate assays. For example, *trxA,* encoding thioredoxin 1, is known to be essential for T7 phage propagation and scored high in the solid plate assays but not in our planktonic growth assays (Fig. 2). Thioredoxin 1 is a processivity factor for T7 RNA polymerase and it is reported that T7 phage can bind and lyse a *trxA* deletion strain, though T7 phage propagation is severely compromised [102]. Similarly, we observed higher fitness scores for *galU* in the presence of T4, P1 and λ phages, and *dnaJ* in the presence of 186, P1 and λ phages, on plates but not in broth. GalU catalyzes the synthesis of UDP-D-glucose, a central precursor for synthesis of cell envelope components, including LPS core, and it is known that the growth of P1 phage is compromised on a *galU* mutant [99]*. dnaJ* codes for a chaperone protein and *dnaJ* insertion mutants are known to inhibit the growth of λ phage [103–105]. We also observed higher fitness scores for a number of genes involved in LPS biosynthesis (*waaC*, *waaD*, *waaE*, *waaF, lpcA*) in the presence of T4, LZ4 and λ phages when grown on solid plates. These results suggest that LPS components either play an important role in efficient phage infection cycle or these LPS truncations lead to a destabilized membrane and probably decrease outer membrane protein levels via envelope stress response [101,106–109]. A detailed description about many of the genes we identified in this study is included as a Supplemental Note. In summary, our results indicate how the abiotic environment can have an important influence on the host fitness and susceptibility, and the type of resistance mechanism selected in the presence of different phages[110–114].

Overall, our RB-TnSeq LOF screen provided a number of top hits that agree with earlier reports and also yielded a set of novel genes whose role in phage infection was not previously known (Supplementary Table S1-S3). Later in this manuscript, we describe follow-up experiments with 18 of these top-scoring hits from the RB-TnSeq screen to validate their role in phage resistance.

A CRISPRi screen provides deeper view of phage resistance determinants

We next employed a rationally designed *E. coli* K-12 MG1655 genome-wide CRISPRi library approach to look for bacterial essential and nonessential genes and genomic regulatory regions important in phage infection and to provide a complementary genetic screen to RB-TnSeq. This CRISPRi library comprises multiple sgRNAs targeting annotated genes, promoter regions and transcription factor binding sites, a total of 13,000 target regions (Rishi et al 2020, Methods). By directing dCas9 to different genomic regions via unique sgRNAs, CRISPRi enables the interrogation of genic and non-genic regions. For the CRISPRi assays, we recovered a frozen aliquot of the library, which was back diluted in fresh media, supplemented with dCas9 and gRNA inducers, and mixed with phage. We assayed T2, T3, T4, T5, T6, N4, 186, CEV1, CEV2, LZ4 and λ phages in planktonic cultures at an MOI of 1, recovered survivors after phage treatment for 3 hrs, isolated plasmid DNA, and the variable gRNA region was PCR amplified and sequenced (Methods). In total, we identified 542 genes (including 75 genes of unknown function), 94 promoter regions and 44 transcription factor binding sites that show high fitness scores across all phages (sgRNAs with log2 fold change >= 2 and false discovery rate-adjusted p-value<0.05) (Supplementary Table S4, Methods).

To confirm our assay system correctly identifies host factors important in phage infection we looked for genes that are known to impact phage infection and also, are in agreement with high fitness scores in our RB-TnSeq results. Indeed, the top-scoring hits included sgRNAs that target both the genes coding for phage receptors and their promoter and transcription factor

binding sites (Fig. 3A, Supplementary Table S4). Our CRISPRi data also confirmed many of the top-scoring genes uncovered in RB-TnSeq screen (Fig. 3A, Supplementary Table S3), thus validating the importance of these genes in specific phage infection pathways. For example, *ompF* was the top-scoring genes for CEV1 phage, *tsx* showed high fitness scores in presence of LZ4 whereas *fhuA* was the top scorer for phage CEV2. We also found that sgRNAs targeting *dgcJ* or its promoter showed high fitness scores in the presence of N4 phage, thereby confirming the RB-TnSeq data (Fig. 2A). In agreement with RB-TnSeq data, we observed that knockdown of *igaA* yields resistance to 10/11 phages we assayed (Figs. 3B-3D).

Among disagreements between the RB-TnSeq and CRISPRi datasets, we found that our CRISPRi screen failed to return some of the highest-scoring genes uncovered in RB-TnSeq dataset. These include genes encoding the EnvZ/OmpR two-component system for T4 and CEV1 phages, and YgaH and WecB for N4 phage. Here we note that, we have only one sgRNA targeting *ygaH*, and no sgRNAs targeting *wecB/nfrC* region in our CRISPRi library. In addition to these genes, the contribution of LPS core biosynthesis genes in phage infection was less clear in our CRISPRi dataset. For example, both OmpF and LPS seem to be crucial for CEV1 infection from our RB-TnSeq dataset (Fig. 2A), while the CRISPRi screen data showed high fitness score for *ompF* and not for all core LPS biosynthetic genes (Fig. 3A). Nevertheless, we find high fitness scores for genes encoding LPS transport system (*lptABC*) and lipid A biosynthesis enzymes (Fig. 3B). In summary, these results indicate that we might be missing few candidates in our CRISPRi screen (as compared to RB-TnSeq) because either some genes lack sufficient sgRNA coverage or even minimal expression of these genes is probably enough for phage infection.

One of the key advantages of CRISPRi is the ability to study the contribution of essential genes on phage infection. We found 11 essential genes (*csrA*, *kdsC*, *lptA*, *lptB*, *lptC*, *lpxA*, *lpxC*, *nusG*, *secE, yejM* and *tsf*) that showed broad fitness advantage (fitness score >=2 in more than one phage assay) when downregulated (Fig. 3B). None of these essential genes are present in our RB-TnSeq library, except for *yejM*, which has transposon insertions in the C-terminal portion (after 5 putative transmembrane helices) of the protein. The putative cardiolipin transporter encoded by *yejM* and its upstream neighbor *yejL* both show enhanced fitness in the presence of T3, T4, T6, CEV1, CEV2 and LZ4 phages in the CRISPRi dataset (Fig. 3A, 3B). These genes have not been previously associated with phage resistance. Although the physiological role of cardiolipin is still emerging, it is known that cardiolipins play an important role in outer membrane protein translocation system and membrane biogenesis [115–117]. A recent study showed that decreased cardiolipin levels activate Rcs envelope stress response [117]. Downregulation of cardiolipin transport probably results in phage resistance via increased colanic acid biosynthesis, but further mechanistic studies are needed.

Although the fitness benefit phenotype conferred by most top scoring essential genes in the presence of phages is challenging to interpret, some of these hits agree with previous work. For example, it is known that the downregulation of genes involved in transcription antitermination (*nusB, nusG*) and Sec translocon subunit E (*secE*) shows high fitness in presence of λ phage, and are crucial for the phage growth cycle [118–120]. *lptABC*, *kdsC*, and *lpxAC* are known to impact outer membrane biogenesis, LPS synthesis and transport [108,109,121,122]; while the RNA binding global regulator CsrA is known to be involved in carbohydrate metabolism and regulation of biofilm [123,124]. Downregulation of these genes likely leads to pleiotropic effects leading to enhanced fitness in the presence of phages. Our CRISPRi screen also identified a number of *E. coli* tRNA related genes showing enhanced

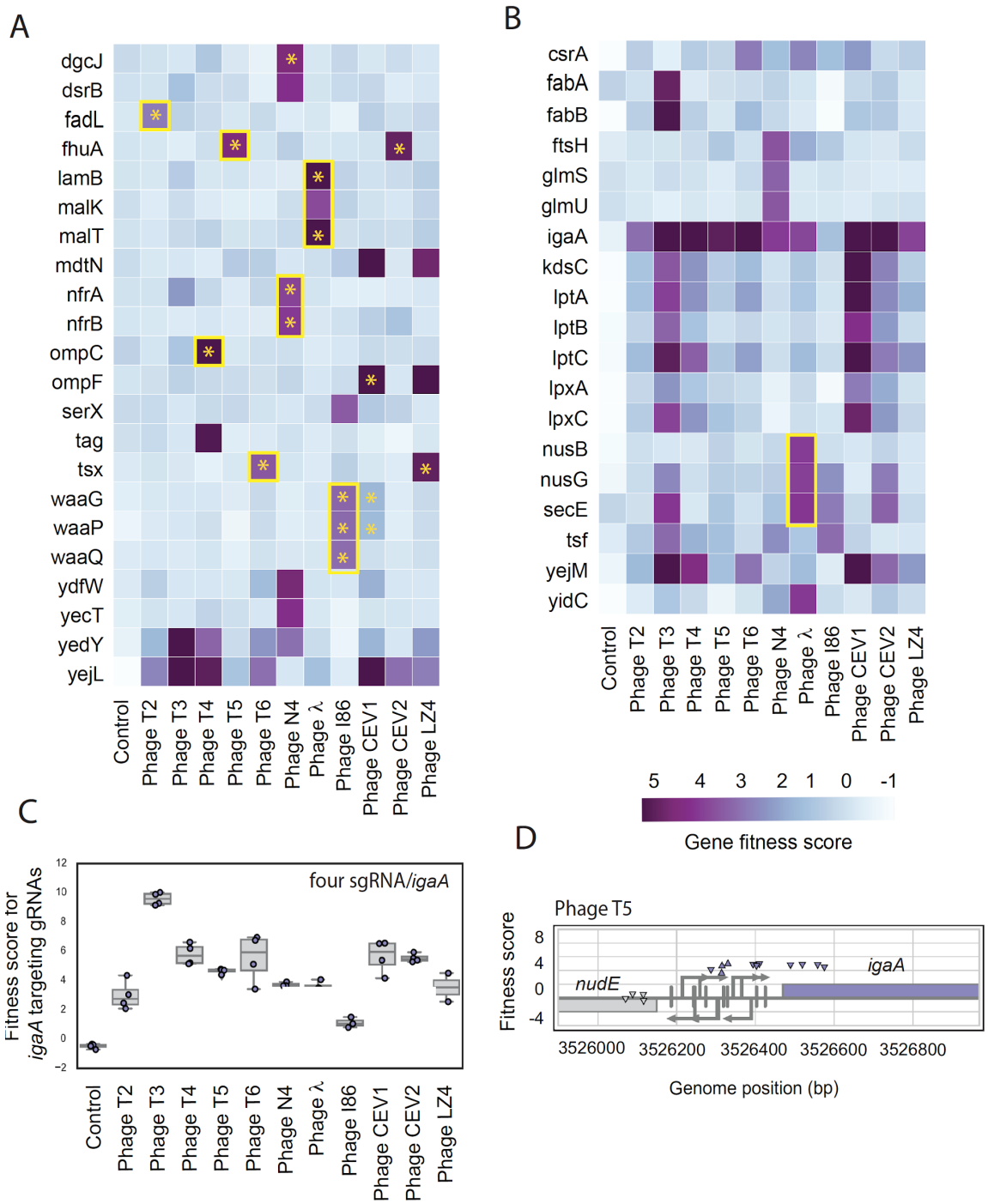


Figure 3. CRISPRi fitness profiles of *E. coli* non-essential and essential genes in the presence of diverse phages (A) Heatmap of top scoring non-essential genes across 11 phages. Yellow boxes highlight genes that encode phage receptors and are known to interfere with phage growth when downregulated. Yellow stars indicate these data points are in agreement with RB-TnSeq results. (B) Heatmap of top scoring essential genes across 11 phages. Yellow boxes highlight genes that are known to interfere with phage growth when downregulated. (C) Boxplot of fitness data for *igaA*-targeting sgRNAs across 11 different phages. Each data point in this plot is a specific sgRNA targeting *igaA*. (D) Genome browser plot for *nudE-igaA* locus with targeting sites for gRNAs and their fitness scores. The downward-facing triangles mean that the sgRNA targeted the non-template strand of the gene. Under each promoter, a vertical bar denotes the +1 for the promoter with the stem for the promoter starting at -60 relative to the transcription start site.

fitness in the presence of diverse phages (Supplementary Table S4). How the down-regulation of host tRNA and tRNA modification genes impact host fitness, phage growth and infection cycle is not clear, although recent studies have shown increased aminoacyl-tRNA synthetase activities in the early phage infection cycle [125,126].

Finally, neither the RB-TnSeq nor CRISPRi screens found high scores for *ompF* in the presence of T2 phage and *cmk* in the presence of T7 phage (Figs. 2A, 3A). Previous reports had indicated that OmpF serves as a secondary receptor to T2 (in addition to the primary receptor FadL)[127] and Cytidine monophosphate kinase (encoded by *cmk*) is an essential host factor in T7 phage infection [50]. Our genome-wide screens did not recapitulate these findings probably because these mutants provide an intermediate fitness in the presence of phages [128] and are swept from the population in the presence of highly resistant phage receptor mutations.

In summary, CRISPRi served as a complementary screening technology to RB-TnSeq, validating many RB-TnSeq hits and provided an avenue to study the role of essential genes and non-genic regions on phage infection.

Dub-seq enables parallel mapping of multicopy suppressors of phage infection

To study the effect of host factor gene dosage or overexpression on phage resistance, we assayed the *E. coli* K-12 Dub-seq library expressed in the *E. coli* K-12 strain. This library is made up of randomly sheared 3 kb genomic DNA cloned into a dual barcoded vector [65]. Similar to pooled fitness assays with the RB-TnSeq library, we recovered a frozen aliquot of the library to mid-log phase, collected a cell pellet for the initial sample, and used the remaining cells to inoculate an LB culture supplemented with different dilutions of a phage in SM buffer (Methods). Similar to LOF RB-TnSeq assays, after 8 hrs of phage infection in planktonic cultures (and overnight incubation in case of plate assays), we collected the surviving phage-resistant strains, isolated plasmid DNA, and sequenced the DNA barcodes (Methods). In total, we performed 67 genome-wide GOF competitive fitness assays in the presence of 13 different phages at varying MOIs, both in planktonic and solid plate assay format and 4 no-phage control experiments (Methods). Overall we identified 233 high scoring positive hits for the *E. coli* K-12 Dub-seq screen made up of 129 unique phages x gene combinations and found more than 5 Dub-seq hits per phage that confer positive growth benefit (fitness score >= 4, FDR of 0.7, Methods, Supplementary Table S5).

The growth benefit phenotypes we observe in Dub-seq assays may not be simply due to overexpression of genes encoded on genomic fragments, but other potential effects such as increased gene copy number (gene dosage) or other indirect dominant negative effects may be playing a role [129–134]. For example, overexpression or higher copy number of a regulatory region might lower the effective concentration of a transcription factor important in the regulation of phage receptor expression, and may lead to a phage-resistant strain. To confirm whether our method captures such host factors that control the expression of a phage receptor, we looked for known regulators. Acetylesterase Aes, transcription repressor Mlc, and *mal* regulon repressor MalY are known to reduce the expression of *mal* regulon activator *malT* or prevent MalT's activation of the λ phage receptor *lamB* [135–140]. Indeed, our Dub-seq screen identified *mlc*, *malY* and *aes* as the top scoring genes in the presence of λ phage, confirming that Dub-seq can identify host factors involved in regulating the expression of a phage receptor (Figs. 4A and 4B). We also found that overexpression of a gene encoding glucokinase (*glk*) and cyclic-AMP (cAMP) phosphodiesterase (*cpdA*) showed enhanced fitness

in the presence of λ phage (Supplementary Fig. S2). Glk has been proposed to inhibit *mal* regulon activator *malT [141]*, while CpdA hydrolyzes cAMP and negatively impacts cAMP-CRP regulated gene expression of *lamB [9,142,143]*.

One of the top candidates that broadly enhanced host fitness in the presence of diverse phages is transcriptional activator RcsA (that increases colanic acid production by inducing capsule synthesis gene cluster [101]). Genomic fragments with *rcsA* showed the highest gene score of +12 to +16 in most experiments (47/51 assays shown in Fig. 4A). In addition, we identified growth advantage with dozens of genes when overexpressed in presence of specific phages. For example, we found genomic fragments encoding *pdeO (dosP)*, *pdeR (gmr)*, *pdeN (rtn)*, *pdeL* (*yahA)*, *pdeC* (*yjcC)*, *pdeB (ylaB)*, *pdeI (yliE)*, *ddpX*, *flhD* and *yhbJ*/*rapZ* all confer resistance to N4 phage (Fig. 4A and 4C, Supplementary Figs. S4). Except for *ddpX*, *flhD* and *yhbJ*/*rapZ*, which encode D-Ala-D-Ala dipeptidase involved in peptidoglycan biosynthesis, flagellar transcriptional regulator and an RNase adaptor protein respectively, all others encode cyclic-di-GMP phosphodiesterases (PDEs). The PDEs are a highly conserved group of proteins in bacteria that catalyze the degradation of cyclic-di-GMP, a key secondary signaling molecule involved in biofilm formation, motility, virulence and other cellular processes [144–146]. Incidentally, increased dosage of *pdeN* (*rtn)* is known to confer resistance to N4 phage, albeit with unknown mechanism [147,148]. In addition to the N4 phage hits, we found that overexpression of *ygbE* and *ompF* showed high fitness scores in the presence of T4 phages, overexpression of small RNA *micF* showed high fitness scores for CEV1 phage, while overexpression of *ompT* gives high fitness scores in the presence of T3 and T7 phages (Fig. 4).

This is the first systematic analysis of how gene overexpression impacts phage resistance. While we do not understand all of the mechanisms leading to phage resistance, a few of these hits are consistent with the known biology of phage receptors. For example, expression of outer membrane porins *ompC* and *ompF* are regulated reciprocally by *ompR*, and increased *ompF* level reduces expression of the T4 phage receptor *ompC* [149–151]. Similarly, increased expression of the sRNA *micF* causing resistance to phage CEV1 (Supplementary Fig. S2) is consistent with a report that elevated levels of *micF* specifically downregulates *ompF* (CEV1 receptor) [152,153]. As *micF* is encoded within the intergenic region of *rcsD* (activator of Rcs pathway and colanic acid biosynthesis) and *ompC*, and also contains OmpR operator sites, the resistance-causing Dub-seq fragments containing *micF* could be acting via a combination of effects that cannot be resolved in our screen. Finally, overexpression of the *lit* gene within the defective prophage element e14 shows high fitness in presence of T6, CEV1, CEV2, 186, λ cI857 and LZ4 phages, but only on plate-based assays (Fig. 4, Supplementary Figs. S3, S4). Overexpression of *lit* is known to interfere with the T4 phage growth [154], though we did not observe high-fitness scores in the presence of T4 phage.

In summary, we identified 129 multicopy suppressors of phage infection that encode diverse functions, and our results indicate that enhanced host fitness (phage resistance) can occur via diverse cellular mechanisms. In the following section, we describe follow-up experiments with 13 of these top-scoring hits from the Dub-seq screen to validate their role in phage resistance.

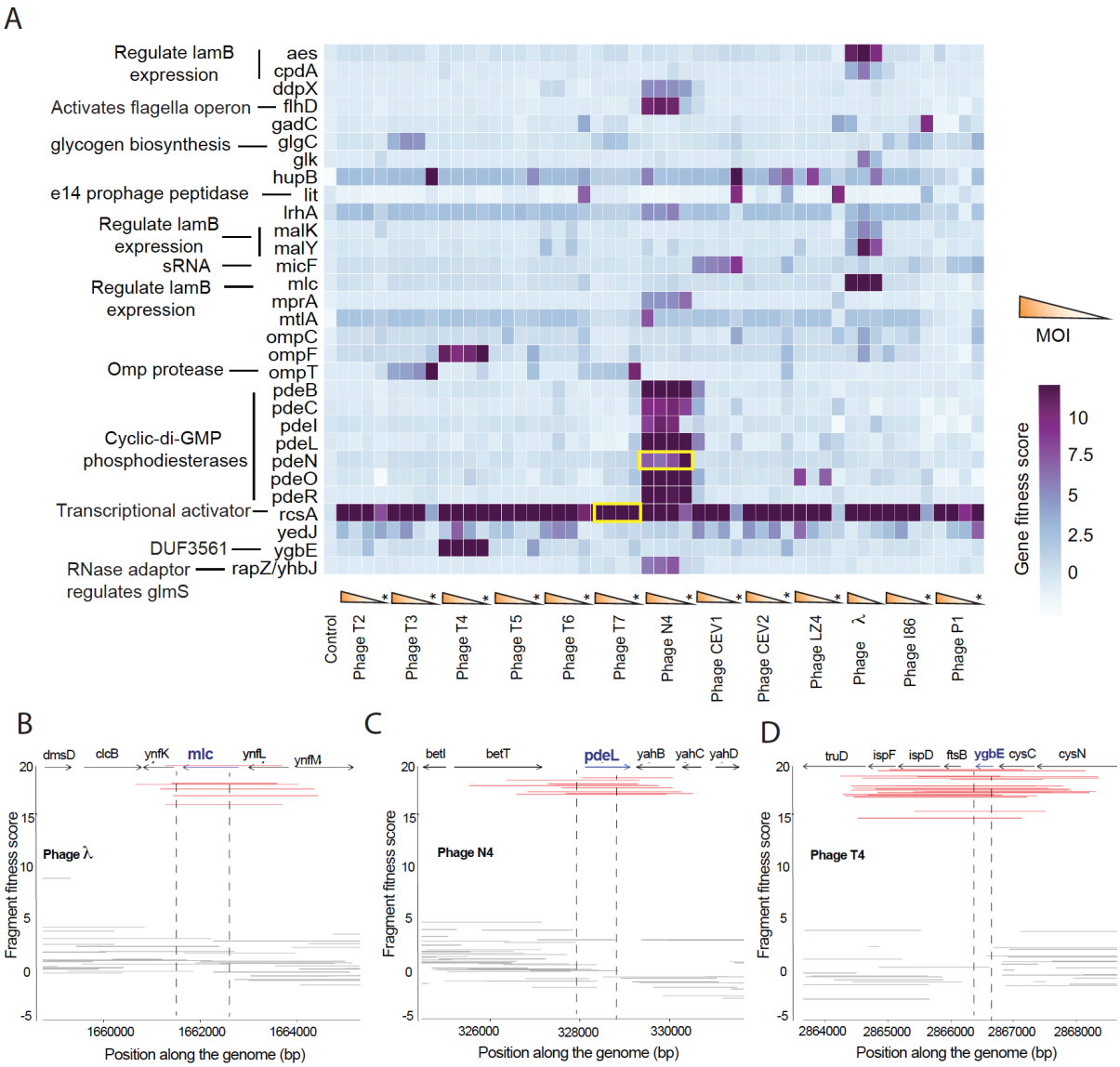


Figure 4 Dub-seq screening data for 13 dsDNA phages. (A) Heatmap of Dub-seq data for 13 dsDNA phages at different MOI and 30 genes with large fitness benefits. Only genes with high-confidence effects and gene fitness score of >=4 in at least one phage assay are shown. Yellow boxes highlight genes that are known to show resistance when overexpressed. The pooled fitness assays performed on solid plate agar are marked with stars. (B to D) Dub-seq viewer plots for high-scoring *mlc* (B), *pdeL* (C) and *ygbE* (D) containing fragments in the presence of λ, N4 and T4 phages respectively. Red lines represent fragments covering highlighted genes completely (start to stop codon), while grey colored fragments either cover the highlighted gene partially or do not cover the highlighted gene completely. Additional Dub-seq viewer plots are provided in Supplementary Figs S2-S4.

