The protein interaction network of bacteriophage lambda with its host, Escherichia coli

Sonja Blasche¹,*, Stefan Wuchty²,*, Seesandra V. Rajagopala³ and Peter Uetz⁴,#

¹ Genomics and Proteomics Core Facilities, German Cancer Research Center, 69120 Heidelberg, Germany
² National Center of Biotechnology Information, National Institutes of Health, Bethesda, MD 20894, USA
³ J Craig Venter Institute, Rockville, MD 20850, USA
⁴ Center for the Study of Biological Complexity, Virginia Commonwealth University, Richmond, VA 23284, USA
⁵ present address: Dept. of Computer Science, University of Miami, Coral Gables, FL 33146, USA
* authors contributed equally to this work
# corresponding author

Peter Uetz
Center for the Study of Biological Complexity
Virginia Commonwealth University
Richmond, VA 23284, USA
Phone: (804) 827-4573
Email: peter@uetz.us

Running title: The protein interactome of E. coli and phage lambda
Abstract
Although most of the 73 open reading frames (ORFs) in bacteriophage λ have been investigated intensively the function of many genes in host-phage interactions remains poorly understood. Using yeast two-hybrid screens of all lambda ORFs for interactions with its host *E. coli*, we determined a raw dataset of 631 host-phage interactions resulting in a set of 62 high-confidence interactions after multiple rounds of retesting. These links suggest novel regulatory interactions between the *E. coli* transcriptional network and lambda proteins. Targeted host proteins and genes required for lambda infection are enriched among highly connected proteins, suggesting that bacteriophages resemble interaction patterns of human viruses. Lambda tail proteins interact with both bacterial fimbrial proteins and *E. coli* proteins homologous to other phage proteins. Lambda appears to dramatically differ from phages like T7 because of its unusually large number of modified and processed proteins which reduces the number of host-virus interactions detectable by yeast two-hybrid screens.

Introduction
More than 60 years ago, Esther Lederberg discovered phage lambda (1). Since this seminal discovery lambda has become a model organism, contributing to our current understanding of the ways genes work and viruses take control of their hosts. However, phage lambda is still far from being completely understood given that the function of several genes in its 48.5 kb genome remain only vaguely defined. For instance, 14 out of 73 predicted lambda proteins are still largely uncharacterized (2).

Some of the best-characterized aspects of lambda biology are the genetic switch that determines whether a phage lyses the cell or integrates into its host genome as a prophage and the mechanism of transcriptional anti-termination (3). Also, lambda continues to provide new insights into its gene regulatory circuits (4, 5) with recent studies of its DNA packaging motor as the vanguard of nanomotor research (6).

Key to lambda biology is a detailed understanding of the phage’s proteins, including their interactions. We recently analyzed 95 protein-protein interactions (PPIs) between ~70 lambda proteins, capturing 16 of 33 previously published interactions (2). While protein interactions are reasonably well characterized within the virus particle, host-
phage interactions remain incompletely known during the replication and recombination stages of the lambda life cycle in the host cell. In addition, the molecular details of virion assembly are still largely mysterious. Altogether, about 30 host-lambda interactions have been discovered in the past 50 years (Table 1).

To improve our understanding of lambda PPIs, we have cloned 68 lambda open reading frames (ORFs) (2) and performed screens against a pooled *E. coli* library using a multi-vector yeast two-hybrid strategy. Out of a total of 631 raw interactions we found 244 reproducible interactions. Furthermore, we confirmed 162 interactions in an additional round of retesting, allowing us to obtain an overlapping set of 62 high-confidence interactions. In our analysis of this set of phage-host interactions we show that lambda preferably targets highly connected *E. coli* proteins, indicating that the phage affects central host proteins. Utilizing a set of genes that are necessary for phage infection in *E. coli*, we elucidated a regulatory subnetwork that suggests new ways the phage interferes with host transcription factors to take control of the cell.

**Materials and Methods**

**Yeast two-hybrid screens**

For the yeast two-hybrid screens, 68 phage lambda ORFs were cloned into the bait vectors pGBK7T7g and pDEST32 (2) and screened against the *E. coli* W3110 library (7) in the prey vectors pGADT7g and pDEST22 (8). As a result we obtained two bait-prey combinations, pGBK7T7g-pGADT7g and pDEST32-pDEST22. The screens were performed in three runs, each including the lambda baits once or multiple times against one library using one screening method (Table 2, Fig. 1).