Experimental validation of LOF and GOF screen hits

To validate the phage resistance phenotypes observed in our LOF screens, we sourced 6 individual deletion strains *lpcA*, *galU*, *ompF*, *dgcJ*, *ygaH* and *dsrB* from *E. coli* K-12 BW25113 KEIO library [100] and disrupted *igaA* in the *E. coli* K-12 BW25113 strain (Supplementary Fig S1). We then determined the efficiency of plating (EOP) for eight phages (Fig. 5). In addition, we performed gene complementation experiments using *E. coli* K-12 ASKA plasmids [155] to

check if the plating defect can be restored (Fig. 5, Supplementary Fig. S5). To validate the hits identified in our Dub-seq GOF screens, we moved 12 individual plasmids expressing *rcsA*, *ygbE*, *yedJ*, *flhD*, *mtlA*, *pdeB*, *pdeC*, *pdeI*, *pdeL*, *pdeN*, *pdeO*, and *pdeR* into *E. coli* K-12 and tested the EOP of 11 phages (total 36 phage x gene combinations) (Fig. 5, Methods). We present these results below.

CEV1 phage requires both OmpF and LPS for K-12 infection

ompF, lpcA/gmhA and additional genes involved in LPS biosynthesis and transport showed the highest fitness scores in the presence of CEV1 phage in the RB-TnSeq and CRISPRi data (Figs. 2A and 3A). In agreement with our LOF screen data, we observed severely reduced EOP of CEV1 on *ompF* and *lpcA* deletion strains compared to the control BW25113 strain (Fig. 5A, Supplementary Fig. S5). These plating efficiency defects could be reverted when the respective deleted genes were expressed in *trans* (Supplementary Fig. S5). These results indicate that CEV1 infection proceeds by recognizing both OmpF and LPS core, and loss of either *ompF* or LPS disruption leads to a resistance phenotype.

Overexpression of colanic acid biosynthesis pathway reduces sensitivity to diverse phages

One of the top scoring genes in our LOF screens with different phages was *igaA* (*yrfF*) which codes for a negative regulator of the Rcs phosphorelay pathway. Likewise, the transcriptional activator RcsA that works downstream from IgaA in the Rcs signaling pathway showed the highest fitness scores in GOF Dub-seq screen against all phages tested in this work. Furthermore, *dsrB* (downstream (from rcsA) region B) showed higher fitness in CRISPRi screens in the presence of N4 phage (Fig. 3A). The Rcs signaling pathway is involved in regulating a large number of genes involved in diverse functions including synthesis of colanic acid and biofilm formation in response to perturbations in the cell envelope [101]. Downregulation of *igaA* has been shown to resist infection by T4, λ and 186 phages [51], while overproduction of transcriptional activator RcsA has been shown to display resistance to T7 phage infection [50]. In agreement with our Dub-seq screen data, we observed the EOP of T2, T3, T4, T5, T6, 186, λ, CEV1, CEV2, LZ4, and N4 phages is compromised when *rcsA* is overexpressed (Figs. 5B). Similarly, a *dsrB* deletion strain showed severe N4 phage plating defects, confirming the high fitness scores in our CRISPRi screen (Fig. 5A).

Despite *igaA's* reported essentiality [156,157], our RB-TnSeq library contained 9 disruptions in *igaA*'s cytoplasmic domain which also overlapped with sgRNA target sites in our CRISPRi screen (Figs. 3A, 5A and Supplementary Fig. S1). To validate that this domain is indeed dispensable for strain viability and also important in phage resistance, we successfully reconstructed the *igaA* insertion mutant (Methods). This mutant strain displayed a mucoidy phenotype indicative of increased activation of the Rcs pathway [101,158]. We observed defective plaque morphologies with T3, T5, T6, T7, P2, 186, CEV1 and N4 phages on the *igaA* insertion mutant, indicating inefficient infection and the plating defect could be reversed by supplying the wild-type *igaA* gene *in trans* (Fig. 5A, Supplementary Fig. S5). To gain further insight into the phage resistance mechanism, we performed RNA-seq analysis on the *igaA* disruption mutant (Methods). We found that multiple components of the Rcs pathway (including *rcsA* itself) were upregulated, with 24 genes from the capsular biosynthesis-related operons *wca* and *yjb* significantly upregulated (log2FC > 2, adjusted p-value<0.001) (Figs. 6A, 6B, Supplementary Table S6 and S7). These results indicate that the *igaA* disruption mutant uncovered in this work activates colanic acid biosynthesis, leads to a mucoidy phenotype, and may be interfering with phage infection by blocking phage receptor accessibility. The IgaA

residues 18-164 we mutate in this work overlap with the N-terminal cytoplasmic domain of IgaA that inhibits Rcs signaling in the absence of stress [159]. Recent studies have highlighted that the activation of the Rcs pathway and capsular biosynthesis are essential to survive in diverse environmental contexts, membrane damage and stress-inducing conditions, including those by antibiotic treatment [101,160]. A number of earlier studies have also highlighted the generation of the mucoid phenotype that probably provides fitness advantage by blocking phage adsorption [16,161–164]. Our results suggest this might be a generalized mechanism that provides cross-resistance to diverse phages.

Overexpression of YgbE confers resistance to phage T4 and down-regulates OmpC

YgbE is a DUF3561-containing inner membrane protein with no known function [165]. In Dub-seq screens, *ygbE* shows highest fitness scores in the presence of T4 phage and our EOP data confirm that T4 phage shows strong plating defects when *ygbE* is overexpressed (Fig. 5B, EOP of 7E-6). To gain insight into the mechanism of phage resistance, we performed RNA-seq analysis on a *ygbE* overexpression strain. Differential expression analysis of this strain revealed strong downregulation of *ompC* (log2FC = -5.7, adjusted p-value<<0.001), the primary T4 phage receptor (Figs. 6C, 6D). In addition, we also noticed a strong downregulation of 26 genes (log2FC = -4.5 adjusted p-value<<0.001) related to flagella structure including RNA polymerase sigma 28 (sigma F) factor, and strong upregulation (log2FC = 4.5 adjusted p-value<<0.001) of *arnBCADT* operon involved lipid-A modification (Figs. 6C, 6D, Supplementary Table S6 and S7). The downregulation of *ompC* and upregulation of LPS modification genes is in agreement with the observed phage resistance phenotype of *ygbE* (Figs. 4A, 4D, 5B), but the mechanism of *ompC* downregulation in the *ygbE* overexpression strain remains to be determined.

Cyclic di-GMP is required for infection by phage N4

The top 8 hits in our LOF and GOF screens for N4 phage resistance included enzymes that catalyze synthesis and degradation of cyclic-di-GMP, a key secondary signaling molecule involved in diverse cellular functions. *dgcJ*, one of the top scorers in both RB-TnSeq and CRISPRi LOF screens encodes diguanylate cyclase-J, which is involved in the biosynthesis of cyclic-di-GMP, whereas the seven c-di-GMP-specific phosphodiesterases (PDEs) that are involved in degradation of c-di-GMP showed the highest fitness scores in the Dub-seq screen. Though the signaling network of cyclic-di-GMP is complex, deletion of diguanylate cyclases (DGCs) or overexpression of PDEs is known to reduce c-di-GMP levels, inhibit curli and biofilm formation, while increasing cellular motility [144–146,166]. The high fitness scores for *dgcJ* in RB-TnSeq and CRISPRi screens is intriguing considering it is one of the 12 DGCs encoded on *E. coli* K-12 genome [167,168], and none of the other DGCs show phenotypes in our screens. Similarly, *E. coli* K-12 genome codes for 13 PDEs in total [167,168], and we find 6 of these PDEs show a phage resistance phenotype when overexpressed. Our EOP estimations with N4 phage showed severe plating defect (EOP <8E-8) on *pdeO (dosP)*, *pdeR (gmr)*, *pdeN (rtn)*, *pdeL* (*yahA)*, *pdeB (ylaB)*, *pdeI (yliE)* overexpressing strains, while minor plating defect on *dgcJ* mutant (EOP of 0.8) (Figs. 5A, 5B). The plating defect of N4 phage on *dgcJ* could be reverted back when the *dgcJ* was provided *in trans* (Supplementary Fig. S5).

To gain insight into how the overexpression of specific genes might affect phage infection pathways, we performed differential RNA-seq experiments (standard growth conditions, in the absence of phage) on five c-di-GMP phosphodiesterases (*pdeL*, *pdeB*, *pdeC*, *pdeN* and *pdeO*) overexpressing strains, each of which shows resistance to N4 phage (Methods). RNA-seq experiments on the *pdeL* overexpression strain revealed large changes in the *E. coli*

transcriptome relative to the wild-type BW25113 strain, with 103 genes significantly upregulated (log2FC > 2, q < 0.001) and 109 genes significantly downregulated (log2FC < -2, q < 0.001) (Fig. 6E, 6F, Supplementary Table S6 and S7). N4 phage receptor genes *nfrA* and *nfrB* were not differentially expressed and therefore phage resistance phenotype appears to be independent of N4 phage receptor transcript levels in a *pdeL* overexpressing strain. The *ddpX* transcript that codes for D-Ala-D-Ala dipeptidase involved in peptidoglycan biosynthesis exhibits upregulation (log2FC = 2.11, q=2e-17) in the *pdeL* overexpression strain. One of the high scoring Dub-seq screen hits in N4 phage assays is *ddpX* (Fig. 4A), though its role in imparting N4 phage resistance is not clear. We note that the fitness effect of Dub-seq fragments encoding *ddpX* might also be because of its upstream *pdeO* (*dosP*) overexpression (Supplementary Fig. S3). Similar to *pdeL* overexpression, overexpression of the other four PDEs did not show substantially different expression of *nfrA* and *nfrB*, suggesting that N4 resistance phenotype in these instances does not appear to be via transcriptional regulation of the N4 phage receptor genes (Supplementary Table S6 and S7).

Validation of other top scoring genes displaying phage resistance phenotype

ygaH is one of the 4 top scoring RB-TnSeq candidates in the presence of N4 phage (Fig. 2A). YgaH is predicted to be a L-valine exporter subunit [165,169] and had not been previously associated with N4 phage infection. In agreement with its strong fitness scores in RB-TnSeq data, our EOP estimations on *ygaH* deletion strain confirmed severe plating defects on N4 phage (Fig. 5A). Although the role of YgaH in N4 phage infection pathway is unclear, Dub-seq fragments encoding *mprA* (a negative regulator of multidrug efflux pump genes), a gene downstream of *ygaH,* also shows an N4 phage resistance phenotype (Fig. 4, Supplementary Fig. S3), further demonstrating the importance of this region in N4 phage infection.

Among other top hits in the GOF Dub-seq screen, we observed defective plating of N4 phage on *lrhA* and *rapZ* (*yhbJ*) overexpression strains (Fig. 5B, Supplementary Fig. S3). *lrhA* codes for a transcription regulator of genes involved in the synthesis of type 1 fimbriae, motility and chemotaxis [170], and has not been previously linked to N4 phage infection. *rapZ* codes for an RNase adaptor protein that negatively regulates the expression level of glucosamine-6-phosphate synthase (GlmS) [170,171]. GlmS catalyzes the first rate-limiting step in the amino sugar pathway supplying precursors for assembly of the cell wall and the outer membrane. Though the role of the essential gene *glmS* in N4 phage infection is unclear, the N4 phage resistance phenotype shown by *rapZ* multicopy expression (which downregulates *glmS* expression) agrees with our CRISPRi screen data wherein knockdown of both *glmS* and *glmU* within *glmUS* operon show strong fitness in the presence of N4 phage (Fig. 3C). These results illustrate the power of using these combination technologies for studying phage infection.

Finally, candidates that showed high fitness scores in our Dub-seq screen but failed to demonstrate strong phage plating defects include strains overexpressing *flhD* or *mtlA* in the presence of N4 phages, and *yedJ* in the presence of multiple phages (Fig. 4A, Supplementary Fig. S6). We speculate other entities such as sRNA coding regions or transcription factor binding sites on Dub-seq fragments encoding these gene loci may be essential for phage resistance phenotype (Supplementary Figs. S6 and S7).

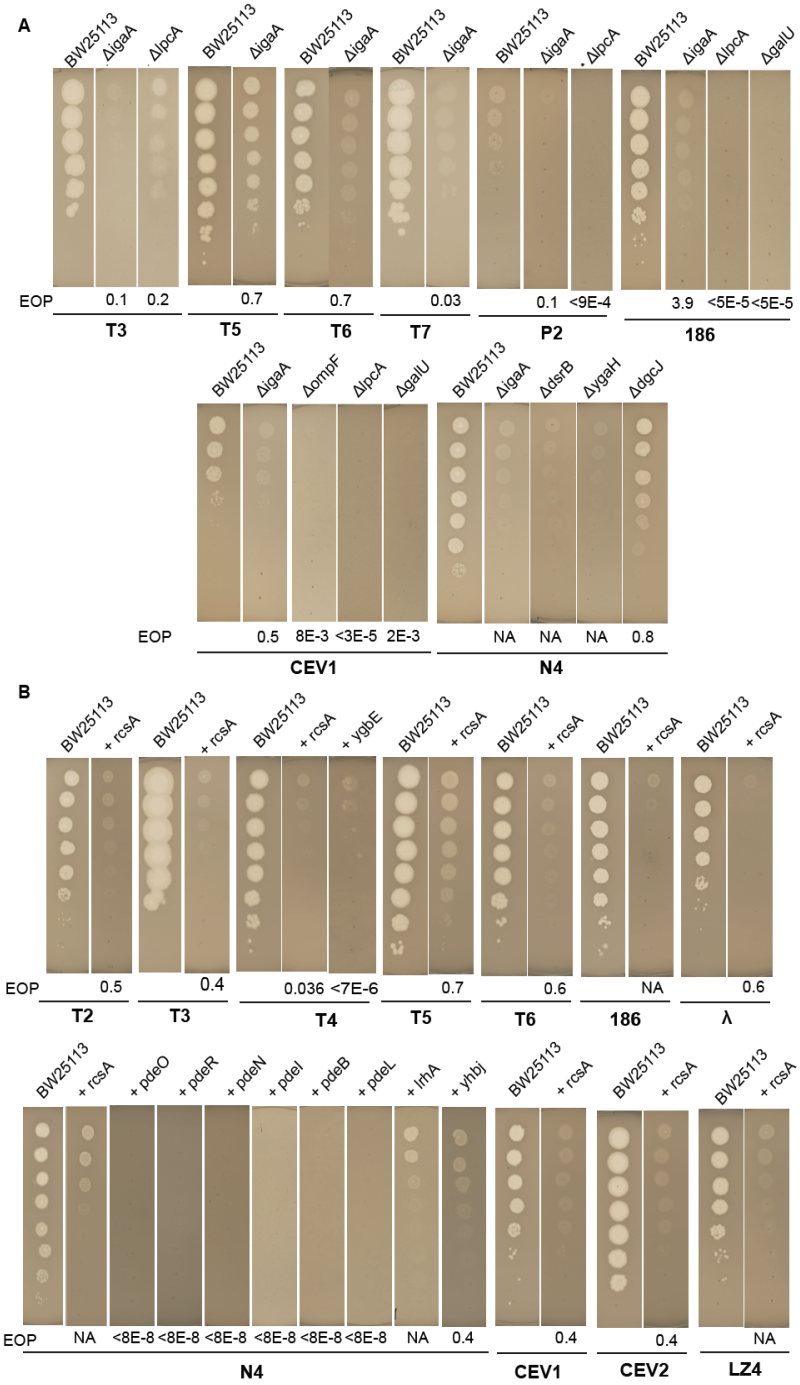


Figure 5. Experimental validations of top scoring gene hits in LOF and GOF screens. (A)Efficiency of plating experiments for LOF screen hits using KEIO [100] library strains. Gene complementation data is presented in Supplementary Fig. 2. (B) Efficiency of plating experiments for GOF screen hits with ASKA plasmid library [155] expressing genes (shown as +gene names) in the presence of different phages. We used no IPTG or 0.1 mM IPTG for inducing expression of genes from ASKA plasmid. We used the BW25113 strain with an empty vector for estimating EOP. The plaque morphology or EOP of T3, T7, P2 and 186 phages on *lpcA* deletion strain indicated inefficient infection. The plating defect was restored to normal when mutants were complemented with a plasmid expressing the respective deleted genes indicating LPS core as the receptor for these phages (Supplementary Fig. S5).

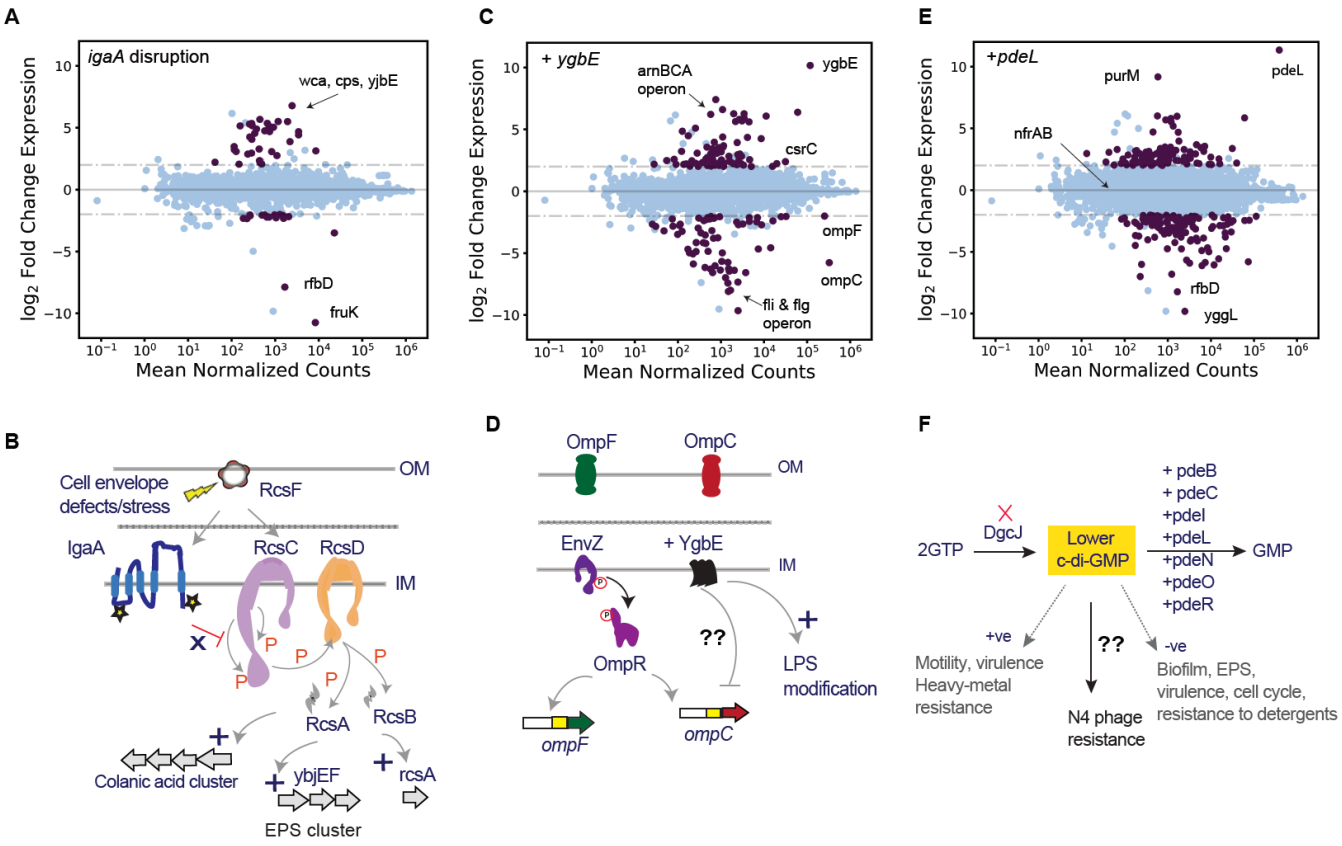


Figure 6. RNA-seq analysis to gain insights into phage resistance mechanisms (A) Upregulation of exopolysaccharide biosynthesis genes observed in *igaA* disruption mutant relative to wt (N=3). (B) Schematic of the Rcs phosphorelay pathway, with mutants *igaA* identified in this work that show a mucoidy phenotype and broad resistance to different phages. (C) Downregulation of *ompC* transcript and upregulation of *arnBCA* operon observed during *ygbE* overexpression relative to wt (N=3). (D) Schematic of *ompF* and *ompC* regulation in the presence of *ygbE* overexpression that displays resistance to phage T4. (E) RNA-seq data of *pdeL* overexpression showed no downregulation of N4 receptor genes (*nfrA* and *nfrB*) and no upregulation of genes involved in EPS or biofilm. (F) Schematic of cyclic-di-GMP pathway with *dgcJ* deletion or overexpression of one of 7 PDE encoding genes (representing decreased c-di-GMP levels) show a high fitness score in presence of N4 phage. In A, C, and E plots, purple filled data points are adjusted p-value < 0.001 and abs(log2FC) > 2. Blue filled is nonsignificant data points. The dashed lines are effect size thresholds of greater than 4 fold.

Extending genome-wide screens to E. coli BL21 strain

To compare the phage resistance phenotypes we observed in *E. coli* K-12 to a closely-related host, we chose *E. coli* BL21 strain as our alternate model system. Historically, both *E. coli* K-12 and B (BL21 ancestor) strains have been used in disparate phage studies and have provided foundational knowledge on phage physiology and growth [10,90,91,93], though phage studies on *E. coli* BL21 are limited relative to those in K-12. Both K-12 and BL21 lack functional CRISPR machinery that functions as one of the common anti-phage systems, and both strains are unique in their genomic content, physiology, and growth characteristics [172–174], thereby providing a valuable reference for comparing closely related host responses to the same phage selection pressures. The genomes of BL21 and K-12 strains have 99% base pair identity over 92% of their genomes, interrupted by deletions or disruptions (by mobile

elements) [172–174]. For example, in comparison to K-12, BL21 has a disruption or deletions in the colanic acid pathway, biofilm formation, flagella formation (a 21 gene cluster including flagella-specific *fliA* sigma factor), genes involved in Rcs signaling pathway (*rcsA*, *rcsB*, *rcsC*), the LPS core gene cluster (BL21 forms truncated LPS core with only 2 hexose units compared to normal five hexose units in K-12, Fig. 2B), and is deficient in Lon and OmpT proteases important components of the protein homeostasis network [172–174]. In addition, BL21 also lacks *ompC* and *nfrAB* genes that code for T4 phage and N4 phage receptors in *E. coli* K-12 respectively [172,175,176].

To investigate BL21 genes essential for phage growth, we constructed an RB-TnSeq library made up of ~97,000 mutants (Methods). For fitness assays, we used the same set of phages that were assayed with K-12 library except phages N4 and 186. Both N4 and 186 phages do not infect BL21 due to lack of N4 phage receptors (NfrA and NfrB) and truncated LPS that probably limits 186 binding. We performed 53 pooled fitness experiments using the BL21 RB-TnSeq library in both liquid and solid plate assay format at varying MOI and 9 no-phage control experiments. In total, we identified 115 high scoring hits, made up of 50 unique phage x gene combinations, and representing 32 unique genes (Supplementary Table S8). All 12 phages have at least one high confidence hit. Largely, the BL21 LOF data is in agreement with our K-12 results for T3, T5, T6, T7, CEV2, LZ4, λ, P1 and P2 phages, especially the high-scoring genes that either code the phage receptor or its regulators (Fig. 7).

The key differences between BL21 and K-12 LOF fitness data are the host factors important in T2, T4 and CEV1 phage infection. For example, *ompC,* and its regulator *EnvZ/OmpR* two-component system and genes involved LPS core biosynthesis, showed high fitness scores in the presence of T4 phage in K-12 LOF screens, while only genes involved in the LPS core biosynthesis (*waaG*, *waaF* and *waaQ*) were important in BL21 screens (Fig. 7). Lack of *ompC* in the BL21 strain probably alleviates the need for the EnvZ/OmpR two-component system for T4 infection. The absolute requirement for LPS R-core structure for T4 phage growth on BL21 is in agreement with the earlier reports [90,93,176–178]. Similarly, CEV1 phage, which showed a strict requirement for OmpF and full length LPS in K-12 screens (Fig. 2), seems to require only OmpF in the BL21 infection cycle (Fig 7). This suggests that CEV1 phages can tolerate truncated LPS of BL21 but not that of K-12 (Fig. 2). Because of the OmpF requirement, CEV1 growth on BL21 also showed strict dependence on the EnvZ/OmpR two-component system, a key regulator of *ompF* expression. Finally, in distinction to K-12 data, our BL21 RB-TnSeq data indicates that T2 phage probably binds preferentially to truncated LPS R-core in BL21 (*waaF* and *waaG* showed high fitness in our screen). Furthermore, this effect seems to be independent of FadL. We confirmed this observation by measuring the EOP on a BL21 *fadL* deletion strain (data not shown). The absence of a FadL requirement for T2 growth on BL21 is intriguing considering its 100% nucleotide identity with K-12 *fadL*. It is possible that either the conformational integrity of FadL is compromised in the absence of full-length LPS or that T2 may be recognizing more than one outer membrane protein [87,179] (Supplementary Notes). These results are consistent with earlier observations on the difficulty in isolating T2 resistant mutants in *E. coli* B cells [15].

Finally, to investigate whether increased gene copy number of host factors interferes with phage growth in BL21, we constructed a BL21 Dub-seq library made up of 65,788 unique fragments (Methods, Supplementary Fig. S8). We then screened the BL21 Dub-seq library in a variant of BL21 as the host (Methods). From 24 pooled fitness experiments in planktonic cultures and on solid media in the presence of 12 phages we identified 39 high-scoring candidates (fitness score >=4, FDR of 0.74, Methods). Other than a few top-scoring

candidates in the presence of λ phage, the BL21 dataset was considerably different from K-12 dataset (Supplementary Table S9).

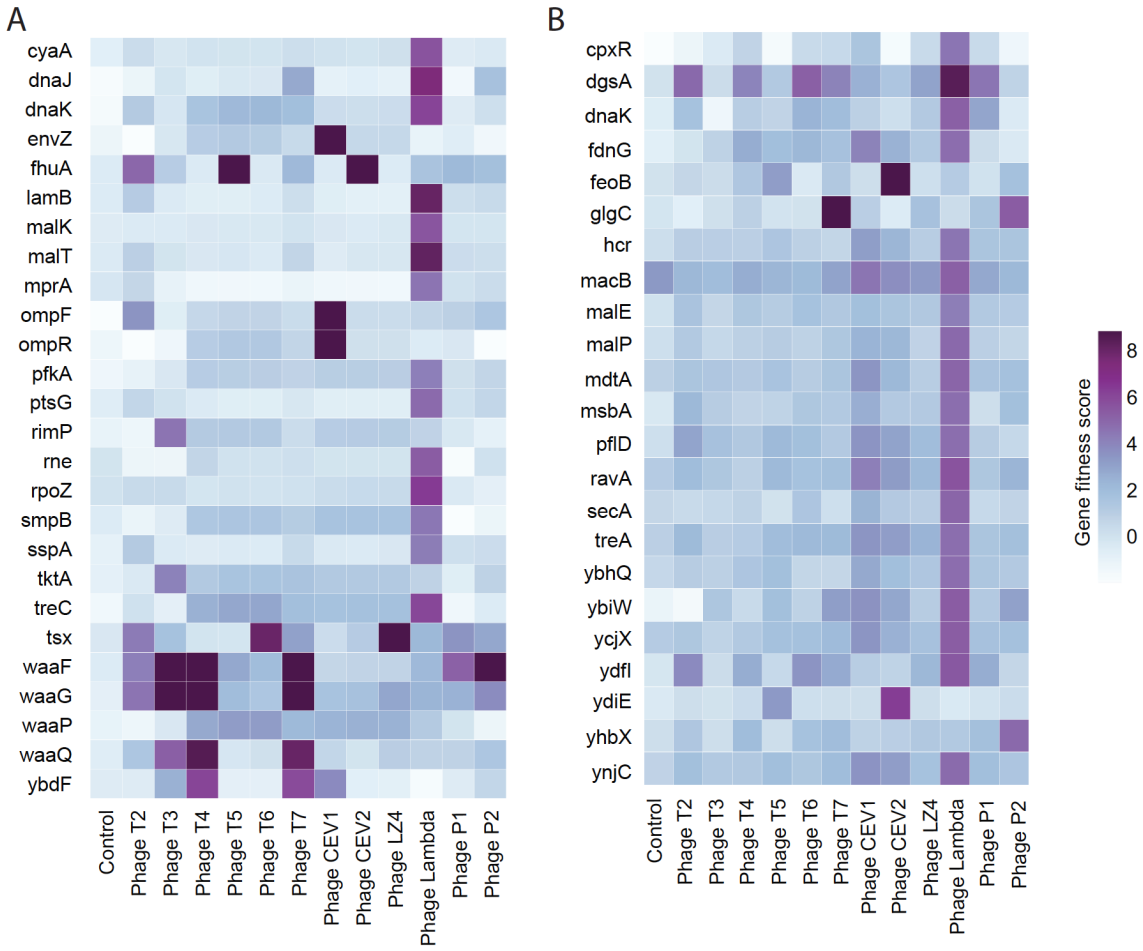


Fig. 7 Genome-wide screens in *E. coli* BL21 strain (A) Heatmap of BL21 LOF RB-TnSeq data for 12 dsDNA phages at a single MOI, and selected genes with high-confidence fitness benefits are shown. Additional data is provided in the Supplementary Table S8. (B) Heatmap of GOF BL21 Dub-seq data for 12 dsDNA phages with high-confidence fitness fitness benefit. Fitness scores of >=4 in at least one phage assay are shown. These assays were performed in planktonic culture. Additional data is provided in the Supplementary Table S9.

Some of the tops scoring hits in BL21 Dub-seq screen showed broad resistance to many different phages while some were phage specific. For example, BL21 Dub-seq fragments with *mlc* (*dgsA*) gave higher fitness in the presence of T2, T4, T6, T7, λ, and P1 phages. This resistance phenotype of Mlc in the presence of λ phage is consistent with its known regulatory role. Mlc, a global regulator of carbohydrate metabolism, is known to negatively regulate the maltose regulon and mannose permease system [180] (via regulating the *lamB* expression activator MalT), both of which are known to play a crucial role in phage λ DNA penetration and infection [181–184]. Another top scoring candidate *glgC* showed higher fitness in the presence of T7 and P2 phages. Overexpression of *glgC* (that encodes Glucose-1-phosphate adenylyltransferase) is known to increase glycogen accumulation [185,186] and titrate out the global carbon storage regulator CsrA [187]. We speculate that the interaction of GlgC and

CsrA probably impacts biofilm formation [124,188–190] and leads to alternations in the LPS profile [191,192] leading to phage resistance phenotype.

Among the candidates that show phage specific resistance phenotype, we observed that overexpression of *argG* showed fitness score of 3.7 in the presence of T4 phages (Supplementary Table S9). Argininosuccinate synthetase enzyme (encoded by *argG*) catalyzes the penultimate step of arginine biosynthesis and has not been associated with phage resistance phenotype before. However, early studies have indicated the inactivating effect of arginine on T4 phages [193]. Finally, Dub-seq fragments encoding ferrous iron uptake system (FeoB) and putative heme trafficking protein (YdiE) yield strong fitness in the presence T5 phage and T5-like CEV2 phage (Fig. 7). It is known that increased uptake of ferrous iron increases Fur-ferrous iron occupancy and Fur-mediated repression of T5 phage receptor *fhuA* [194–198]. Most of these top-scoring candidates were missing in the K-12 dataset, probably because of strong selection for *rcsA* overexpressing strains in all of our K-12 Dub-seq experiments. This highlights the importance of studying how even closely related hosts can have nuanced interactions with the same bacteriophage.

Summary and Conclusions:

We applied high-throughput LOF and GOF screening methods to two different *E. coli* strains to map the landscape of genetic determinants important in host-phage interactions. We demonstrate how these methods can rapidly identify phage receptors and both novel and previously described host factors involved in resistance across a wide panel of dsDNA phage types. By using LOF RB-TnSeq and CRISPRi methods, we extensively map both non-essential and essential host genes along with non-genic regions such as promoter and transcription factor binding sites implicated in phage infection and resistance. The Dub-seq methodology uncovers dozens of multicopy suppressors that encode diverse functions and point to myriad of ways how host gene dosage can influence phage resistance. This global survey of host factors that play an important role in phage growth across two widely studied *E. coli* strains provides a detailed view of cross-resistance patterns for diverse phages and will be a rich dataset for deeper biological insights and machine learning approaches.

The strong positive fitness scores we observe due to the selection pressure during these pooled fitness experiments is both a strength and limitation of our methods. The strength of these screens is that we can rapidly identify host factors crucial in phage infectivity cycle because they display a very strong fitness score when disrupted (for example, a phage receptor and its regulators in LOF screen). This strong positive selection, however, limits the detection of intermediate phage resistance routes whose disruption may lead to strains with relatively poor or neutral fitness and will be swept from the population. Our data from both RB-TnSeq and Dub-seq screens shows consistency across a range of MOI, and also suggests that different assay formats (solid and liquid) allow increased discovery of diverse phage resistance mechanisms. These assay platforms could be expanded to phage-banks made up of hundreds of phages at just one MOI, and may be sufficient to rapidly discover the phage receptors in the target host. In addition, extending the screening methods to closely related but different strains would be highly valuable. For example, we used two laboratory *E. coli* strains that have rough LPS architecture where core oligosaccharide is the terminal part of the LPS. However, it is known that the genetic and structural diversity of LPS and the repeat structure O-polysaccharide attached to LPS (to form smooth LPS) is very large in pathogenic and environmental isolates of Enterobacteriaceae, and may impact phage infectivity [47,57,199–

202]. The genetic screens presented in this work may aid in filling the knowledge gap on phage interaction with different O-antigens and its impact on phage infectivity and resistance.

Our results also highlight that phage infectivity depends on the host cellular physiology and, in particular, membrane characteristics of the host imparted by LPS and outer membrane protein biogenesis pathways [108,109,203]. For instance, our results indicate that the disruptions in LPS (for example, deletions or downregulation of *waa* genes Fig. 2A, 3A, 7A), LPS transport pathways (for example, downregulation of *lptABC*, Fig. 3B) and signal resembling membrane stress (for example, disruption in *igaA* in Fig 2A, 3B; overexpression of *rcsA*, *ompF, micF* in Fig 4A) can influence phage infectivity (additional discussion in Supplementary Notes).

The screening technologies presented in this work are scalable to study phage resistance in diverse conditions that simulate the natural environment and may provide valuable insights on host fitness and host-phage co-evolutionary dynamics under more ecologically relevant conditions. For example, recent studies highlighted the evidence of subdued phage resistance in the natural environment, probably because of the fitness cost associated with resistance mechanisms [89,204–209]. In addition, these methods have the capability to identify fitness costs associated with broadly seen phage resistance phenotypes in a competitive natural environment, and thus improve our understanding of microbial ecology in general [13,114,206,207,210–212]. Such systems-level insights will be valuable both in uncovering new mechanisms in host-phage interaction and perhaps in developing different design strategies for targeted microbial community interventions, engineering highly virulent or extended host-range phages and rationally formulated phage-cocktails for therapeutic applications [89,204,208,213–226]. Alternatively, identifying phage resistance determinants may also enable engineering of bacterial strains with phage defense systems crucial in a number of bioprocesses such as in the dairy industry [227,228], biocontainment strategies for bioproduction industry [229,230] or to facilitate bacterial vaccine discovery and development [231–233].

There is a clear applied interest in utilizing combinations/cocktails of phages to regulate or eliminate bacterial populations due to the reduced likelihood of evolved multi-phage resistance [89,214,218]. However, designing such cocktails relies on a better understanding of cross-resistance among phages [28,60,234]. In particular, identification of phages that differ in their receptor use or against which cross-resistance is unlikely to evolve would allow for better design of such therapies [23,36]. Moreover, identifying phages that select for resistance that have interrelated phenotypic consequences with, for example, antibiotic sensitivity is a recent advancement in the field that could directly benefit from these screening approaches [52,89]. By combining fitness datasets for phages and antibiotics or phage-antibiotic combination therapies [235–237], such screens could provide an avenue for performing high-throughput search for genetic trade-offs or ‘evolutionary traps’ [52,89] that could provide a much-needed solution to overcome the antibiotic resistance pandemic.

Author contributions

V.K.M., A.M.D. and A.P.A. conceived the project.

V.K.M. led the experimental work and supervised the project.

V.K.M., B.A.A., H.S.R., D.P., and C.Z. performed experiments and collected data.

V.K.M., B.A.A., H.S.R., D.P., P.S.N., M.N.P. and A.M.D. processed and analyzed the data.

R.C., provided critical reagents and advice.

V.K.M., B.A.A., H.S.R., D.P., P.S.N., M.N.P., B.K., A.M.D. and A.P.A. wrote the paper.

Acknowledgements

Authors would like to thank Elizabeth B. Kutter, Studier F William, Lucia B. Rothman-Denes, Ian J. Molineux, Jason J. Gill and Paul Turner for sharing reagents and helpful discussions.

The initial concepts for this project were funded by ENIGMA, a Scientific Focus Area Program at Lawrence Berkeley National Laboratory, supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under contract DE-AC02-05CH11231. This project was funded by the Microbiology Program of the Innovative Genomics Institute, Berkeley.

Sequencing was performed at: Vincent J. Coates Genomics Sequencing Laboratory (University of California at Berkeley), supported by NIH S10 Instrumentation Grants S10RR029668, S10RR027303, and OD018174;

Competing interest

VKM, AMD, and APA are holders of a patent on the genetic screening technology.

Supplementary Figures

S1. Map of *igaA* mutants
S2. Dub-seq viewer plots for high-scoring genomic fragments in the presence of phages CEV1 and λ
S3. Dub-seq viewer plots for high-scoring genomic fragments in the presence of phages T2, T3, T4, T6 and N4
S4. Dub-seq viewer plots for high-scoring genomic fragments in the presence of phages N4, 186 and LZ4.
S5. Estimation of efficiency of plating (EOP) with KEIO deletion strains and genetic complementations
S6. Examples of candidates that showed high fitness scores in our Dub-seq screen but failed in EOP validation experiments
S7. Dub-seq and RB-TnSeq data uncovers *flhD* upstream loci important in N4 phage infection
S8. Description of BL21 Dub-seq library

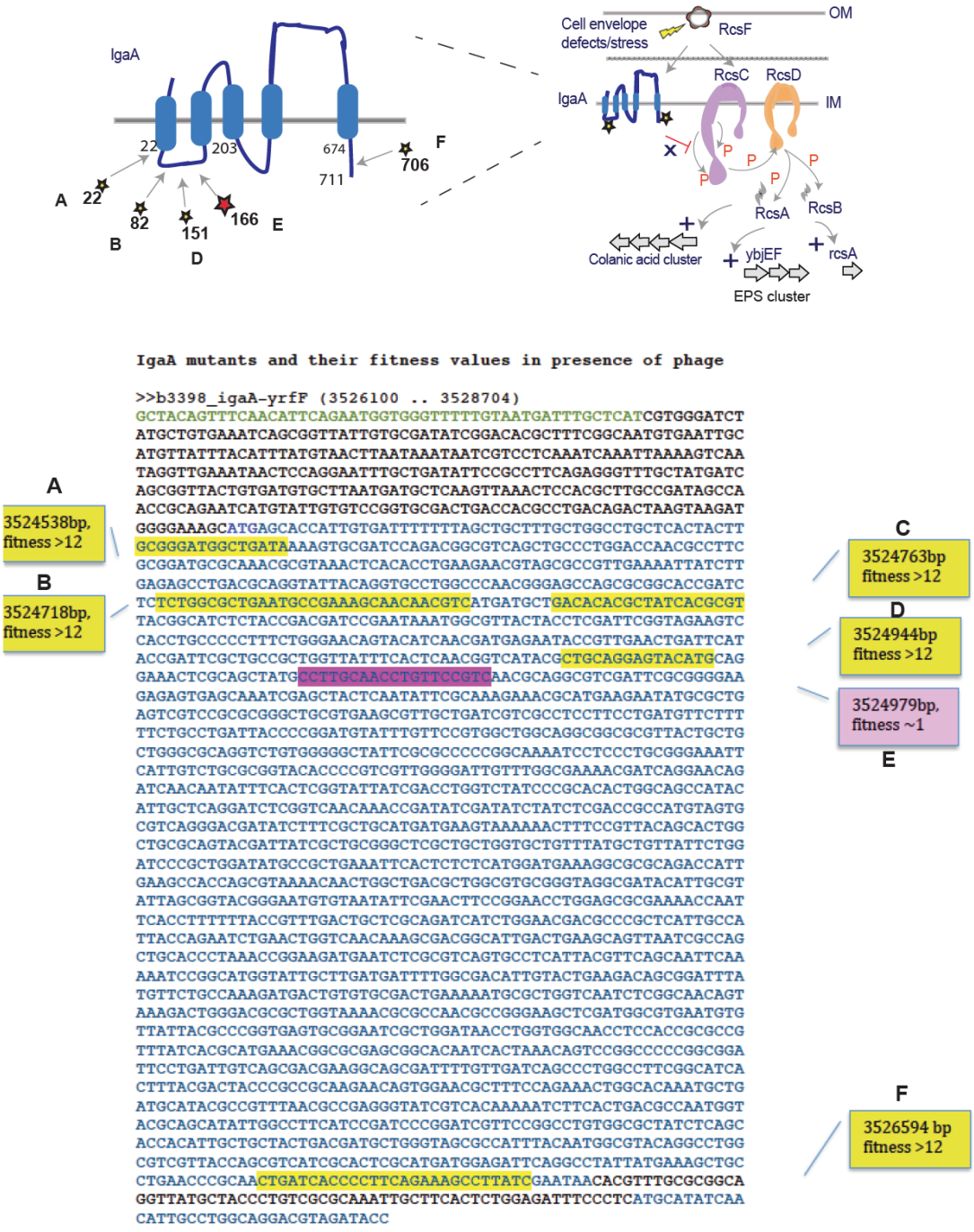
Supplementary Tables

Supplementary Tables are deposited here: https://doi.org/10.6084/m9.figshare.11859216.v1