In the first run, arrayed pool screens were performed using the whole *E. coli* pGADT7g library arranged on two plates where each well contained a different prey pool of 24 clones. Each bait was mated with each prey pool, and the diploids were arrayed on selective agarose plates. A flowchart of the procedure is shown in Fig. 1. In the second and third runs the lambda bait proteins were screened against pooled *E. coli* pGADT7g and pDEST22 libraries, respectively. The diploid yeasts were plated on selective agarose plates, and prey plasmids were identified by colony PCR and sequencing.
The pGBKT7g/pGADT7g and pDEST32/22 vector systems differ mainly by copy number and, as a consequence, in the amount of protein expressed in the yeast cell. As previously shown (9), the number of obtained screening hits varies clearly between the vector systems, and the number of common pairs detected in both systems is usually low. However, the two vector systems complement each other, suggesting that using both led to an increase in identified protein-protein interactions.

Promiscuous baits

Potentially promiscuous baits were identified using two different approaches: (1) Binary tests of all baits against empty prey vector and (2) generation of a “bait promiscuity factor” using pairwise Y2H tests of phage baits against 184 preys that were found in our initial screen. These preys were retested to make sure that they are reproducible. Some of the baits turned out to show nearly random interactions and were therefore discarded.

In the first approach, the lambda baits in pGBKT7g and pDEST32 bait vectors were tested against empty pGADT7g and pDEST22 prey vectors (i.e. without prey protein), respectively. This step resulted in the identification of four random activators (lambdap18 = gpL, lambdap29 = Tfa, lambdap47 = Orf28, and lambdap67 = NinF) in the pGBKT7g/pGADT7g vector system and five random activators (lambdap05 = gpC, lambdap06 = Nu3, lambdap29 = Tfa, lambdap47 = Orf28 and lambdap67 (NinF) yielded positive results (“pseudo-interactions”) with empty vectors in both systems. All baits producing “interactions” with empty prey vector were deemed non-screenable and were thus excluded.

In the second approach, pairwise tests of phage lambda bait proteins against different prey proteins suggested that lambdap16 (gpH), lambdap36 (Ea8.5), lambdap45 (Ea10), lambdap65 (NinD) and lambdap83 (Ea22) were putative random activators. A bait protein was defined as random activator if it binds to more than 30 out of 184 preys that were not identified previously in the phage/host pool screens.

According to their tendency to interact randomly, a promiscuity score ranging from 1 (non-promiscuous) to 3 was assigned to all screenable bait proteins. Baits of categories 2 (lambdap16 = gpH, lambdap36 = Ea8.5, and lambdap83 = Ea22) and 3 (lambdap45 =
Ea10 and lambdap65 = NinD) had ≤45 and ≥60 host interaction partners, respectively. No protein had between 45 and 60 partners.

Promiscuous preys
A prey protein was tagged as potentially promiscuous or ‘sticky’ if it interacted with 5 or more different lambda baits reproducibly and/or if it interacted randomly in pairwise yeast two-hybrid tests. A list of potential promiscuous preys and the number of phage lambda baits they interacted with is provided in Suppl. Table S1.

Most promiscuous preys were found by the pGBKT7g/pGADT7g vector system. The pDEST system revealed only interactions with two sticky proteins, LeuB and YajI, both of which were found as single hits. We acknowledge that the pGBKT7g/pGADT7g system is less stringent than the pDEST system since it also detected most interactions in (10). Prey proteins identified as promiscuous were excluded from further evaluation.

Verification of single hits by pairwise Y2H assays
144 interacting lambda-host pairs were found as multiple hits in our initial screens. This set and selected additional interactions from the initial screen were retested using binary Y2H tests. Out of these 162 were tested positive, yielding an overlap of 62 “high-confidence” interactions (Suppl. Table S1). The retests were performed as previously described in (8). A summary of all screens is provided in Table 3.

Molecular interactions data of E. coli
We utilized 3,559 experimentally determined, physical interactions between 1,987 proteins of E. coli (Rajagopala et al., submitted). As for regulatory interactions, we utilized data from the RegulonDB database (11). In particular, we used 4,442 regulatory interactions between 187 transcription factors and 1,638 genes in E. coli that RegulonDB collected from the literature.