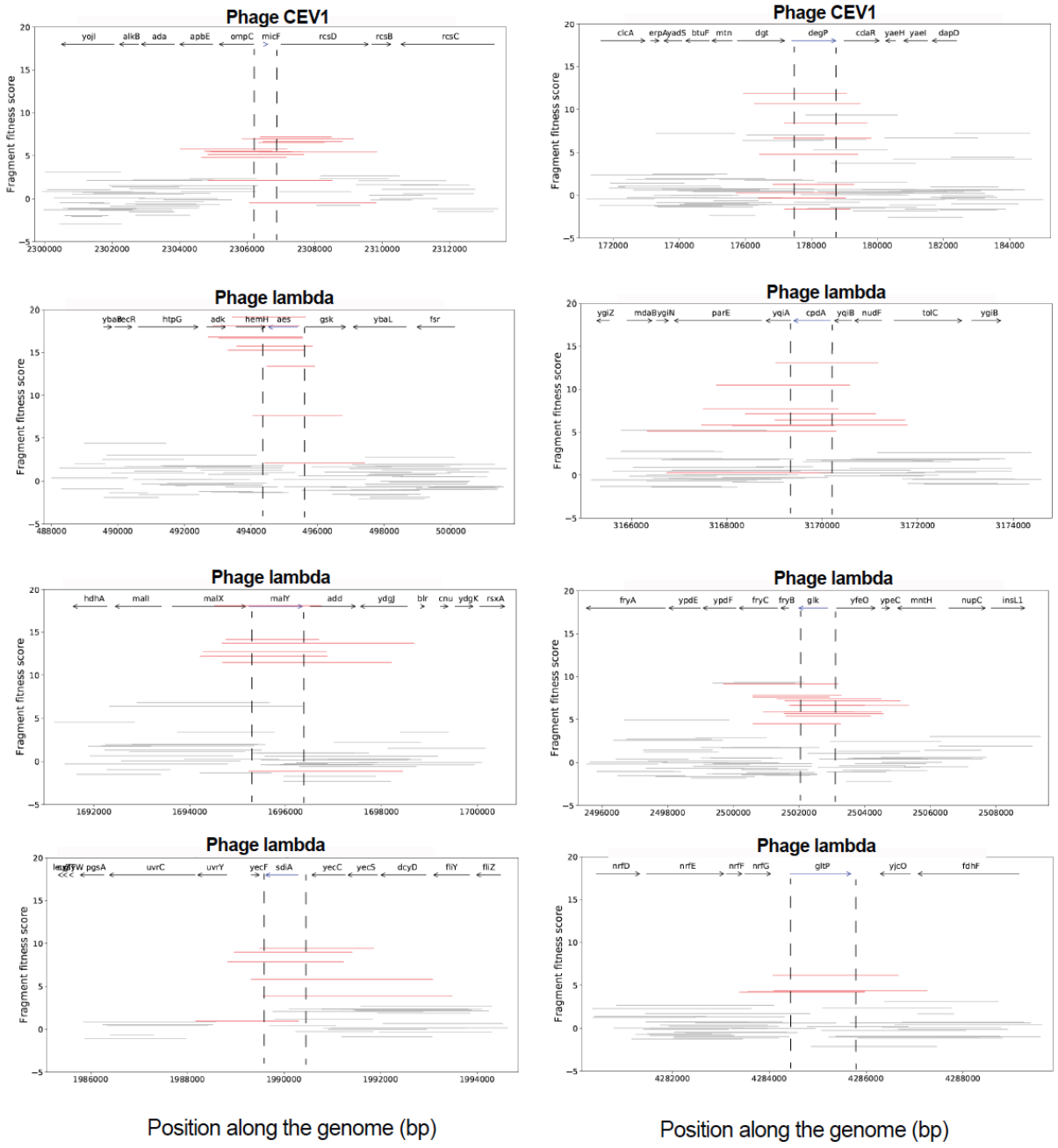
1. Supplementary Table S1: RB-TnSeq K-12 dataset
2. Supplementary Table S2: Summary of known and new high fitness score hits per phage per screen at any MOI.
3. Supplementary Table S3: Literature table: Mapping RB-TnSeq, CRISPRi and Dub-seq hits to reported phage resistance data
4. Supplementary Table S4: CRISPRi K-12 dataset
5. Supplementary Table S5: Dub-seq K-12 dataset
6. Supplementary Table S6: RNA-seq dataset
7. Supplementary Table S7: RNA-seq DESeq2_Summary
8. Supplementary Table S8: RB-TnSeq BL21 dataset
9. Supplementary Table S9: Dub-seq BL21 dataset
10. Supplementary Table S10: List of primers
11. Supplementary Table S11: List of plasmids
12. Supplementary Table S12: List of strains: phage and bacteria

Supplementary notes:

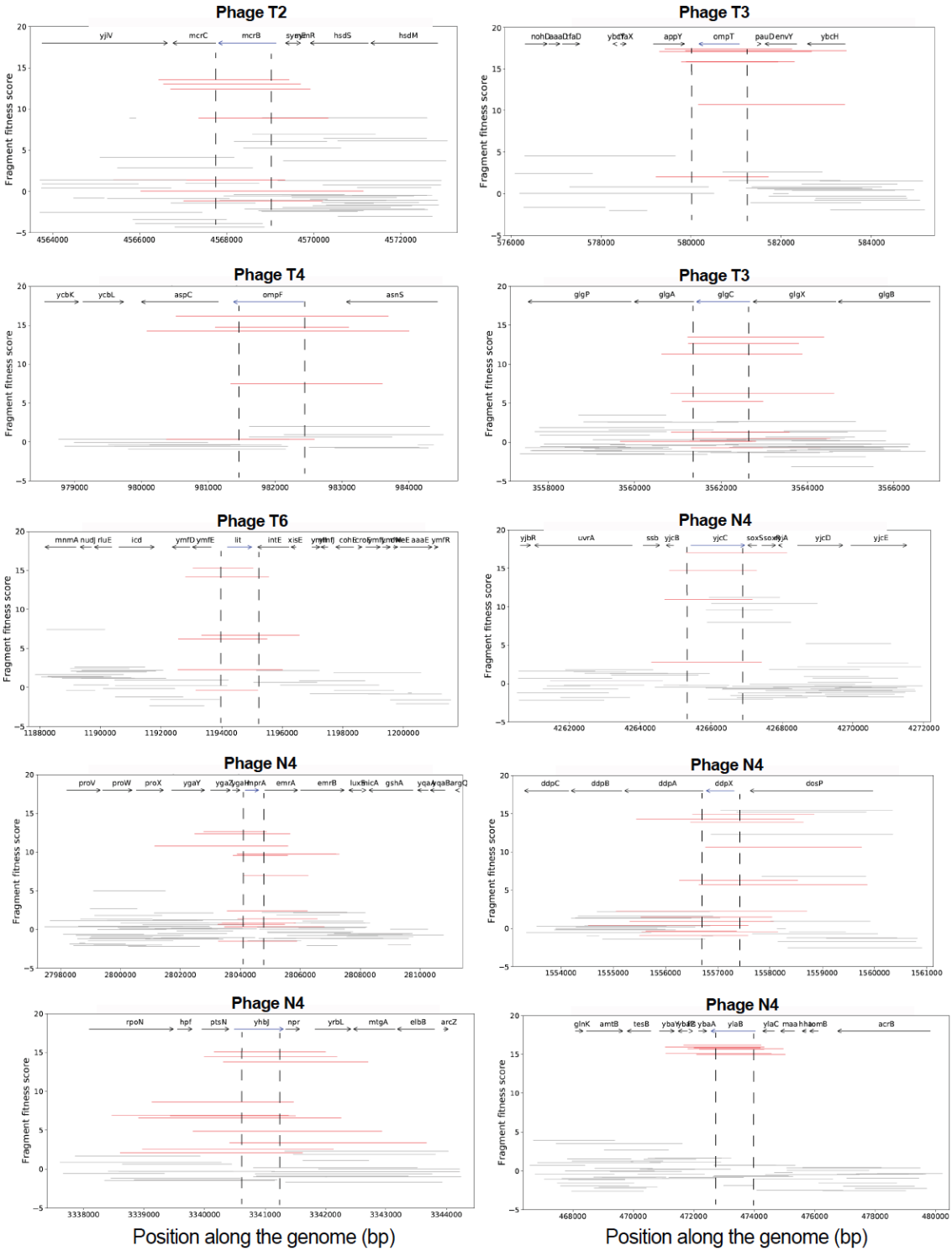
Detailed information on each phage, known host factors important in phage infection and top-scoring hits for each screen. Supplementary notes will be made available upon request.



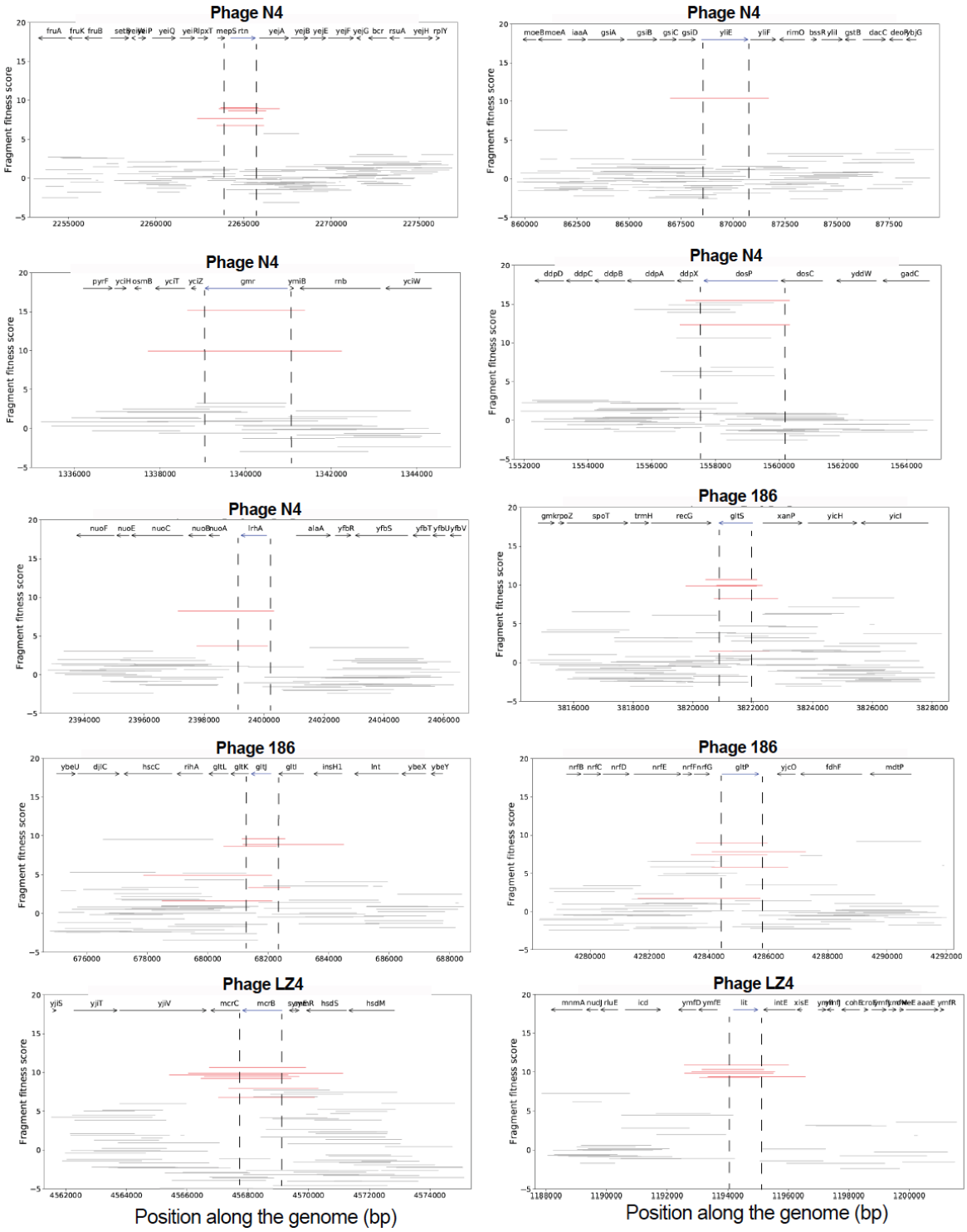
Supplementary Fig. S1: **Map of *igaA* mutants:** *igaA* mutants and their fitness scores from *E. coli* K-12 RB-TnSeq screen. Schematic of rcs phosphorelay with the predicted topology of IgaA [159] is shown at the top. RB-TnSeq mutants in IgaA are mapped to the predicted topology of IgaA (top) and also mapped on to *igaA* nucleotide sequence, with mutant position and fitness scores. Mutants A, B, and D were constructed and their mucoidy and phage resistance phenotype were confirmed (mutant A data is shown in Fig 5A, mutant B and D data is not shown). Though full-length deletion of *igaA* has not been possible, our results indicate that disruption between aa 22-151 is dispensable. Strain with *igaA*-mutant A was further subjected to EOP and RNA-seq analysis presented in the main text.



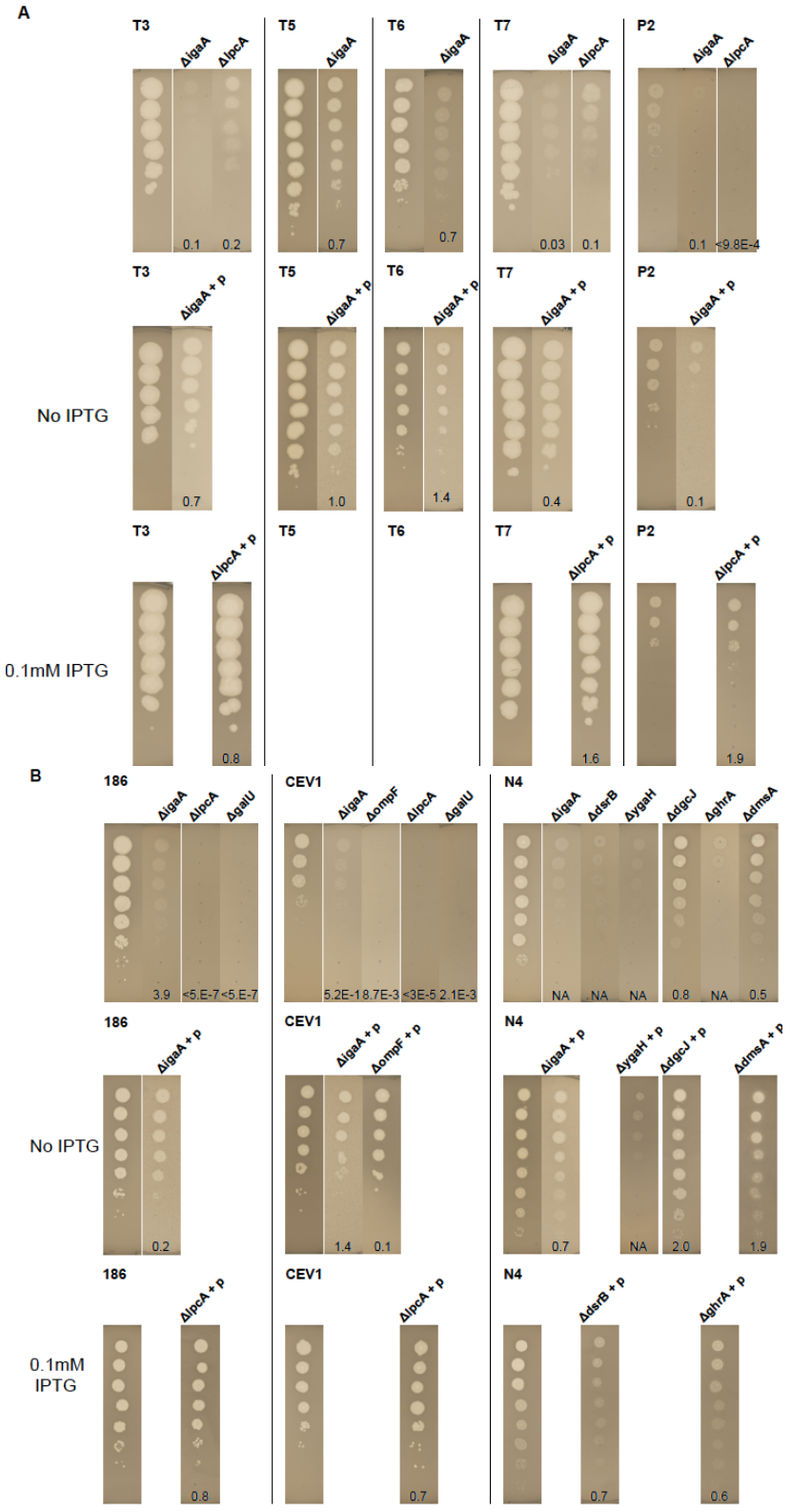
Supplementary Fig. S2: **Dub-seq viewer plots for high-scoring genomic fragments in the presence of phages.** Following top candidates are shown: high scoring fragments encoding *micF* and *degP* for CEV1 phage; *aes, cpdA, malY, glk, sdiA* and *gltP* for λ phage cI857. Red lines represent fragments covering highlighted genes completely (start to stop codon), while grey colored fragments either cover the highlighted gene partially or do not cover the highlighted gene completely.



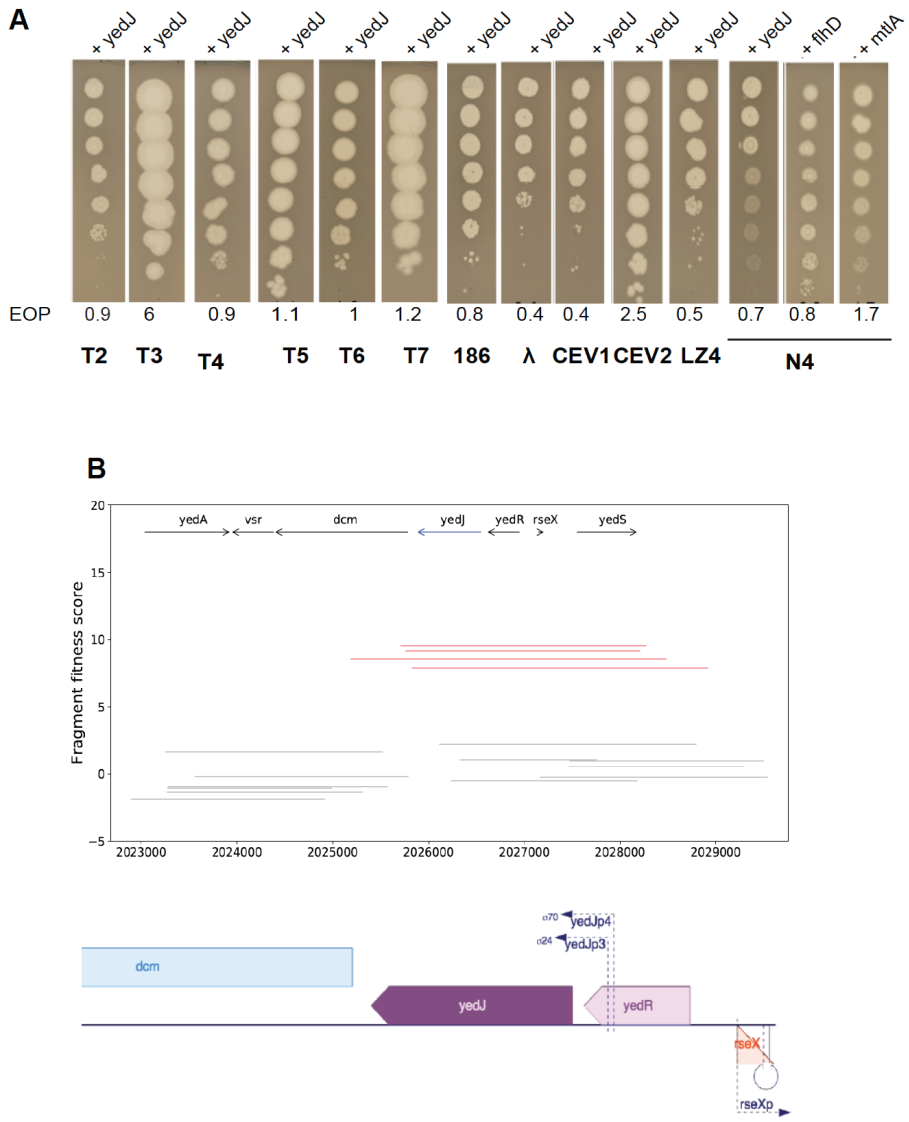
Supplementary Fig. S3: **Dub-seq viewer plots for high-scoring genomic fragments in the presence of phages**. Following top candidates are shown: high scoring fragments encoding *mcrB* for T2 phage; *ompT* and *glgC* for T3 phage; *ompF* for T4 phage; lit for T6 phage; *yjcC (pdeC), mprA, ddpX, yhbJ(rapZ), ylaB (pdeB)* fo N4 phage. Red lines represent fragments covering highlighted genes completely (start to stop codon), while grey colored fragments either cover the highlighted gene partially or do not cover the highlighted gene completely.



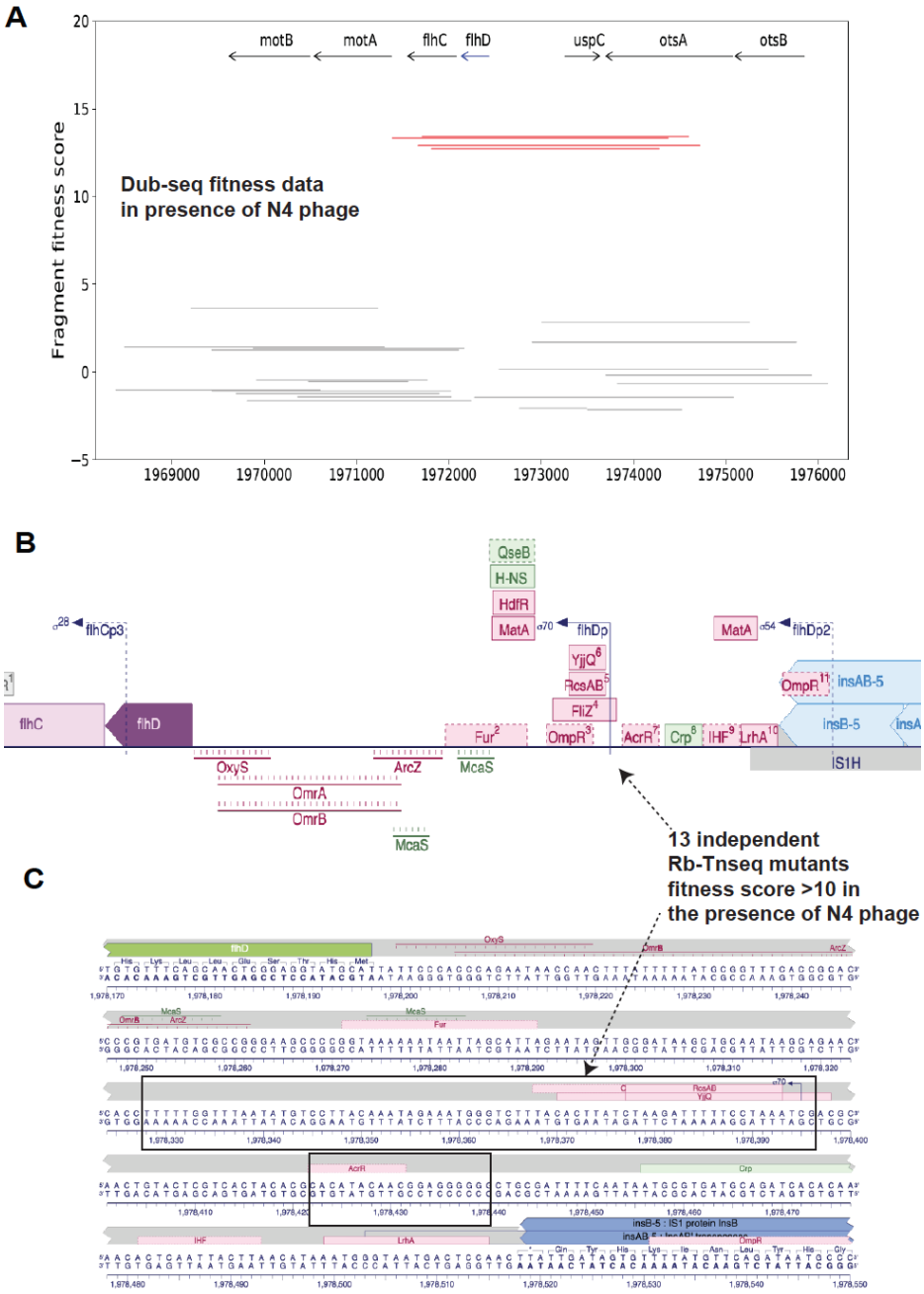
Supplementary Fig. S4: **Dub-seq viewer plots for high-scoring genomic fragments in the presence of phages.** Following top candidates are shown: high scoring fragments encoding *rtn(pdeN), yliE (pdeI), gmr (pdeR), dosP (pdeO)* and *lrhA* for N4 phage; *gltS, gltJ* and *gltP* for 186 phage; *mcrB* and *lit* for LZ4 phage. Red lines represent fragments covering highlighted genes completely (start to stop codon), while grey colored fragments either cover the highlighted gene partially or do not cover the highlighted gene completely.



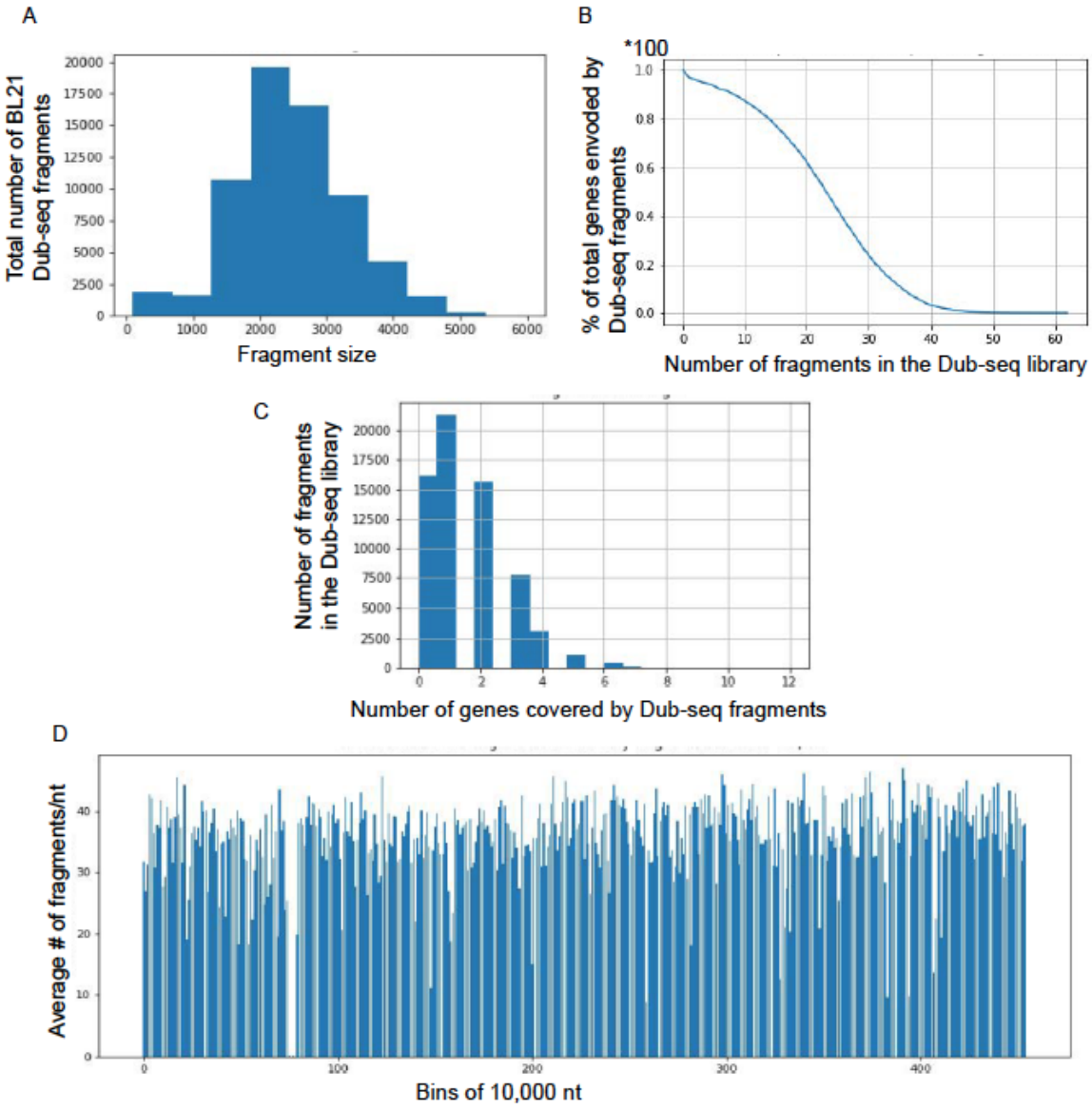
Supplementary Fig. S5: **Estimation of efficiency of plating (EOP) with KEIO deletion strains and genetic complementation** A and B: Efficiency of plating experiments with KEIO strains and ASKA plasmid complementation of respective genes (indicated by +p) in the presence of different phages. We used No IPTG or 0.1 mM IPTG for inducing expression of genes from ASKA plasmid. All experiments were in *E. coli* K-12 BW25113 strain.



Supplementary Fig. S6: **Examples of candidates that showed high fitness scores in our Dub-seq screen but failed in EOP validation experiments** (A) Efficiency of plating experiments with ASKA plasmid expressing genes (shown as +gene names) in the presence of different phages. We used No IPTG and 0.1 mM IPTG for inducing expression of genes from ASKA plasmid. We used BW25113 strain with an empty vector for EOP calculations. (B) Dub-seq viewer plots for high scoring fragments (red bars) encoding *yedJ*, along with neighboring genes (grey bars) with fitness score in the presence of T4 phage on y-axis. The Ecocyc operon [165] view for *yedJ* regions is on the bottom. The genomic fragments encoding *yedJ* also encode a small RNA *rseX* (RNA suppressor of extracytoplasmic stress protease) that binds to RNA-binding protein Hfq (a global regulator) and specifically targets *ompA* and *ompC* mRNAs. Our library does not have genomic fragments that can resolve *rseX* contribution to fitness independently. Overexpression of RseX has also been shown to increase biofilm formation [238] indicating the role of *yedJ-rseX* locus on resistance to T4 phage and other phages.



Supplementary Fig. S7: **Dub-seq and RB-TnSeq data uncovers *flhD* upstream loci important in N4 phage infection** (A) Dub-seq viewer plots for high scoring fragments (red bars) encoding *flhD* and upstream region along with neighboring genes (grey bars) with fitness score on the y-axis. Overexpression of *flhD* failed to demonstrate strong phage plating defects in our EOP validation experiments (Supplementary Fig. S6A) (B) The Ecocyc operon view of the regulatory region upstream of *flhD* with multiple transcription factor binding sites [165] (C) Zoom-in view of the upstream region of *flhD* operon with arrows identifying the location of 13 RB-TnSeq insertion mutants that have fitness score >10 in the presence of N4 phage. These results indicate the role of *flhD* regulatory region on N4 phage growth. We speculate that these TnSeq mutants of *flhD* regulatory region may not be overexpressing *flhD* as our EOP experiments failed to validate *flhD* overexpression as the cause of N4 phage resistance.



Supplementary Fig. S8. Description of *E. coli* BL21 Dub-seq library (A) The fragment insert size distribution in the *E. coli* BL21 Dub-seq library. (B) Cumulative distribution plot showing the percentage of genes in the *E. coli* BL21 genome (y-axis) covered by a number of independent genomic fragments (x-axis). (C) The distribution of the number of genes that are completely covered (start to stop codon) per genomic fragment in the *E. coli* BL21 Dub-seq library. (D) Genome coverage of E. *coli* BL21 Dub-seq library in 10,000 kB windows mapped to *E. coli* BL21-DE3

Methods

Bacterial strains and growth conditions

The primers and plasmids used in this study are listed in Supplementary Table S10 and Table S11, respectively. The bacterial strains and their sources are listed in Supplementary Table S12. All plasmid manipulations were performed using standard molecular biology techniques [239]. All enzymes were obtained from New England Biolabs (NEB) and oligonucleotides were received from Integrated DNA Technologies (IDT). Unless noted, all strains were grown in LB supplemented with appropriate antibiotics at 37ºC in the Multitron shaker. All bacterial strains were stored at -80ºC for long-term storage in 15% sterile glycerol (Sigma). The genotypes of *E. coli* strains used in the assays include BW25113 (K-12 lacI+rrnBT14 Δ(araB–D)567 Δ(rhaD–B)568 ΔlacZ4787(::rrnB-3) hsdR514 rph-1), *E. coli* MG1655 (K-12 F– λ– ilvG– rfb-50 rph-1) and *E. coli* BL21 (B F– ompT gal dcm lon hsdSB(rB–mB–) [malB+]K-12(λS)). The genetic and phenotypic differences between these strains are well documented [100,172–174,240].

Bacteriophages and propagation

The bacteriophages used in this study and their sources are listed in Supplementary Table S12. All phages except P2 phage and N4 phage were propagated on *E. coli* BW25113 strain. To propagate P2 phage and N4 phage we used *E. coli* C and *E. coli* W3350 strains respectively. We used a host-range mutant of T3 (from our in-house phage stock that can grow on both *E. coli* K-12 and BL21 strains), mutant λ phage (temperature-sensitive mutant allele cI857) [9,241] and a strictly virulent strain of P1 phage (P1*vir*) [242] that favors a lytic phage growth cycle. We followed standard protocols for propagating phages [224]. Phage titer was estimated by spotting 2 ul of a 10-fold serial dilution of each phage in SM buffer (Teknova) on a lawn of *E. coli* BW25113 via top agar overlay method using 0.7% LB-agar. SM buffer was supplemented with 10 mM calcium chloride and magnesium sulphate (Sigma). We routinely stored phages as filter-sterilized (0.22um) lysates at 4ºC.

Construction of BL21 RB-TnSeq library

We created the *E. coli* BL21-ML4 transposon mutant library by conjugating *E. coli* BL21 strain with *E. coli* WM3064 harboring the pKMW3 mariner transposon vector library (APA752) [63]. We grew *E. coli* BL21 at 30ºC to mid-log-phase and combined equal cell numbers of BL21 and donor strain APA752, conjugated them for 5 hrs at 30°C on 0.45-m nitrocellulose filters (Millipore) overlaid on LB agar plates containing diaminopimelic acid (DAP) (Sigma). The conjugation mixture was then resuspended in LB and plated on LB agar plates with 50 ug/ml kanamycin to select for mutants. After 1 day of growth at 30°C, we scraped the kanamycin-resistant colonies into 25 ml LB, determined the OD600 of the mixture, and diluted the mutant library back to a starting OD600 of 0.2 in 250 ml of LB with 50 ug/ml kanamycin. We grew the diluted mutant library at 30°C for a few doublings to a final OD600 of 1.0, added glycerol to a final volume of 15%, made multiple 1-ml 80°C freezer stocks, and collected cells for genomic DNA extraction. To link random DNA barcodes to transposon insertion sites, we isolated the

genomic DNA from cell pellets of the mutant libraries with the DNeasy kit (Qiagen) and followed published protocol to generate Illumina compatible sequencing libraries [63,67]. We then performed single-end sequencing (150 bp) with the HiSeq 2500 system (Illumina). Mapping the transposon insertion locations and the identification of their associated DNA barcodes was performed as described previously [63]. Of 4,195 protein-coding genes in *E. coli* BL21, our BL21 RB-TnSeq library has fitness estimates for 3,083. 12 independent strains were used to compute fitness for the typical protein-coding gene.

Construction of BL21 Dub-seq library

To construct *E. coli* BL21 Dub-seq library, we used the dual barcoded pFAB5526 plasmid library with a kanamycin resistance marker (https://benchling.com/s/seq-1Gkg3lrrSno4EF0Ye11k). The Dub-seq backbone plasmid pFAB5526 is the same in design and genetic composition to the original pFAB5491 Dub-seq plasmid (https://benchling.com/s/seq-39Hoh4d1AResilOUPVJ9)[65] except for the kanamycin resistance marker and a mobility gene present on pFAB5526. We mapped the dual barcodes of the pFAB5526 library via the Barcode-Pair sequencing (BPseq) protocol [63,65]. We sequenced BPseq samples on HiSeq 2500 system with 150 bp single-end runs. We then cloned 3 kbps of *E. coli* BL21 genomic fragments between UP and DOWN barcodes by ligating the end-repaired genomic fragments with Pml1 restriction digested pFAB5526, electroporating the ligation into *E. coli* DH10B cells (NEB) and selecting the transformants on LB agar plates supplemented with kanamycin (50 ug/ml). We scraped the kanamycin resistant colonies into 25 ml LB, and diluted the transformant mixture to a starting OD600 of 0.2 in 250 ml of LB with 50 ug/ml kanamycin and grew the library at 30°C for few doublings to a final OD600 of 1.0. Finally, we added glycerol to a final volume of 15%, made multiple 1-ml 80°C freezer stocks, and collected cells for plasmid DNA extraction (Qiagen). Next, we mapped the cloned genomic fragment and its pairings with neighboring dual barcodes via Barcode-Association-with Genome fragment sequencing or BAGseq [63,65]. We sequenced the BAGseq samples on HiSeq 2500 system with 150 bp single-end runs using the reported sample preparation steps [63,65]. The data processing of BPseq and BAGseq steps was performed using *Dub-seq* python library with default parameters (https://github.com/psnovichkov/DubSeq) as detailed earlier [63,65]. We mapped *E. coli* BL21 Dub-seq library to *E. coli* BL21-DE3 genome sequence [173]. BL21 Dub-seq library was then electroporated into BL21DE3C43 strain and the transformants were collected into 25 ml LB. The transformant mixture was then diluted to a starting OD600 of 0.2 in 250 ml of LB with 50 g/ml kanamycin and grew the library at 30°C for a few doubling to a final OD600 of 1.0. Finally, we added glycerol to a final volume of 15%, made multiple 1-ml 80°C freezer stocks. These stocks were further used for pooled fitness assays as described below. The BL21 Dub-seq library is made up of 65,788 unique barcoded fragments. The average fragment size of the library is 2.5 kb and the majority of fragments covered 2–3 genes in their entirety (Supplementary Fig. S8). Similar to *E. coli* BW25113 Dub-seq library [63,65], BL21 Dub-seq library covers 85% of genes from start to stop codon by at least 5 independent genomic fragments and 97% of all genes are covered by at least one fragment. In total, 132 genes are not covered in their entirety by any Dub-seq fragment.

Competitive growth experiments with RB-TnSeq library

A single aliquot (1-ml) of a mutant library was thawed, inoculated into 25 ml of medium supplemented with kanamycin (50 ug/ml), and grown to OD600 of 0.6-0.8 at 37°C. After the mutant library recovered and reached mid-log phase, we collected cell pellets as a common reference for BarSeq (termed time-zero or start samples) and used the remaining cells to set up competitive mutant fitness assays in the presence of different phages at different multiplicity of infection. For performing planktonic culture assays, we diluted the recovered mutant library stock to a starting OD600 of 0.04 in 2X LB media (350 ul) and mixed in equal volume (350 ul) of phages diluted in phage dilution buffer. We also set up control 'no-phage' competitive mutant fitness assays wherein we replaced phages with simply the phage dilution buffer. The mutant library experiments were grown in the wells of a 48-well microplate (700 ul per well). We grew the microplates in Tecan Infinite F200 readers with orbital shaking and OD600 readings every 15 min for 8 hrs. After the experiment, survivors were collected, pelleted and stored at 80°C prior to genomic DNA extraction. For solid plate-based assays, we incubated the mixture of phage and diluted mutant library at room temperature for 15 minutes, and then plated the mixture on LB-agar supplemented with kanamycin plates, and incubated at 37°C overnight. The next day, the resistant colonies were scraped, resuspended in 1 ml LB media, and pelleted. Assay pellets were typically stored at -80°C prior to genomic DNA extraction. The 8 hr assay period was decided based on our preliminary time-course experiment results for a specific phage at different MOIs, wherein we took intermittent samples between 2-10 hrs, processed as detailed below and found to yield consistent top scoring hits (results not shown). In addition, assay samples after 8 hrs provided sufficient DNA for sample processing step.

Competitive growth experiments with Dub-seq library

Similar to RB-TnSeq competitive fitness assays, a single aliquot of the *E. coli* Dub-seq library (*E. coli* BW25113 library expressed in *E. coli* BW25113 or *E. coli* BL21 Dub-seq library in *E. coli* BL21DE3C43 strain) was thawed, inoculated into 25 ml of LB medium supplemented with appropriate antibiotics (chloramphenicol 30 ug/ml for *E. coli* BW25113 library and kanamycin (50 ug/ml) for *E. coli* BL21 Dub-seq library) and grown to OD600 of 0.6-0.8. At mid-log phase, we collected cell pellets as a common reference for BarSeq (termed start or time-zero samples) and we used the remaining cells to set up competitive fitness assays in the presence of different phages at different multiplicity of infection. For performing planktonic culture assays, we diluted the recovered Dub-seq library stock to a starting OD600 of 0.04 in 2X LB media and mixed in equal volume (350 ul) with phages diluted in phage dilution buffer. We also set up control 'no-phage' competitive mutant fitness assays wherein we replaced phages with simply the phage dilution buffer. The Dub-seq library cultures were grown in the wells of a 48-well microplate (700 ul per well) and grew the microplates in Tecan Infinite F200 readers with orbital shaking and OD600 readings every 15 min for 3-8 hrs. After the experiment, survivors were collected, pelleted and stored at 80°C prior to plasmid DNA extraction. For solid plate-based assays, we incubated the mixture of phage and diluted Dub-seq library at room temperature for 15 minutes, and then plated the mixture on LB-agar plates supplemented with appropriate antibiotics, and incubated at 37°C overnight. Next day, the resistant colonies were

scraped, resuspended in 1ml LB media, and pelleted. Assay pellets were typically stored at 80°C prior to plasmid DNA extraction. We also performed these assays using *E. coli* strains with empty plasmid (used in Dub-seq library creation).

BarSeq of RB-TnSeq and Dub-seq pooled fitness assay samples

We isolated genomic DNA from RB-TnSeq library samples using the DNeasy Blood and Tissue kit (Qiagen). Plasmid DNA from the Dub-seq library samples was extracted either individually using the Plasmid miniprep kit (Qiagen) or in 96-well format with a QIAprep 96 Turbo miniprep kit (Qiagen). We performed 98°C BarSeq PCR protocol as described previously [63,65] with the following modifications. BarSeq PCR in a 50 ul total volume consisted of 20 umol of each primer and 150 to 200 ng of template genomic DNA or plasmid DNA. For the HiSeq4000 runs, we used an equimolar mixture of BarSeq_P2 primers along with new Barseq3_P1 primers. The BarSeq_P2 primer contains the tag that is used for demultiplexing by Illumina software, while the new Barseq3_P1 primer contained an additional sequence to verify that it came from the expected sample. The new Barseq3_P1 primer contains the same sequence as BarSeq_P1 reported earlier with 1 to 4 N's (which varies with the index) followed by the reverse (not the reverse complement) of the 6-nucleotide index sequence [243]. This modification to earlier BarSeq PCR protocol was done to eliminate the barcode bleed-through problem in sequencing and also to aid in cluster and sample discrimination on the HiSeq4000. All experiments done on the same day and sequenced on the same lane are considered as a 'set'. Equal volumes (5 ul) of the individual BarSeq PCRs were pooled, and 50 ul of the pooled PCR product was purified with the DNA Clean and Concentrator kit (Zymo Research). The final BarSeq library was eluted in 40 ul water. The BarSeq libraries were sequenced on Illumina HiSeq4000 instrument with 50 SE runs. We usually multiplexed 96 Barseq samples per lane for both RB-TnSeq and Dub-seq library samples.