Functional groups
All lambda-phage proteins and corresponding E. coli interactors were assigned to 9 and 14 different functional groups, respectively. As for lambda, these classes included virion
head, virion tail, superinfection exclusion, transcription, replication, recombination, lysis, inhibition of host replication and proteins of unknown function. The categories for *E. coli* involved biosynthesis of cofactors prosthetic groups and carriers, cell envelope, phage origin, cellular processes, DNA metabolism, energy metabolism, fimbrial proteins, protein fate, protein synthesis, regulatory functions, transcription, transport and binding proteins, unknown/uncharacterized/hypothetical proteins and others with known function. Functional annotations of phage lambda were obtained from the reference proteome set in Uniprot (http://www.uniprot.org). Functional groups of *E. coli* were obtained from EcoGene 3.0 (http://ecogene.org).

**Phage host-dependencies and essential genes**

We used a set of 57 genes of *E. coli* that are required for the phage infection as assessed by measuring their effect on lambda replication when knocked out in *E. coli* (12). 705 essential *E. coli* genes were obtained from the Database of Essential Genes (DEG) (13).

**Enrichment analysis**

We grouped proteins of *E. coli* according to their number of interaction partners in an underlying protein interactions network. We represented each group by $N_{\geq k}$ proteins that had at least $k$ interactions. In each group we calculated the number of proteins with a certain feature $i$ such as being targeted by the phage or required for the infection process $N_{i, \geq k}$. Randomly assigning feature $i$ to proteins we defined $E_{i, \geq k} = \frac{N_{i, \geq k}}{N_{\geq k}}$ as the enrichment of proteins with feature $i$ where $N_{i, \geq k}$ was the corresponding random number of proteins with feature $i$ among all $N_{\geq k}$ proteins. After averaging $E_i$ over 10,000 randomizations $E_i > 1$ pointed to an enrichment and *vice versa*, while $E_i \sim 1$ indicated a random process (14).

**Results**

**Interactions of phage lambda with its host: overview**
Over the past decades about 30 interactions between proteins of phage lambda and its host have been identified using a wide range of approaches and studies (15) (Table 1). After our successful study of lambda-lambda interactions (2), we decided to use a similar approach to screen all lambda ORFs against a library of *E. coli* clones using the yeast two-hybrid system. In short, we screened each lambda ORF as Y2H bait against an array of *E. coli* preys as well as a pooled library of prey clones (Fig. 1, Table 2, see methods for details). These screens provided a raw data set of 631 unique bait/prey pairs, involving 294 distinct *E. coli* proteins. 244 of these interactions were found in multiple hits and after removal of promiscuous proteins and repeated retesting we finally obtained a set of 62 “high-confidence” interactions (Table 4, Suppl. Table S1).

As expected (10), different Y2H vectors resulted in very different interactions: 573 raw pairs were found using the pGBK7T7g/pGADT7g vector system while the pDEST32/22 system yielded only 56 interactions. Two were found in both systems (lambda gpG – *E. coli* ClpP and lambda gpA – *E. coli* NohB), and their significance is discussed below. The results of all phage lambda-host screens are summarized in Table 4 and Suppl. Table S1.

**Phage-phage vs. phage-host interactions and genome organization**

In an interaction network of phage proteins, we previously observed that structural proteins of the virion were involved in most interactions (2). By contrast, patterns of interactions of phage proteins with their corresponding host proteins are less clear (Fig. 2A). Most structural proteins do not appear to interact with host proteins, except for gpH and gpG, the tape measure and the tail assembly chaperone, respectively. Furthermore, the proteins encoded by the early left operon and early right operon significantly interact while the structural proteins of the late operon have only few interactions with the host.

**Interactions among functional groups**

The 62 high-confidence phage lambda – host interactions were combined in a complete network along with 31 previously known interactions. A comparison of these two sets indicated an overlap of one interaction. Furthermore, we annotated each host and phage protein with its corresponding function (Fig. 2B). Unexpectedly, we found few clear
patterns, indicating that only a few phage proteins interacted with very specific groups of target proteins. Most notably, the *E. coli* group ‘transcription’ (NusA, RpoABD, RpoS, Fis, ihfAB) was targeted by phage lambda proteins involved in transcription (gpN, cI), replication (xis, gpO, gpP), recombination (int) and lysis (cII), while lambda transcription factors (gpN, gpQ) targeted most functional groups of the host. Similarly, many proteins involved in phage transcription interacted with host proteins involved in energy metabolism. Interestingly, an *E. coli* fimbrial protein, YehD, was exclusively linked to phage lambda proteins that were involved in virion tail formation such as the tape measure protein gpH. Even though host fimbrial proteins do not have any homologs in lambda, the tape measure protein appears to be capable of selectively targeting host fimbrial proteins. We also found that the *E. coli* group ‘protein synthesis’ was targeted mainly by the putative single strand binding protein Ea10, a protein that we observed to bind the ribosomal proteins RpmA and RpsG as well as the ribosome modulation factor Rmf. ‘Transport and binding’ host proteins interacted mainly with virion tail proteins. Given that transport proteins are expected to be transmembrane proteins, these proteins may help the phage to attach to the host cells. Interestingly, gpH appears to be injected with the DNA and thus may interact with intracellular proteins. Finally, proteins of unknown function (or hypothetical proteins) of both host and phage were involved in interactions with several functional categories. Still, most functionally unknown proteins of the host interacted with unknown phage proteins.

**Protein-protein interactions and host factors required for lambda infection**

Maynard et al. (12) screened the *E. coli* KEIO collection (16) and found 57 host proteins that are required for lambda infection and reproduction. However, only 4 of these have been previously known to physically interact with lambda proteins, namely LamB, the lambda receptor, IhfAB, the integration host factor, and HflD, which interacts with the cII protein, catalyzing its degradation and DNA binding ability.

In the combined network of high-confidence and previously known phage lambda – host interactions we found that lambda significantly targeted a total of 5 genes that are necessary for the phage infection (P = 1.4 × 10⁻³, Fisher’s exact test). However, the significance of these interactions is not clear. For instance, HldD is involved in the
synthesis of the lippopolysaccharide (LPS) and is required for lambda to infect its host (12). In our set of 62 high-confidence host-phage interactions, HldD (= RfaD) interacts with gpQ, a phage protein involved in transcription anti-termination (Fig. 3).

**Phage lambda targets essential *E. coli* genes and genes involved in transcription**

It has been suggested that both bacterial and viral proteins tend to interact preferentially with highly connected host proteins, so-called hubs (17). Similarly, such hubs have been proposed to be more likely to be essential for a cell (18). In the combined network we observed a similar significant enrichment among the 22 essential genes in *E. coli* targeted by lambda (P = 4.7 × 10^-4, Fisher’s exact test). Ten *E. coli* genes that were previously known to interact with 8 lambda proteins are essential, while 12 essential host genes were found to interact with 6 phage proteins in our screens (Fig. 3). Furthermore, we observed that lambda targeted 9 transcription factors (1.7 × 10^-3, Fisher’s exact test). In particular, 4 transcription factors interacting with 3 lambda proteins were found in previously published interactions. In turn, our screens detected 5 transcription factors that were targeted by 6 lambda proteins (Fig. 3).

**Interaction of lambda with lambda homologs in the *E. coli* genome**

Many *E. coli* proteins derive from prophages and while many of these proteins have mutated or become pseudogenes, some still seem to be active (19). Utilizing all obtained host-phage interactions, we found seven *E. coli* proteins of phage origin that were involved in 10 phage host interactions (Table 5). Six out of these 7 proteins were highly similar to phage lambda proteins with a sequence similarity of up to 100%. Previously, lambda gpA was found to interact with lambda Nu1 and Orf79 interacted with itself (2). Since all three proteins have *E. coli* homologs of phage origin, it was not surprising that the lambda proteins also interact with the cellular homologues of Nu1 (*E. coli* NohA and NohB) and Orf79 (*E. coli* YbcW), respectively (Table 5). The Nu1 homologue NohA was also found to interact with the phage antiterminator gpN. Even though the biological relevance of this interaction remains unclear, binding of gpN to lambda Nu1 has been previously reported (2).
Interactions of the phage lambda defective tail fiber proteins Orf314 and Orf401 with the host-encoded tail fiber assembly proteins TfaR and TfaQ and the tail fiber protein StfR, respectively, appear plausible. In turn, the biological relevance of other interactions remain obscure, like lambda integrase-NohB, lambda gpO-TfaQ and gpP-YdaG, given that phage lambda has no homologues of YdaG, a protein of Rac prophage origin.