Data processing and analysis of BarSeq reads

RB-TnSeq fitness data was analysed as previously described [63] with additional filters as presented below. Briefly, the fitness value of each strain (an individual transposon mutant) is the normalized log2(strain barcode abundance at end of experiment/strain barcode abundance at the start of the experiment). The fitness value of each gene is the weighted average of the fitness of its strains. We applied filters on experiments such that the mean reads per gene is >= 10. As we have reported earlier [63], in a typical experiment without stringent positive selection, gene fitness score (fit) > 1 and associated t-like statistic> 4 suffices to give a low rate of false positives. Because of the stringent positive selection in phage assays, our standard quality metrics reported earlier [63] were not suitable. Because most of the sequencing reads were from a handful of phage-resistant mutants in the population, the majority of the strains in the library did not have enough reads to accurately calculate fitness values. To determine suitable filters, we compared fitness data between two halves of each gene. As we have

several insertions in most genes, we can compute fitness values for the two halves of a gene separately and then plot the first half and second half fitness values for each gene (that has sufficient coverage) against each other as described earlier [63]. These plots are available in the figshare file https://doi.org/10.6084/m9.figshare.11413128. Phage T2 assays on solid agar showed poor consistency and were dropped out of the analysis. We observed from the first half and second half fitness plots that fitness values < 5 were often not consistent. To reduce false positives, we required that fit≥ 5; t ≥ 5; standard error = fit/t ≤ 2; and fit ≥ maxFit - 8, where maxFit is the highest fitness value of any gene in an experiment. The limit of maxFit - 8 was chosen based on the fitness score of positive controls (host factors that are known to interfere phage growth when deleted) and from experimental validations of new hits.

For the Dub-seq library, fitness data was analyzed using *barseq* script from the *Dub-seq* python library with default parameters as previously described [65]. From a reference list of barcodes mapped to the genomic regions (BPSeq and BAGSeq), and their counts in each sample (BarSeq), we estimate fitness values of each genomic fragment (strain) using *fscore* script from the Dub-seq python library. The *fscore* script identifies a subset of barcodes mapped to the genomic regions that are well represented in the time-zero samples for a given experiment set. We require that a barcode to have at least 10 reads in at least one time-zero (sample before the experiment) sample to be considered a valid barcode for a given experiment set. Then the *fscore* script calculates fitness score (normalized ratio of counts between the treatment sample and sum of counts across all time-zero samples) only for the strains with valid barcodes. The fitness scores calculated for all Dub-seq fragments, we estimate a fitness score for each individual gene *gscore* that is covered by at least one fragment as detailed earlier using non-negative least squares regression [65]. The non-negative regression determines if the high fitness of the fragments covering the gene is due to a particular gene or its nearby gene, and avoids overfitting. We applied the same data filters as reported earlier [65] to ensure that the fragments covering the gene had a genuine benefit. Briefly, we identify a subset of the effects to be reliable, if the fitness effect was large relative to the variation between start samples (|score| >= 2); the fragments containing the gene showed consistent fitness (using a *t* test); and the number of reads for those fragments was sufficient for the gene score to have little noise. As in RB-TnSeq BarSeq analysis, Dub-seq data also showed strong positive selection in the presence of phages, where few strains accounting for the most reads and displaying very high fitness scores. To reduce false positives, we applied an additional filter of gene fitness score >= 4 threshold. This threshold cut-off was chosen by analyzing the fragment fitness scores against the number of supported reads, and that the cut off value agrees with experimental validations carried out in this work. Finally, we estimate the false discovery rate for high-confidence effects as detailed earlier [65]. Briefly, to estimate the number of hits in the absence of genuine biological effects, we randomly shuffled the mapping of barcodes to fragments, recomputed the mean scores for each gene in each experiment, and identified high-confidence effects as for the genuine data. We repeated the shuffle procedure 10 times. The complete dataset of fragment and gene scores are provided at https://doi.org/10.6084/m9.figshare.11838879.v2. Dub-seq data for *E. coli* strain with empty

plasmid (used in Dub-seq library creation) was consistent with *E. coli* with no plasmid (data not shown) indicating we do not see any plasmid-mediated fitness effects.

E. coli MG1655 CRISPRi Library

The design and construction of *E. coli* MG1655 CRISPRi library are detailed elsewhere (Rishi et al 2020 submitted). Briefly, the *E. coli* MG1655 CRISPRi library is made up of 32,992 unique sgRNAs targeting 4,457 genes (including small RNA genes, insertion elements and prophages), 7,442 promoter regions, and 1,060 transcription factor binding sites. The sgRNA library is driven by pBAD promoter (induced by arabinose) and is expressed from a high copy plasmid (ColE1 replication origin) in *E. coli* K-12 MG1655 strain harboring a genomically encoded aTc-inducible *dCas9* gene.

To perform pooled fitness experiments using *E. coli* MG1655 CRISPRi library, a single aliquot of the library was thawed, inoculated into 5 ml of LB medium supplemented with carbenicillin, kanamycin, and glucose and grown to OD600 of 0.5. At mid-log phase, we collected cell pellets as an initial time point (time zero) of the library and diluted the remaining culture in induction media (LB broth with arabinose (0.1%), aTc (200 ng/mL), carbenicillin 100 µg/mL, and kanamycin 30 µg/mL) to initiate dCas9 and sgRNA expression for about 6 doublings (to OD600 about 0.5). Finally, we diluted the library to OD600 of 0.1 in 2X LB supplemented with induction media and incubated 350 ul of the library with 350 ul of diluted phage stocks (MOI of 1) in 5 ml test tubes at room temperature for 15 minutes. We also set up a control ‘no-phage’ culture wherein we simply mixed the phage dilution buffer with the library. We then moved these competitive fitness assay cultures to 37°C with shaking (200 rpm) for 2 hrs, collected the entire cell pellet after 2 hrs and stored at -80°C prior to plasmid DNA extraction to isolate the library. The plasmid library was extracted using QIAprep Spin Miniprep Kit, used for PCR to generate Illumina sequencing samples and sequenced on Illumina HiSeq2500 instrument with 100 SE runs. Only samples that yielded more than 2 million reads (average library read depth of ~60) were used in the downstream analysis.

Illumina reads were quality filtered, trimmed, and mapped using the same procedure as Rishi et al 2020 (submitted) to generate strain abundances (i.e. read counts) for each sgRNA library member (recall that each strain in the library is uniquely determined by the 20 bp variable region of the sgRNA it harbors). The fitness of each sgRNA library member was calculated as the log2 fold-change in abundance of the sgRNA after the experiment vs before the experiment using edgeR[244,245]. For a given enrichment comparison, sgRNAs with fewer than 10 read counts in each replicate of the time-zero and end of experiment samples were filtered out of the analysis. In the edgeR pipeline, each sample was normalized for sequencing depth using the edgeR function calcNormFactors, pseudocounts and dispersions were calculated using the edgeR functions estimateCommonDisp and estimateTagwiseDisp, and the log2 fold-change and corresponding p-values were calculated using the edgeR function exactTest, which is based on a qCML (quantile-adjusted conditional maximum likelihood) method. FDR-adjusted p-values were calculated using the Benjamini-Hochberg method. We performed a post-hoc analysis on these fitness scores and picked a threshold of fitness >= 2

and FDR-adjusted p-value < 0.05 for new hits of interest based on the satisfaction of these criteria by both positive controls (known phage receptors for example *fadL* for T2 phage and *waa* operon for 186 phage) and validated hits (for example, *dsrB* in the presence of N4 phage). Finally, gene fitness scores were calculated by taking the median fitness scores of (filtered) sgRNAs targeting a given gene.

Construction of *igaA* mutant strains

We created our *igaA* genetic disruption mutant through a modified recombineering method using pSIM5 [65,246] and approximate insertion site based off of RB-TnSeq mapping (Supplementary Fig. S1). Primers were designed to amplify a kanamycin resistant (kanR) selection marker with 50 bp homology to the insertion site of interest corresponding to a kanR insertion at 51 bp internal to *igaA*. PCRs were generated and gel-purified through standard molecular biology techniques [239] and stored at -20°C until use. Deletions were performed by incorporating the above dsDNA template into the BW25113 genome through standard pSIM5-mediated recombineering methods. First temperature-sensitive recombineering vector, pSIM5, was introduced into BW25113 through standard electroporation protocols and grown with chloramphenicol at 30°C. Recombination was performed through electroporation with an adapted pSIM5 recombineering protocol. Post-recombination, clonal isolates were streaked onto kanamycin plates without chloramphenicol at 37°C to cure the strain of pSIM5 vector, outgrown at 37˚C and stored at -80°C until use. A detailed protocol is available upon request. Disruptions were verified by colony PCR followed by Sanger sequencing at the targeted locus and 16S regions. The *igaA* mutant map is given here https://benchling.com/s/seq-pZEspfGx8K9tYzixzMoJ

Experimental validation of individual phage resistance phenotype

To validate the phage resistance phenotypes from both LOF and GOF screens, we performed efficiency of plating (EOP) assays on select *E. coli* mutants from the KEIO collection [100] and overexpression ASKA library [100,155], respectively. The complete list of primers, plasmids and strains used in validation assays are provided in Supplementary Tables S10-S12. All deletion strains and plasmids used in this work were confirmed via Sanger sequencing. An isogenic deletion mutant in *E. coli* BL21 for *fadL* was generated by phage P1 mediated transduction of kanamycin resistance from individual KEIO mutants [100]. The gene deletion strains from KEIO library and the BL21 transductants were cultured in LB media or LB agar supplemented with kanamycin (25 ug/ml), while ASKA strains were cultured in LB media or LB agar supplemented with chloramphenicol (30 ug/ml). The plasmids from ASKA library were extracted using QIAprep Miniprep kit (Qiagen), and electroporated into *E. coli* BW25113 strains for GOF validations or respective Keio mutants for complementation of LOF phenotypes. The bottom agar was supplemented with 0.1 mM IPTG to induce protein production from ASKA plasmids. Based on toxicity associated with gene overexpression, IPTG levels (0 to 0.1 mM) were adjusted and cultures were grown overnight at 30°C.

Phages were quantified by spot titration method. 2 ul of serially ten-fold diluted phages were spotted on a solidified lawn of ~4 ml 0.5 % top-agar inoculated with 100 µl of a fresh overnight bacterial culture and incubated overnight at RT, 30°C or 37°C. The EOP was calculated as the ratio of the number of plaques on mutant or overexpression strain to the number of plaques on the parental strains (BW25113 or BL21). The EOP's were calculated at least by two biological replicates. Supplementary Fig. S5 lists all EOP estimation numbers and plaque images.

RNA-seq experiments

Transcriptomes were collected and analyzed for strains BW25113 (N=3), knockout mutants for *igaA* (N=3), and overexpression strains for *pdeB* (N=1), *pdeC* (N=1), *pdeL* (N=3), *pdeN* (N=1), *pdeO* (N=1), and *ygbE* (N=3). All cultures for RNAseq experiments were performed on the same day from unique overnights and subsequent outgrowths.

Strains were grown overnight in LB with an appropriate selection marker at 30˚C. Strains were diluted to OD600 ~0.02 in 10 mL LB with the appropriate selection marker and, for overexpression strains, 0.1mM IPTG. Cultures were grown at 30˚C at 180 RPM until they reached an OD600 0.3-0.4. Samples were collected as follows: 400 µL of culture was added to 800 µL RNAProtect (Qiagen), incubated for 5 minutes at room temperature, and centrifuged for 10 minutes at 5000xg. RNA was purified using RNeasy RNA isolation kit (Qiagen) and quantified and quality-assessed by Bioanalyzer. Library preparation was performed by the Functional Genomics Laboratory (FGL), a QB3-Berkeley Core Research Facility at UC Berkeley. Illumina Ribo-Zero rRNA Removal Kit was used to deplete ribosomal RNA. Subsequent library preparation steps of fragmentation, adapter ligation and cDNA synthesis were performed on the depleted RNA using the KAPA RNA HyperPrep kit (KK8540). Truncated universal stub adapters were used for ligation, and indexed primers were used during PCR amplification to complete the adapters and to enrich the libraries for adapter-ligated fragments. Samples were checked for quality on an Agilent Fragment Analyzer. Sequencing was performed at the Vincent Coates Sequencing Center on a HiSeq4000 using 100PE runs.

RNA-seq Data Analysis

For all RNA-seq experiments, analyses were performed through a combination of KBase [247] and custom jupyter notebook-based methods. Briefly, Illumina reads were trimmed using Trimmomatic [248] v0.36 and assessed for quality using FASTQC. Trimmed reads were mapped to the *E. coli* BW25113 genome (NCBI Accession: CP009273) with HISAT2 [249]. Alignments were quality-assessed with BAMQC. From this alignment, transcripts were assembled and abundance-estimated with StringTie [250]. For experiments with sufficient biological replicates, tests for differential expression were performed on normalized gene counts by DESeq2 (negative binomial generalized linear model) [251]. Additional analyses for all experiments were performed in Python3 and visualized employing matplotlib and seaborn packages. For analyses involving DESeq2, conservative thresholds were employed for assessing differentially expressed genes. Genes were considered differentially expressed if

they possessed a Bonferoni-corrected p-value below a threshold of 0.001 and an absolute log2 fold change greater than 2. For supplemental analyses without tests for differential expression, data were analyzed by comparing log2 fold-changes of Transcripts Per Kilobase Million (TPM) counts. The complete datasets are provided Supplementary Table S6 and S7.

Data availability

Complete data from RB-TnSeq experiments is deposited here https://doi.org/10.6084/m9.figshare.11413128 and Dub-seq experiments is deposited here: https://doi.org/10.6084/m9.figshare.11838879.v2

REFERENCES

1. Breitbart M, Rohwer F. Here a virus, there a virus, everywhere the same virus? Trends in Microbiology. 2005. pp. 278–284. doi:10.1016/j.tim.2005.04.003
2. Wilhelm SW, Suttle CA. Viruses and Nutrient Cycles in the Sea. BioScience. 1999. pp. 781–788. doi:10.2307/1313569
3. Koskella B, Taylor TB. Multifaceted Impacts of Bacteriophages in the Plant Microbiome. Annu Rev Phytopathol. 2018;56: 361–380.
4. Shkoporov AN, Clooney AG, Sutton TDS, Ryan FJ, Daly KM, Nolan JA, et al. The Human Gut Virome Is Highly Diverse, Stable, and Individual Specific. Cell Host Microbe. 2019;26: 527–541.e5.
5. Mirzaei MK, Maurice CF. Ménage à trois in the human gut: interactions between host, bacteria and phages. Nature Reviews Microbiology. 2017. pp. 397–408. doi:10.1038/nrmicro.2017.30
6. Suttle CA. Marine viruses — major players in the global ecosystem. Nature Reviews Microbiology. 2007. pp. 801–812. doi:10.1038/nrmicro1750
7. De Sordi L, Lourenço M, Debarbieux L. The Battle Within: Interactions of Bacteriophages and Bacteria in the Gastrointestinal Tract. Cell Host Microbe. 2019;25: 210–218.
8. Shkoporov AN, Hill C. Bacteriophages of the Human Gut: The “Known Unknown” of the Microbiome. Cell Host Microbe. 2019;25: 195–209.
9. Casjens SR, Hendrix RW. Bacteriophage lambda: Early pioneer and still relevant. Virology. 2015;479-480: 310–330.
10. Karam JD, Drake JW. Molecular biology of bacteriophage T4. American Society for Microbiology; 1994.
11. Molineux I. T7 Bacteriophages. Encyclopedia of Molecular Biology. 2002. doi:10.1002/047120918x.emb1510
12. Heller KJ. Molecular interaction between bacteriophage and the gram-negative cell envelope. Archives of Microbiology. 1992. pp. 235–248. doi:10.1007/bf00245239
13. De Smet J, Hendrix H, Blasdel BG, Danis-Wlodarczyk K, Lavigne R. Pseudomonas predators: understanding and exploiting phage-host interactions. Nat Rev Microbiol. 2017;15: 517–530.
14. Sandulache R, Prehm P, Kamp D. Cell wall receptor for bacteriophage Mu G(+). J Bacteriol. 1984;160: 299–303.
15. Demerec M, Fano U. Bacteriophage-Resistant Mutants in Escherichia Coli. Genetics. 1945;30: 119–136.
16. Hancock RE, Reeves P. Bacteriophage resistance in Escherichia coli K-12: general pattern of resistance. J Bacteriol. 1975;121: 983–993.

17. Chirakadze I, Perets A, Ahmed R. Phage Typing. Methods in Molecular Biology. 2009. pp. 293–305. doi:10.1007/978-1-60327-565-1_17
18. Bondy-Denomy J, Qian J, Westra ER, Buckling A, Guttman DS, Davidson AR, et al. Prophages mediate defense against phage infection through diverse mechanisms. ISME J. 2016;10: 2854–2866.
19. Moller AG, Lindsay JA, Read TD. Determinants of Phage Host Range in Staphylococcus Species. Appl Environ Microbiol. 2019;85. doi:10.1128/AEM.00209-19
20. Trudelle DM, Bryan DW, Hudson LK, Denes TG. Cross-resistance to phage infection in Listeria monocytogenes serotype 1/2a mutants. Food Microbiol. 2019;84: 103239.
21. Kauffman KM, Hussain FA, Yang J, Arevalo P, Brown JM, Chang WK, et al. A major lineage of non-tailed dsDNA viruses as unrecognized killers of marine bacteria. Nature. 2018. pp. 118–122. doi:10.1038/nature25474
22. Porter NT, Hryckowian AJ, Merrill BD, Gardner JO, Singh S, Sonnenburg JL, et al. Multiple phase-variable mechanisms, including capsular polysaccharides, modify bacteriophage susceptibility in Bacteroides thetaiotaomicron. doi:10.1101/521070
23. Schooley RT, Biswas B, Gill JJ, Hernandez-Morales A, Lancaster J, Lessor L, et al. Development and Use of Personalized Bacteriophage-Based Therapeutic Cocktails To Treat a Patient with a Disseminated Resistant Acinetobacter baumannii Infection. Antimicrob Agents Chemother. 2017;61. doi:10.1128/AAC.00954-17
24. Dedrick RM, Guerrero-Bustamante CA, Garlena RA, Russell DA, Ford K, Harris K, et al. Engineered bacteriophages for treatment of a patient with a disseminated drug-resistant Mycobacterium abscessus. Nat Med. 2019;25: 730–733.
25. Dedrick RM, Jacobs-Sera D, Bustamante CAG, Garlena RA, Mavrich TN, Pope WH, et al. Prophage-mediated defence against viral attack and viral counter-defence. Nat Microbiol. 2017;2: 16251.
26. Mathieu A, Dion M, Deng L, Tremblay D, Moncaut E, Shah SA, et al. Virulent coliphages in 1-year-old children fecal samples are fewer, but more infectious than temperate coliphages. Nat Commun. 2020;11: 378.
27. Bourdin G, Navarro A, Sarker SA, Pittet A-C, Qadri F, Sultana S, et al. Coverage of diarrhoea-associated Escherichia coli isolates from different origins with two types of phage cocktails. Microb Biotechnol. 2014;7: 165–176.
28. Wright RCT, Friman V-P, Smith MCM, Brockhurst MA. Cross-resistance is modular in bacteria–phage interactions. PLOS Biology. 2018. p. e2006057. doi:10.1371/journal.pbio.2006057
29. Denes T, den Bakker HC, Tokman JI, Guldimann C, Wiedmann M. Selection and Characterization of Phage-Resistant Mutant Strains of Listeria monocytogenes Reveal Host Genes Linked to Phage Adsorption. Applied and Environmental Microbiology. 2015. pp. 4295–4305. doi:10.1128/aem.00087-15
30. Leprohon P, Gingras H, Ouennane S, Moineau S, Ouellette M. A genomic approach to understand interactions between Streptococcus pneumoniae and its bacteriophages. BMC Genomics. 2015;16: 972.
31. Duerkop BA, Huo W, Bhardwaj P, Palmer KL, Hooper LV. Molecular Basis for Lytic Bacteriophage Resistance in Enterococci. MBio. 2016;7. doi:10.1128/mBio.01304-16
32. Sumrall ET, Shen Y, Keller AP, Rismondo J, Pavlou M, Eugster MR, et al. Phage resistance at the cost of virulence: Listeria monocytogenes serovar 4b requires galactosylated teichoic acids for InlB-mediated invasion. PLoS Pathog. 2019;15: e1008032.
33. Goldfarb T, Sberro H, Weinstock E, Cohen O, Doron S, Charpak□Amikam Y, et al. BREX is a novel phage resistance system widespread in microbial genomes. The EMBO