**Topological properties of the phage-host interaction network**

To investigate the location of targeted proteins, we analysed a network of protein-protein interactions of *E. coli*. Utilizing the combined network of high-confidence and previously known host-phage interactions, we calculated the enrichment of targeted proteins as a function of their degree (Fig. 4A), suggesting that host proteins with an increased number of interaction partners are prime targets for the phage. Similarly, genes that are required for the phage infection are enriched among higher connected host proteins as well.

As a corollary to the observed phage’s preference to target central positions in the protein interaction network of *E. coli* we hypothesized that targeted proteins allow the pathogen to reach other proteins efficiently. In particular, we calculated shortest paths from reference proteins to remaining proteins in the underlying network of interaction between *E. coli* proteins. Fig. 4B suggests that lengths of paths from targeted proteins to all other proteins are significantly shorter than the corresponding paths from required (P < 10^{-10}, Student’s t-test) and non-targeted and non-required proteins (P < 10^{-15}), respectively. As for the relative placement of targeted proteins and genes that lambda requires for its infection and proliferation, we calculated shortest paths between such sets of proteins. In Fig. 4C, we observed that the distribution of path lengths between targeted and required genes was significantly shifted to smaller values than paths from targeted proteins to non-required, other proteins (P = 2.9 × 10^{-4}, Student’s t-test).

**Interaction of protein-protein and gene regulatory networks**

We utilized PPIs and regulatory interactions between transcription factors and their target genes in *E. coli* to identify paths from targeted genes to genes that are required for the infection of the phage. Specifically, we observed 5 required host genes that are directly targeted by phage PPIs (hflD, hldD, ihfAB, and lamB). As a mean to control their
expression we identified the shortest path from required genes to targeted genes that started with a transcription factor – gene interaction, prompting us to focus on 27 required genes involving 78 shortest paths. In Fig. 5, we show that all required genes involved in lamB regulation were either directly targeted (lamB) or were under control of transcription factor Crp (malT, malI and cyaA). As for other genes that play a role in the ability of the phage to penetrate the cell, Crp controls the expression of manZ, a required gene and encoding a subunit of the ManXYZ mannose permease. Similarly, genes that are involved in CII degradation were directly targeted by the phage (hflD) or were controlled by Crp (hflK and hflC).

**Discussion**

**Overlap with previously known interactions**

Given our limited knowledge about lambda-host interactions, we identified 62 mostly new high confidence interactions (15). Although we were initially disappointed by finding few known PPIs, we can explain many of the false negatives in retrospect. Nine out of the more than 30 previously published interactions involve host proteases and their phage substrates. Such interactions are hard to detect with a high-throughput approach since they are probably weak, short and transient, although we did find one protease, ClpP, to interact with gpO. The other detected, known interaction appears between gpN and NusA, a well-known interaction involved in anti-termination of lambda transcription (15). We did not find the five previously known interactions involving transcription factors, namely those among CI, CII, gpQ, and RNA polymerase or sigma-70 (15). Multiple proteins may be required to stabilize such proteins in a quaternary structure, a condition our Y2H assay did not capture. Out of the remaining known interactions, two appear between phage tail proteins and the phage receptors LamB and OmpC. Since we did not clone and screen gpJ in our Y2H assay we cannot assess if these interactions are amenable to our assay. Although we found an interaction between replication proteins gpO and gpP in our previous study (2), we were not able to detect interactions among gpP and DnaA or DnaB. The latter interaction may require the hexameric structure of DnaB. Furthermore, we did not find an interaction between gpP and DnaA, possibly because this interaction
may necessitate DnaA to bind DNA. The remaining four undetected interactions are all involved in recombination (Xis-Fis, Int-IHF, Gam-RecB, NinB-SSB). Xis is one of the few proteins that we were not able to clone and screen. Since IHF is a heterodimer, its interaction with Int may require three proteins and therefore may have eluded our experimental approach. As for the remaining two missed interactions (Gam-RecB, NinB-SSB) we also surmise that they may require DNA-binding to be stabilized.