Journal. 2015. pp. 169–183. doi:10.15252/embj.201489455
34. Bernheim A, Sorek R. The pan-immune system of bacteria: antiviral defence as a community resource. Nat Rev Microbiol. 2019. doi:10.1038/s41579-019-0278-2
35. Ofir G, Sorek R. Contemporary Phage Biology: From Classic Models to New Insights. Cell. 2018;172: 1260–1270.
36. Young R, Gill JJ. MICROBIOLOGY. Phage therapy redux--What is to be done? Science. 2015;350: 1163–1164.
37. Rostøl JT, Marraffini L. (Ph)ighting Phages: How Bacteria Resist Their Parasites. Cell Host Microbe. 2019;25: 184–194.
38. Samson JE, Magadán AH, Sabri M, Moineau S. Revenge of the phages: defeating bacterial defences. Nature Reviews Microbiology. 2013. pp. 675–687. doi:10.1038/nrmicro3096
39. Campbell A. The future of bacteriophage biology. Nat Rev Genet. 2003;4: 471–477.
40. Weitz JS, Poisot T, Meyer JR, Flores CO, Valverde S, Sullivan MB, et al. Phage-bacteria infection networks. Trends Microbiol. 2013;21: 82–91.
41. Moller AG, Lindsay JA, Read TD. Determinants of Phage Host Range in Staphylococcus Species. Applied and Environmental Microbiology. 2019. doi:10.1128/aem.00209-19
42. Jonge PA de, de Jonge PA, Nobrega FL, Brouns SJJ, Dutilh BE. Molecular and Evolutionary Determinants of Bacteriophage Host Range. Trends in Microbiology. 2019. pp. 51–63. doi:10.1016/j.tim.2018.08.006
43. Dy RL, Richter C, Salmond GPC, Fineran PC. Remarkable Mechanisms in Microbes to Resist Phage Infections. Annu Rev Virol. 2014;1: 307–331.
44. Nobrega FL, Vlot M, de Jonge PA, Dreesens LL, Beaumont HJE, Lavigne R, et al. Targeting mechanisms of tailed bacteriophages. Nature Reviews Microbiology. 2018. pp. 760–773. doi:10.1038/s41579-018-0070-8
45. van Houte S, Buckling A, Westra ER. Evolutionary Ecology of Prokaryotic Immune Mechanisms. Microbiol Mol Biol Rev. 2016;80: 745–763.
46. Brüssow H. Bacteriophage-host interaction: from splendid isolation into a messy reality. Curr Opin Microbiol. 2013;16: 500–506.
47. Broeker NK, Barbirz S. Not a barrier but a key: How bacteriophages exploit host's O-antigen as an essential receptor to initiate infection. Molecular Microbiology. 2017. pp. 353–357. doi:10.1111/mmi.13729
48. Christen M, Beusch C, Bösch Y, Cerletti D, Flores-Tinoco CE, Del Medico L, et al. Quantitative Selection Analysis of Bacteriophage φCbK Susceptibility in Caulobacter crescentus. J Mol Biol. 2016;428: 419–430.
49. Maynard ND, Birch EW, Sanghvi JC, Chen L, Gutschow MV, Covert MW. A forward-genetic screen and dynamic analysis of lambda phage host-dependencies reveals an extensive interaction network and a new anti-viral strategy. PLoS Genet. 2010;6: e1001017.
50. Qimron U, Marintcheva B, Tabor S, Richardson CC. Genomewide screens for Escherichia coli genes affecting growth of T7 bacteriophage. Proc Natl Acad Sci U S A. 2006;103: 19039–19044.
51. Rousset F, Cui L, Siouve E, Becavin C, Depardieu F, Bikard D. Genome-wide CRISPR-dCas9 screens in E. coli identify essential genes and phage host factors. PLoS Genet. 2018;14: e1007749.
52. Chan BK, Sistrom M, Wertz JE, Kortright KE, Narayan D, Turner PE. Phage selection restores antibiotic sensitivity in MDR Pseudomonas aeruginosa. Scientific Reports. 2016. doi:10.1038/srep26717
53. Bohm K, Porwollik S, Chu W, Dover JA, Gilcrease EB, Casjens SR, et al. Genes affecting

progression of bacteriophage P22 infection in Salmonella identified by transposon and single gene deletion screens. Mol Microbiol. 2018;108: 288–305.
54. 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: 1605–1618.
55. Roucourt B, Lavigne R. The role of interactions between phage and bacterial proteins within the infected cell: a diverse and puzzling interactome. Environmental Microbiology. 2009. pp. 2789–2805. doi:10.1111/j.1462-2920.2009.02029.x
56. Doron S, Melamed S, Ofir G, Leavitt A, Lopatina A, Keren M, et al. Systematic discovery of antiphage defense systems in the microbial pangenome. Science. 2018. p. eaar4120. doi:10.1126/science.aar4120
57. Cowley LA, Low AS, Pickard D, Boinett CJ, Dallman TJ, Day M, et al. Transposon Insertion Sequencing Elucidates Novel Gene Involvement in Susceptibility and Resistance to Phages T4 and T7 in O157. MBio. 2018;9. doi:10.1128/mBio.00705-18
58. Chatterjee A, Willett JLE, Nguyen UT, Monogue B, Palmer KL, Dunny GM, et al. Parallel genomics uncover novel enterococcal-bacteriophage interactions. doi:10.1101/858506
59. Pickard D, Kingsley RA, Hale C, Turner K, Sivaraman K, Wetter M, et al. A genomewide mutagenesis screen identifies multiple genes contributing to Vi capsular expression in Salmonella enterica serovar Typhi. J Bacteriol. 2013;195: 1320–1326.
60. Shin H, Lee J-H, Kim H, Choi Y, Heu S, Ryu S. Receptor diversity and host interaction of bacteriophages infecting Salmonella enterica serovar Typhimurium. PLoS One. 2012;7: e43392.
61. Filippov AA, Sergueev KV, He Y, Huang X-Z, Gnade BT, Mueller AJ, et al. Bacteriophage-resistant mutants in Yersinia pestis: identification of phage receptors and attenuation for mice. PLoS One. 2011;6: e25486.
62. Piya D, Lessor L, Koehler B, Stonecipher A, Cahill J, Gill JJ. Genome-wide Screens Reveal Escherichia Coli Genes Required for Growth of T1-like Phage LL5 and rV5-like Phage LL12. doi:10.20944/preprints201809.0340.v1
63. Wetmore KM, Price MN, Waters RJ, Lamson JS, He J, Hoover CA, et al. Rapid quantification of mutant fitness in diverse bacteria by sequencing randomly bar-coded transposons. MBio. 2015;6: e00306–15.
64. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell. 2013;152: 1173–1183.
65. Mutalik VK, Novichkov PS, Price MN, Owens TK, Callaghan M, Carim S, et al. Dual-barcoded shotgun expression library sequencing for high-throughput characterization of functional traits in bacteria. Nat Commun. 2019;10: 308.
66. Smith AM, Heisler LE, Mellor J, Kaper F, Thompson MJ, Chee M, et al. Quantitative phenotyping via deep barcode sequencing. Genome Research. 2009. pp. 1836–1842. doi:10.1101/gr.093955.109
67. Price MN, Wetmore KM, Jordan Waters R, Callaghan M, Ray J, Liu H, et al. Mutant phenotypes for thousands of bacterial genes of unknown function. Nature. 2018. pp. 503–509. doi:10.1038/s41586-018-0124-0
68. Peters JM, Koo B-M, Patino R, Heussler GE, Hearne CC, Qu J, et al. Enabling genetic analysis of diverse bacteria with Mobile-CRISPRi. Nat Microbiol. 2019;4: 244–250.
69. Rock JM, Hopkins FF, Chavez A, Diallo M, Chase MR, Gerrick ER, et al. Programmable transcriptional repression in mycobacteria using an orthogonal CRISPR interference platform. Nature Microbiology. 2017. doi:10.1038/nmicrobiol.2016.274
70. Liu X, Gallay C, Kjos M, Domenech A, Slager J, Kessel SP, et al. High□throughput

CRISPRi phenotyping identifies new essential genes in Streptococcus pneumoniae. Molecular Systems Biology. 2017. p. 931. doi:10.15252/msb.20167449
71. Cui L, Vigouroux A, Rousset F, Varet H, Khanna V, Bikard D. A CRISPRi screen in E. coli reveals sequence-specific toxicity of dCas9. Nature Communications. 2018. doi:10.1038/s41467-018-04209-5
72. Wang T, Guan C, Guo J, Liu B, Wu Y, Xie Z, et al. Pooled CRISPR interference screening enables genome-scale functional genomics study in bacteria with superior performance. Nat Commun. 2018;9: 2475.
73. Hogan AM, Zisanur AS, Lightly TJ, Cardona ST. A Broad-Host-Range CRISPRi Toolkit for Silencing Gene Expression in Burkholderia. ACS Synthetic Biology. 2019. pp. 2372–2384. doi:10.1021/acssynbio.9b00232
74. Müh U, Pannullo AG, Weiss DS, Ellermeier CD. A Xylose-Inducible Expression System and a CRISPR Interference Plasmid for Targeted Knockdown of Gene Expression in Clostridioides difficile. Journal of Bacteriology. 2019. doi:10.1128/jb.00711-18
75. Qu J, Prasad NK, Yu MA, Chen S, Lyden A, Herrera N, et al. Modulating Pathogenesis with Mobile-CRISPRi. J Bacteriol. 2019;201. doi:10.1128/JB.00304-19
76. Guzzo M, Castro LK, Reisch CR, Guo MS, Laub MT. A CRISPR Interference System for Efficient and Rapid Gene Knockdown in Caulobacter crescentus. MBio. 2020;11. doi:10.1128/mBio.02415-19
77. Häuser R, Blasche S, Dokland T, Haggård-Ljungquist E, von Brunn A, Salas M, et al. Bacteriophage protein-protein interactions. Adv Virus Res. 2012;83: 219–298.
78. Blount ZD. The unexhausted potential of E. coli. Elife. 2015;4: e05826.
79. Zimmer C. Microcosm: E. coli and the New Science of Life. Vintage; 2008.
80. Ghatak S, King ZA, Sastry A, Palsson BO. The y-ome defines the 35% of Escherichia coli genes that lack experimental evidence of function. Nucleic Acids Res. 2019;47: 2446–2454.
81. Sarker SA, Sultana S, Reuteler G, Moine D, Descombes P, Charton F, et al. Oral Phage Therapy of Acute Bacterial Diarrhea With Two Coliphage Preparations: A Randomized Trial in Children From Bangladesh. EBioMedicine. 2016;4: 124–137.
82. Brüssow H. Hurdles for Phage Therapy to Become a Reality—An Editorial Comment. Viruses. 2019. p. 557. doi:10.3390/v11060557
83. Kaper JB, Nataro JP, Mobley HL. Pathogenic Escherichia coli. Nat Rev Microbiol. 2004;2: 123–140.
84. Croxen MA, Brett Finlay B. Molecular mechanisms of Escherichia coli pathogenicity. Nature Reviews Microbiology. 2010. pp. 26–38. doi:10.1038/nrmicro2265
85. Smith H. Virulence determinants of Escherichia coli: present knowledge and questions. Canadian Journal of Microbiology. 1992. pp. 747–752. doi:10.1139/m92-121
86. Calendar R. The Bacteriophages. Oxford University Press on Demand; 2006.
87. Lenski RE. Dynamics of Interactions between Bacteria and Virulent Bacteriophage. Advances in Microbial Ecology. 1988. pp. 1–44. doi:10.1007/978-1-4684-5409-3_1
88. Bertani G. Lysogeny at mid-twentieth century: P1, P2, and other experimental systems. J Bacteriol. 2004;186: 595–600.
89. Kortright KE, Chan BK, Koff JL, Turner PE. Phage Therapy: A Renewed Approach to Combat Antibiotic-Resistant Bacteria. Cell Host Microbe. 2019;25: 219–232.
90. Schwartz M. Interaction of Phages with their Receptor Proteins. Virus Receptors. 1980. pp. 59–94. doi:10.1007/978-94-011-6918-9_4
91. Lindberg AA. Bacteriophage receptors. Annu Rev Microbiol. 1973;27: 205–241.
92. Bertozzi Silva J, Storms Z, Sauvageau D. Host receptors for bacteriophage adsorption. FEMS Microbiol Lett. 2016;363. doi:10.1093/femsle/fnw002

93. Wright A, McConnell M, Kanegasaki S. Lipopolysaccharide as a Bacteriophage Receptor. Virus Receptors. 1980. pp. 27–57. doi:10.1007/978-94-011-6918-9_3
94. Heller KJ. Molecular interaction between bacteriophage and the gram-negative cell envelope. Archives of Microbiology. 1992. pp. 235–248. doi:10.1007/bf00245239
95. Kiino DR, Rothman-Denes LB. Genetic analysis of bacteriophage N4 adsorption. J Bacteriol. 1989;171: 4595–4602.
96. Letarov AV, Kulikov EE. Adsorption of Bacteriophages on Bacterial Cells. Biochemistry . 2017;82: 1632–1658.
97. Raya RR, Oot RA, Moore-Maley B, Wieland S, Callaway TR, Kutter EM, et al. Naturally resident and exogenously applied T4-like and T5-like bacteriophages can reduce Escherichia coli O157:H7 levels in sheep guts. Bacteriophage. 2011;1: 15–24.
98. Kutter E, Gachechiladze K, Poglazov A, Marusich E, Shneider M, Aronsson P, et al. Evolution of T4-related phages. Virus Genes. 1995;11: 285–297.
99. Franklin NC. Mutation in gal U gene of E. coli blocks phage P1 infection. Virology. 1969. pp. 189–191. doi:10.1016/0042-6822(69)90144-5
100. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, et al. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol. 2006;2: 2006.0008.
101. Wall E, Majdalani N, Gottesman S. The Complex Rcs Regulatory Cascade. Annu Rev Microbiol. 2018;72: 111–139.
102. Chamberlin M. Isolation and characterization of prototrophic mutants of Escherichia coli unable to support the intracellular growth of T7. J Virol. 1974;14: 509–516.
103. Friedman DI, Olson ER, Georgopoulos C, Tilly K, Herskowitz I, Banuett F. Interactions of bacteriophage and host macromolecules in the growth of bacteriophage lambda. Microbiol Rev. 1984;48: 299–325.
104. Friedman DI. Interaction between bacteriophage λ and its Escherichia coli host. Current Opinion in Genetics & Development. 1992. pp. 727–738. doi:10.1016/s0959-437x(05)80133-9
105. Polissi A, Goffin L, Georgopoulos C. The Escherichia coli heat shock response and bacteriophage lambda development. FEMS Microbiol Rev. 1995;17: 159–169.
106. Karow M, Raina S, Georgopoulos C, Fayet O. Complex phenotypes of null mutations in the htr genes, whose products are essential for Escherichia coli growth at elevated temperatures. Res Microbiol. 1991;142: 289–294.
107. Pagnout C, Sohm B, Razafitianamaharavo A, Caillet C, Offroy M, Leduc M, et al. Pleiotropic effects of rfa-gene mutations on Escherichia coli envelope properties. Sci Rep. 2019;9: 9696.
108. Klein G, Raina S. Regulated Assembly of LPS, Its Structural Alterations and Cellular Response to LPS Defects. Int J Mol Sci. 2019;20. doi:10.3390/ijms20020356
109. Mitchell AM, Silhavy TJ. Envelope stress responses: balancing damage repair and toxicity. Nat Rev Microbiol. 2019;17: 417–428.
110. Meaden S, Paszkiewicz K, Koskella B. The cost of phage resistance in a plant pathogenic bacterium is context□dependent. Evolution. 2015. pp. 1321–1328. doi:10.1111/evo.12652
111. Díaz-Muñoz SL, Koskella B. Bacteria–Phage Interactions in Natural Environments. Advances in Applied Microbiology. 2014. pp. 135–183. doi:10.1016/b978-0-12-800259-9.00004-4
112. Mirzaei MK, Maurice CF. Ménage à trois in the human gut: interactions between host, bacteria and phages. Nature Reviews Microbiology. 2017. pp. 397–408. doi:10.1038/nrmicro.2017.30

113. Padfield D, Castledine M, Buckling A. Temperature-dependent changes to host-parasite interactions alter the thermal performance of a bacterial host. ISME J. 2020;14: 389–398.
114. Alseth EO, Pursey E, Luján AM, McLeod I, Rollie C, Westra ER. Bacterial biodiversity drives the evolution of CRISPR-based phage resistance. Nature. 2019;574: 549–552.
115. Shibuya I. Metabolic regulations and biological functions of phospholipids in Escherichia coli. Prog Lipid Res. 1992;31: 245–299.
116. Dalebroux ZD, Edrozo MB, Pfuetzner RA, Ressl S, Kulasekara BR, Blanc M-P, et al. Delivery of cardiolipins to the Salmonella outer membrane is necessary for survival within host tissues and virulence. Cell Host Microbe. 2015;17: 441–451.
117. Nepper JF, Lin YC, Weibel DB. Rcs phosphorelay activation in cardiolipin-deficient Escherichia coli reduces biofilm formation. bioRxiv. 2019. p. 522219. doi:10.1101/522219
118. Zhou Y, Filter JJ, Court DL, Gottesman ME, Friedman DI. Requirement for NusG for transcription antitermination in vivo by the lambda N protein. J Bacteriol. 2002;184: 3416–3418.
119. Sullivan SL, Gottesman ME. Requirement for E. coli NusG protein in factor-dependent transcription termination. Cell. 1992;68: 989–994.
120. Henthorn KS. Multilevel regulation of gene expression in lambda and related bacteriophages. 1994.
121. Bertani B, Ruiz N. Function and Biogenesis of Lipopolysaccharides. EcoSal Plus. 2018;8. doi:10.1128/ecosalplus.ESP-0001-2018
122. Konovalova A, Kahne DE, Silhavy TJ. Outer Membrane Biogenesis. Annu Rev Microbiol. 2017;71: 539–556.
123. Potts AH, Vakulskas CA, Pannuri A, Yakhnin H, Babitzke P, Romeo T. Global role of the bacterial post-transcriptional regulator CsrA revealed by integrated transcriptomics. Nat Commun. 2017;8: 1596.
124. Romeo T, Babitzke P. Global Regulation by CsrA and Its RNA Antagonists. Microbiol Spectr. 2018;6. doi:10.1128/microbiolspec.RWR-0009-2017
125. Mojardín L, Salas M. Global Transcriptional Analysis of Virus-Host Interactions between Phage ϕ29 and Bacillus subtilis. Journal of Virology. 2016. pp. 9293–9304. doi:10.1128/jvi.01245-16
126. Chevallereau A, Blasdel BG, De Smet J, Monot M, Zimmermann M, Kogadeeva M, et al. Next-Generation "-omics" Approaches Reveal a Massive Alteration of Host RNA Metabolism during Bacteriophage Infection of Pseudomonas aeruginosa. PLoS Genet. 2016;12: e1006134.
127. Hantke K. Major outer membrane proteins of E. coli K12 serve as receptors for the phages T2 (protein Ia) and 434 (protein Ib). Molecular and General Genetics MGG. 1978. pp. 131–135. doi:10.1007/bf00267377
128. Grigonyte AM, Harrison C, MacDonald PR, Montero-Blay A, Tridgett M, Duncan J, et al. Comparison of CRISPR and Marker-Based Methods for the Engineering of Phage T7. Viruses. 2020;12. doi:10.3390/v12020193
129. Prelich G. Gene overexpression: uses, mechanisms, and interpretation. Genetics. 2012;190: 841–854.
130. Veitia RA, Bottani S, Birchler JA. Gene dosage effects: nonlinearities, genetic interactions, and dosage compensation. Trends Genet. 2013;29: 385–393.
131. van Leeuwen J, Pons C, Boone C, Andrews BJ. Mechanisms of suppression: The wiring of genetic resilience. Bioessays. 2017;39. doi:10.1002/bies.201700042
132. Holzmayer TA, Pestov DG, Roninson IB. Isolation of dominant negative mutants and inhibitory antisense RNA sequences by expression selection of random DNA fragments. Nucleic Acids Res. 1992;20: 711–717.

133. Sopko R, Huang D, Preston N, Chua G, Papp B, Kafadar K, et al. Mapping pathways and phenotypes by systematic gene overexpression. Mol Cell. 2006;21: 319–330.
134. Cardarelli L, Maxwell KL, Davidson AR. Assembly mechanism is the key determinant of the dosage sensitivity of a phage structural protein. Proc Natl Acad Sci U S A. 2011;108: 10168–10173.
135. Plumbridge J. Control of the expression of the manXYZ operon in Escherichia coli: Mlc is a negative regulator of the mannose PTS. Mol Microbiol. 1998;27: 369–380.
136. Decker K, Plumbridge J, Boos W. Negative transcriptional regulation of a positive regulator: the expression of malT, encoding the transcriptional activator of the maltose regulon of Escherichia coli, is negatively controlled by Mlc. Mol Microbiol. 1998;27: 381–390.
137. Joly N, Danot O, Schlegel A, Boos W, Richet E. The Aes protein directly controls the activity of MalT, the central transcriptional activator of the Escherichia coli maltose regulon. J Biol Chem. 2002;277: 16606–16613.
138. Lee SJ, Boos W, Bouché JP, Plumbridge J. Signal transduction between a membrane-bound transporter, PtsG, and a soluble transcription factor, Mlc, of Escherichia coli. EMBO J. 2000;19: 5353–5361.
139. Schlegel A, Danot O, Richet E, Ferenci T, Boos W. The N Terminus of the Escherichia coli Transcription Activator MalT Is the Domain of Interaction with MalY. Journal of Bacteriology. 2002. pp. 3069–3077. doi:10.1128/jb.184.11.3069-3077.2002
140. Reidl J, Boos W. The malX malY operon of Escherichia coli encodes a novel enzyme II of the phosphotransferase system recognizing glucose and maltose and an enzyme abolishing the endogenous induction of the maltose system. J Bacteriol. 1991;173: 4862–4876.
141. Lengsfeld C, Schönert S, Dippel R, Boos W. Glucose- and glucokinase-controlled mal gene expression in Escherichia coli. J Bacteriol. 2009;191: 701–712.
142. Hong JS, Smith GR, Ames BN. Adenosine 3':5'-cyclic monophosphate concentration in the bacterial host regulates the viral decision between lysogeny and lysis. Proc Natl Acad Sci U S A. 1971;68: 2258–2262.
143. Grodzicker T, Arditti RR, Eisen H. Establishment of repression by lambdoid phage in catabolite activator protein and adenylate cyclase mutants of Escherichia coli. Proc Natl Acad Sci U S A. 1972;69: 366–370.
144. Jenal U, Reinders A, Lori C. Cyclic di-GMP: second messenger extraordinaire. Nat Rev Microbiol. 2017;15: 271–284.
145. Römling U, Galperin MY, Gomelsky M. Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. Microbiol Mol Biol Rev. 2013;77: 1–52.
146. Hengge R. Principles of c-di-GMP signalling in bacteria. Nat Rev Microbiol. 2009;7: 263–273.
147. Hall BG. The rtn gene of Proteus vulgaris is actually from Escherichia coli. Journal of Bacteriology. 1997. pp. 2433–2434. doi:10.1128/jb.179.7.2433-2434.1997
148. Chae KS, Yoo OJ. Cloning of the lambda resistant genes from Brevibacterium albidum and Proteus vulgaris into Escherichia coli. Biochem Biophys Res Commun. 1986;140: 1101–1105.
149. Norioka S, Ramakrishnan G, Ikenaka K, Inouye M. Interaction of a transcriptional activator, OmpR, with reciprocally osmoregulated genes, ompF and ompC, of Escherichia coli. J Biol Chem. 1986;261: 17113–17119.
150. Yoshida T, Qin L, Egger LA, Inouye M. Transcription regulation of ompF and ompC by a single transcription factor, OmpR. J Biol Chem. 2006;281: 17114–17123.
151. Mizuno T, Mizushima S. Signal transduction and gene regulation through the

phosphorylation of two regulatory components: the molecular basis for the osmotic regulation of the porin genes. Mol Microbiol. 1990;4: 1077–1082.
152. Andersen J, Forst SA, Zhao K, Inouye M, Delihas N. The function of micF RNA. micF RNA is a major factor in the thermal regulation of OmpF protein in Escherichia coli. J Biol Chem. 1989;264: 17961–17970.
153. Deighan P, Free A, Dorman CJ. A role for the Escherichia coli H-NS-like protein StpA in OmpF porin expression through modulation of micF RNA stability. Mol Microbiol. 2000;38: 126–139.
154. Kao C, Snyder L. The lit gene product which blocks bacteriophage T4 late gene expression is a membrane protein encoded by a cryptic DNA element, e14. J Bacteriol. 1988;170: 2056–2062.
155. Kitagawa M, Ara T, Arifuzzaman M, Ioka-Nakamichi T, Inamoto E, Toyonaga H, et al. Complete set of ORF clones of Escherichia coli ASKA library (a complete set of E. coli K-12 ORF archive): unique resources for biological research. DNA Res. 2005;12: 291–299.
156. Cano DA, Domínguez-Bernal G, Tierrez A, Garcia-Del Portillo F, Casadesús J. Regulation of capsule synthesis and cell motility in Salmonella enterica by the essential gene igaA. Genetics. 2002;162: 1513–1523.
157. Cho S-H, Szewczyk J, Pesavento C, Zietek M, Banzhaf M, Roszczenko P, et al. Detecting envelope stress by monitoring β-barrel assembly. Cell. 2014;159: 1652–1664.
158. Stevenson G, Andrianopoulos K, Hobbs M, Reeves PR. Organization of the Escherichia coli K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid. Journal of bacteriology. 1996. pp. 4885–4893. doi:10.1128/jb.178.16.4885-4893.1996
159. Hussein NA, Cho S-H, Laloux G, Siam R, Collet J-F. Distinct domains of Escherichia coli IgaA connect envelope stress sensing and down-regulation of the Rcs phosphorelay across subcellular compartments. PLoS Genet. 2018;14: e1007398.
160. Majdalani N, Gottesman S. The Rcs phosphorelay: a complex signal transduction system. Annu Rev Microbiol. 2005;59: 379–405.
161. Chaudhry W, Lee E, Worthy A, Weiss Z, Grabowicz M, Vega N, et al. Mucoidy, a general mechanism for maintaining lytic phage in populations of bacteria. doi:10.1101/775056
162. Scanlan PD, Buckling A. Co-evolution with lytic phage selects for the mucoid phenotype of Pseudomonas fluorescens SBW25. ISME J. 2012;6: 1148–1158.
163. Martin DR. Mucoid Variation in Pseudomonas AeruginosaInduced by the Action of Phage. Journal of Medical Microbiology. 1973. pp. 111–118. doi:10.1099/00222615-6-1-111
164. Mizoguchi K, Morita M, Fischer CR, Yoichi M, Tanji Y, Unno H. Coevolution of bacteriophage PP01 and Escherichia coli O157:H7 in continuous culture. Appl Environ Microbiol. 2003;69: 170–176.
165. Keseler IM, Mackie A, Santos-Zavaleta A, Billington R, Bonavides-Martínez C, Caspi R, et al. The EcoCyc database: reflecting new knowledge about Escherichia coli K-12. Nucleic Acids Res. 2017;45: D543–D550.
166. Reinders A, Hee C-S, Ozaki S, Mazur A, Boehm A, Schirmer T, et al. Expression and Genetic Activation of Cyclic Di-GMP-Specific Phosphodiesterases in Escherichia coli. J Bacteriol. 2016;198: 448–462.
167. Hengge R, Galperin MY, Ghigo J-M, Gomelsky M, Green J, Hughes KT, et al. Systematic Nomenclature for GGDEF and EAL Domain-Containing Cyclic Di-GMP Turnover Proteins of Escherichia coli. J Bacteriol. 2016;198: 7–11.
168. Povolotsky TL, Hengge R. Genome-Based Comparison of Cyclic Di-GMP Signaling in

Pathogenic and Commensal Escherichia coli Strains. J Bacteriol. 2016;198: 111–126.
169. Park JH, Lee KH, Kim TY, Lee SY. Metabolic engineering of Escherichia coli for the production of L-valine based on transcriptome analysis and in silico gene knockout simulation. Proc Natl Acad Sci U S A. 2007;104: 7797–7802.
170. Lehnen D, Blumer C, Polen T, Wackwitz B, Wendisch VF, Unden G. LrhA as a new transcriptional key regulator of flagella, motility and chemotaxis genes in Escherichia coli. Mol Microbiol. 2002;45: 521–532.
171. Göpel Y, Khan MA, Görke B. Ménage à trois: post-transcriptional control of the key enzyme for cell envelope synthesis by a base-pairing small RNA, an RNase adaptor protein, and a small RNA mimic. RNA Biol. 2014;11: 433–442.
172. Studier FW, Daegelen P, Lenski RE, Maslov S, Kim JF. Understanding the differences between genome sequences of Escherichia coli B strains REL606 and BL21(DE3) and comparison of the E. coli B and K-12 genomes. J Mol Biol. 2009;394: 653–680.
173. Jeong H, Barbe V, Lee CH, Vallenet D, Yu DS, Choi S-H, et al. Genome sequences of Escherichia coli B strains REL606 and BL21(DE3). J Mol Biol. 2009;394: 644–652.
174. Daegelen P, Studier FW, Lenski RE, Cure S, Kim JF. Tracing ancestors and relatives of Escherichia coli B, and the derivation of B strains REL606 and BL21(DE3). J Mol Biol. 2009;394: 634–643.
175. McPartland J, Rothman-Denes LB. The tail sheath of bacteriophage N4 interacts with the Escherichia coli receptor. J Bacteriol. 2009;191: 525–532.
176. Washizaki A, Yonesaki T, Otsuka Y. Characterization of the interactions between Escherichia coli receptors, LPS and OmpC, and bacteriophage T4 long tail fibers. Microbiologyopen. 2016;5: 1003–1015.
177. Hattman S, Fukasawa T. HOST-INDUCED MODIFICATION OF T-EVEN PHAGES DUE TO DEFECTIVE GLUCOSYLATION OF THEIR DNA. Proc Natl Acad Sci U S A. 1963;50: 297–300.
178. Prehm P, Jann B, Jann K, Schmidt G, Stirm S. On a bacteriophage T3 and T4 receptor region within the cell wall lipopolysaccharide of escherichia coli B. J Mol Biol. 1976;101: 277–281.
179. Lenski RE. Two-step resistance by Escherichia coli B to bacteriophage T2. Genetics. 1984;107: 1–7.
180. Boos W, Shuman H. Maltose/maltodextrin system of Escherichia coli: transport, metabolism, and regulation. Microbiol Mol Biol Rev. 1998;62: 204–229.
181. Plumbridge J. Control of the expression of the manXYZ operon in Escherichia coli: Mlc is a negative regulator of the mannose PTS. Mol Microbiol. 1998;27: 369–380.
182. Decker K, Plumbridge J, Boos W. Negative transcriptional regulation of a positive regulator: the expression of malT, encoding the transcriptional activator of the maltose regulon of Escherichia coli, is negatively controlled by Mlc. Mol Microbiol. 1998;27: 381–390.
183. Lee SJ, Boos W, Bouché JP, Plumbridge J. Signal transduction between a membrane-bound transporter, PtsG, and a soluble transcription factor, Mlc, of Escherichia coli. EMBO J. 2000;19: 5353–5361.
184. Lee SJ, Boos W, Bouché JP, Plumbridge J. Signal transduction between a membrane-bound transporter, PtsG, and a soluble transcription factor, Mlc, of Escherichia coli. EMBO J. 2000;19: 5353–5361.
185. Danese PN, Pratt LA, Kolter R. Exopolysaccharide production is required for development of Escherichia coli K-12 biofilm architecture. J Bacteriol. 2000;182: 3593–3596.
186. Eydallin G, Morán-Zorzano MT, Muñoz FJ, Baroja-Fernández E, Montero M, Alonso-

Casajús N, et al. An Escherichia coli mutant producing a truncated inactive form of GlgC synthesizes glycogen: further evidences for the occurrence of various important sources of ADPglucose in enterobacteria. FEBS Lett. 2007;581: 4417–4422.
187. Baker CS, Morozov I, Suzuki K, Romeo T, Babitzke P. CsrA regulates glycogen biosynthesis by preventing translation of glgC in Escherichia coli. Mol Microbiol. 2002;44: 1599–1610.
188. Romeo T, Gong M, Liu MY, Brun-Zinkernagel AM. Identification and molecular characterization of csrA, a pleiotropic gene from Escherichia coli that affects glycogen biosynthesis, gluconeogenesis, cell size, and surface properties. J Bacteriol. 1993;175: 4744–4755.
189. Wang X, Dubey AK, Suzuki K, Baker CS, Babitzke P, Romeo T. CsrA post-transcriptionally represses pgaABCD, responsible for synthesis of a biofilm polysaccharide adhesin of Escherichia coli. Mol Microbiol. 2005;56: 1648–1663.
190. Jackson DW, Suzuki K, Oakford L, Simecka JW, Hart ME, Romeo T. Biofilm formation and dispersal under the influence of the global regulator CsrA of Escherichia coli. J Bacteriol. 2002;184: 290–301.
191. Palaniyandi S, Mitra A, Herren CD, Lockatell CV, Johnson DE, Zhu X, et al. BarA-UvrY two-component system regulates virulence of uropathogenic E. coli CFT073. PLoS One. 2012;7: e31348.
192. Klein G, Raina S. Regulated Control of the Assembly and Diversity of LPS by Noncoding sRNAs. Biomed Res Int. 2015;2015: 153561.
193. Murata A, Odaka M, Mukuno S. The Bacteriophage-Inactivating Effect of Basic Amino Acids; Arginine, Histidine, and Lysine. Agric Biol Chem. 1974;38: 477–478.
194. Lau CKY, Krewulak KD, Vogel HJ. Bacterial ferrous iron transport: the Feo system. FEMS Microbiol Rev. 2016;40: 273–298.
195. Kammler M, Schön C, Hantke K. Characterization of the ferrous iron uptake system of Escherichia coli. J Bacteriol. 1993;175: 6212–6219.
196. Hantke K. Regulation of ferric iron transport in Escherichia coli K12: isolation of a constitutive mutant. Mol Gen Genet. 1981;182: 288–292.
197. Carpenter BM, Whitmire JM, Merrell DS. This Is Not Your Mother's Repressor: the Complex Role of Fur in Pathogenesis. Infection and Immunity. 2009. pp. 2590–2601. doi:10.1128/iai.00116-09
198. Braun V. FhuA (TonA), the career of a protein. J Bacteriol. 2009;191: 3431–3436.
199. Heinrichs DE, Yethon JA, Whitfield C. Molecular basis for structural diversity in the core regions of the lipopolysaccharides of Escherichia coli and Salmonella enterica. Mol Microbiol. 1998;30: 221–232.
200. Ornellas EP, Stocker BA. Relation of lipopolysaccharide character to P1 sensitivity in Salmonella typhimurium. Virology. 1974;60: 491–502.
201. Ho TD, Waldor MK. Enterohemorrhagic Escherichia coli O157:H7 gal mutants are sensitive to bacteriophage P1 and defective in intestinal colonization. Infect Immun. 2007;75: 1661–1666.
202. Raetz CRH, Whitfield C. Lipopolysaccharide Endotoxins. Annual Review of Biochemistry. 2002. pp. 635–700. doi:10.1146/annurev.biochem.71.110601.135414
203. Nikaido H. Molecular basis of bacterial outer membrane permeability revisited. Microbiol Mol Biol Rev. 2003;67: 593–656.
204. Oechslin F. Resistance Development to Bacteriophages Occurring during Bacteriophage Therapy. Viruses. 2018;10. doi:10.3390/v10070351
205. Hernandez CA, Koskella B. Phage resistance evolution in vitro is not reflective of in vivo outcome in a plant-bacteria-phage system. Evolution. 2019;73: 2461–2475.

206. Abedon ST. Bacteriophage Ecology: Population Growth, Evolution, and Impact of Bacterial Viruses. Cambridge University Press; 2008.
207. Díaz-Muñoz SL, Koskella B. Bacteria-phage interactions in natural environments. Adv Appl Microbiol. 2014;89: 135–183.
208. Oechslin F, Piccardi P, Mancini S, Gabard J, Moreillon P, Entenza JM, et al. Synergistic Interaction Between Phage Therapy and Antibiotics Clears Pseudomonas Aeruginosa Infection in Endocarditis and Reduces Virulence. J Infect Dis. 2017;215: 703–712.
209. Hung C-H, Kuo C-F, Wang C-H, Wu C-M, Tsao N. Experimental phage therapy in treating Klebsiella pneumoniae-mediated liver abscesses and bacteremia in mice. Antimicrob Agents Chemother. 2011;55: 1358–1365.
210. Taylor VL, Fitzpatrick AD, Islam Z, Maxwell KL. The Diverse Impacts of Phage Morons on Bacterial Fitness and Virulence. Adv Virus Res. 2019;103: 1–31.
211. Koskella B, Brockhurst MA. Bacteria-phage coevolution as a driver of ecological and evolutionary processes in microbial communities. FEMS Microbiol Rev. 2014;38: 916–931.
212. Castillo D, Rørbo N, Jørgensen J, Lange J, Tan D, Kalatzis PG, et al. Phage defense mechanisms and their genomic and phenotypic implications in the fish pathogen Vibrio anguillarum. FEMS Microbiol Ecol. 2019;95. doi:10.1093/femsec/fiz004
213. Wang X, Wei Z, Yang K, Wang J, Jousset A, Xu Y, et al. Phage combination therapies for bacterial wilt disease in tomato. Nat Biotechnol. 2019;37: 1513–1520.
214. McCallin S, Oechslin F. Bacterial Resistance to Phage and Its Impact on Clinical Therapy. Phage Therapy: A Practical Approach. 2019. pp. 59–88. doi:10.1007/978-3-030-26736-0_3
215. Pires DP, Cleto S, Sillankorva S, Azeredo J, Lu TK. Genetically Engineered Phages: a Review of Advances over the Last Decade. Microbiol Mol Biol Rev. 2016;80: 523–543.
216. Kilcher S, Loessner MJ. Engineering Bacteriophages as Versatile Biologics. Trends Microbiol. 2019;27: 355–367.
217. Ando H, Lemire S, Pires DP, Lu TK. Engineering Modular Viral Scaffolds for Targeted Bacterial Population Editing. Cell Syst. 2015;1: 187–196.
218. Chan BK, Abedon ST, Loc-Carrillo C. Phage cocktails and the future of phage therapy. Future Microbiol. 2013;8: 769–783.
219. Yehl K, Lemire S, Yang AC, Ando H, Mimee M, Torres MDT, et al. Engineering Phage Host-Range and Suppressing Bacterial Resistance through Phage Tail Fiber Mutagenesis. Cell. 2019;179: 459–469.e9.
220. Yosef I, Goren MG, Globus R, Molshanski-Mor S, Qimron U. Extending the Host Range of Bacteriophage Particles for DNA Transduction. Mol Cell. 2017;66: 721–728.e3.
221. Yosef I, Manor M, Kiro R, Qimron U. Temperate and lytic bacteriophages programmed to sensitize and kill antibiotic-resistant bacteria. Proc Natl Acad Sci U S A. 2015;112: 7267–7272.
222. Wright RCT, Friman V-P, Smith MCM, Brockhurst MA. Resistance Evolution against Phage Combinations Depends on the Timing and Order of Exposure. MBio. 2019;10. doi:10.1128/mBio.01652-19
223. Kilcher S, Studer P, Muessner C, Klumpp J, Loessner MJ. Cross-genus rebooting of custom-made, synthetic bacteriophage genomes in L-form bacteria. Proc Natl Acad Sci U S A. 2018;115: 567–572.
224. Kutter E, Sulakvelidze A. Bacteriophages: Biology and Applications. CRC Press; 2004.
225. Dunne M, Rupf B, Tala M, Qabrati X, Ernst P, Shen Y, et al. Reprogramming Bacteriophage Host Range through Structure-Guided Design of Chimeric Receptor Binding Proteins. Cell Rep. 2019;29: 1336–1350.e4.

226. Nobrega FL, Costa AR, Kluskens LD, Azeredo J. Revisiting phage therapy: new applications for old resources. Trends in Microbiology. 2015. pp. 185–191. doi:10.1016/j.tim.2015.01.006
227. Sturino JM, Klaenhammer TR. Engineered bacteriophage-defence systems in bioprocessing. Nat Rev Microbiol. 2006;4: 395–404.
228. Mahony J, Ainsworth S, Stockdale S, van Sinderen D. Phages of lactic acid bacteria: The role of genetics in understanding phage-host interactions and their co-evolutionary processes. Virology. 2012. pp. 143–150. doi:10.1016/j.virol.2012.10.008
229. Ma NJ, Isaacs FJ. Genomic Recoding Broadly Obstructs the Propagation of Horizontally Transferred Genetic Elements. Cell Syst. 2016;3: 199–207.
230. Lajoie MJ, Rovner AJ, Goodman DB, Aerni H-R, Haimovich AD, Kuznetsov G, et al. Genomically recoded organisms expand biological functions. Science. 2013;342: 357–360.
231. Spears PA, Mitsu Suyemoto M, Hamrick TS, Wolf RL, Havell EA, Orndorff PE. In VitroProperties of a Listeria monocytogenes Bacteriophage-Resistant Mutant Predict Its Efficacy as a Live Oral Vaccine Strain. Infection and Immunity. 2011. pp. 5001–5009. doi:10.1128/iai.05700-11
232. Capparelli R, Nocerino N, Lanzetta R, Silipo A, Amoresano A, Giangrande C, et al. Bacteriophage-resistant Staphylococcus aureus mutant confers broad immunity against staphylococcal infection in mice. PLoS One. 2010;5: e11720.
233. Capparelli R, Nocerino N, Iannaccone M, Ercolini D, Parlato M, Chiara M, et al. Bacteriophage therapy of Salmonella enterica: a fresh appraisal of bacteriophage therapy. J Infect Dis. 2010;201: 52–61.
234. Dąbrowska K, Abedon ST. Pharmacologically Aware Phage Therapy: Pharmacodynamic and Pharmacokinetic Obstacles to Phage Antibacterial Action in Animal and Human Bodies. Microbiol Mol Biol Rev. 2019;83. doi:10.1128/MMBR.00012-19
235. Torres-Barceló C, Hochberg ME. Evolutionary Rationale for Phages as Complements of Antibiotics. Trends Microbiol. 2016;24: 249–256.
236. Tagliaferri TL, Jansen M, Horz H-P. Fighting Pathogenic Bacteria on Two Fronts: Phages and Antibiotics as Combined Strategy. Front Cell Infect Microbiol. 2019;9: 22.
237. Abedon ST. Phage-Antibiotic Combination Treatments: Antagonistic Impacts of Antibiotics on the Pharmacodynamics of Phage Therapy? Antibiotics (Basel). 2019;8. doi:10.3390/antibiotics8040182
238. Bak G, Lee J, Suk S, Kim D, Young Lee J, Kim K-S, et al. Identification of novel sRNAs involved in biofilm formation, motility, and fimbriae formation in Escherichia coli. Sci Rep. 2015;5: 15287.
239. Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, et al. Short Protocols in Molecular Biology: A Compendium of Methods from Current Protocols in Molecular Biology. John Wiley & Sons; 1995.
240. Grenier F, Matteau D, Baby V, Rodrigue S. Complete Genome Sequence of Escherichia coli BW25113. Genome Announc. 2014;2. doi:10.1128/genomeA.01038-14
241. Ptashne M. A Genetic Switch: Phage Lambda Revisited. CSHL Press; 2004.
242. Ikeda H, Tomizawa J-I. Transducing fragments in generalized transduction by phage P1 I. Molecular origin of the fragments. Journal of Molecular Biology. 1966. p. 226. doi:10.1016/s0022-2836(66)80067-0
243. Price MN, Ray J, Iavarone AT, Carlson HK, Ryan EM, Malmstrom RR, et al. Oxidative Pathways of Deoxyribose and Deoxyribonate Catabolism. mSystems. 2019;4. doi:10.1128/mSystems.00297-18

244. McCarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res. 2012;40: 4288–4297.
245. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26: 139–140.
246. Datta S, Costantino N, Court DL. A set of recombineering plasmids for gram-negative bacteria. Gene. 2006;379: 109–115.
247. Arkin AP, Cottingham RW, Henry CS, Harris NL, Stevens RL, Maslov S, et al. KBase: The United States Department of Energy Systems Biology Knowledgebase. Nat Biotechnol. 2018;36: 566–569.
248. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30: 2114–2120.
249. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol. 2019;37: 907–915.
250. Pertea M, Pertea GM, Antonescu CM, Chang T-C, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol. 2015;33: 290–295.
251. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15: 550.