In summary, results of our screens are clearly impaired by false negatives (i.e. known but undetected interactions). The most plausible explanations for this finding are threefold: First, many host-phage interactions appear to require additional factors such as other proteins or DNA. Second, lambda appears to have an unusually high fraction of proteins whose interactions with the host require post-translational modification such as proteolytic cleavage. This observation is in stark contrast to our lambda-lambda protein interaction screen: lambda proteins do not appear to involve many modified proteins and thus we found about half of all previously reported interactions (2). Finally, many intraviral interactions are of structural nature making them more stable than interactions involved in transcription, translation and regulatory processes. Thus, they may be easier to detect by Y2H screens than PPIs involved in the virus-host interplay.

**Differences between phages**

Our results may suggest that the Y2H system works well for phage-phage interactions but not for phage-host interactions, given the sophisticated interaction mechanisms that involve protein processing or other post-translational mechanisms inside the host. However, this conclusion is not supported by results from other phages. For instance, while lambda has 9 interactions involving proteases, no such interactions appear to occur in phage T7 (15). Similarly, while a genetic screen of *E. coli* mutants has identified 57 host genes required for lambda infection, a similar screen with T7 identified only 11, of which 9 (82%) are involved in LPS biosynthesis (20). In the lambda screen only 8 out of 57 genes (14%) are involved in LPS biosynthesis (12). Clearly, different phages have evolved many different ways to conquer the same host, suggesting that there may be even more infection mechanisms that are currently unknown (Fig. 6).
**Interactions with required genes**

In our screens and among previously known interactions we found only 5 targeted host genes that are required for the infection process of lambda (namely HflD, LamB, IhfAB, and RfaD, Tables 1 and 4). Although small, this number is not entirely surprising: in fact, almost half of those E. coli proteins that interact with lambda are essential proteins and are therefore not present in the KEIO collection (i.e. Maynard et al. could not find them). Unexpectedly, however, we found only one required gene in our high-confidence Y2H data (namely RfaD = HldD) that interacted with lambda gpQ. Although small, the size of the overlap between targeted and required genes is comparable to previously published interactions: only 3 of the genes found by Maynard et al. (HflD, LamB, and IHF) have been reported in previous studies to directly interact with lambda proteins. Another reason why we may have not found many required genes is the fact that roughly half of them are involved in metabolism such as LPS synthesis or tRNA thiolation (possibly affecting the translation of lambda transcripts). That is, while our screens are biased towards structural or signaling proteins, genetic screens as in Maynard et al. (12) are biased towards genes that are involved in metabolism. However, metabolic enzymes are much more rarely involved in PPIs (except for homo- and heterodimerization) but rather in enzyme-substrate binding, resulting in largely non-overlapping hits.

**Topological characteristics of host-phage interactions**

Analyzing their network topology we observed that proteins with an increasing number of interaction partners were preferably targeted by the phage. We observed a similar yet weaker trend for genes required for infection (Fig. 4A), observations that have been reported for human host-viral (17, 21) as well as host-parasite interactions (22, 23). Another topological measure that reflects central placement of targeted proteins is the shortest path between proteins in a protein interaction network. We found that targeted proteins have shorter paths to other proteins, suggesting that these targets have a more central role in the cell than others (24). Furthermore, paths from targeted proteins to required genes were significantly shorter compared to remaining proteins, further indicating that required genes are more closely placed to targeted host proteins. These
results suggest that host-phage interactions allows the phage to invade the host cell effectively on different levels of cellular organization, similar to human viruses (21).

**Phage Interference with regulatory interactions**

The availability of a large set of host-phage interactions enabled us to identify potential paths the phage uses to exert control on a small set of required genes. Determining connections from the phage to required host genes we identified transcription factors that are central to the regulation of required genes (such as Crp or Fis) and are at least indirectly targeted by the phage. While our results are based on topological considerations only our data makes several predictions, e.g. that transcriptions factors such as Crp and Fis should have a significant effect on lambda (Fig. 5). This predictions has been confirmed in recent experiments (25).

**Acknowledgements**

We thank Sherwood Casjens for constructive comments on the manuscript. This project was partly funded by by NIH grant R01GM79710 and the European Union (grant HEALTH-F3-2009-223101, “AntipathoGN”). Research at NCBI was supported by the National Institutes of Health/Department of Health and Human Service (DHHS) (Intramural Research program of the National Library of Medicine).

**References**


Figure Legends

Figure 1: Screening strategy. (A) Arrayed pool screening. (B) Typical array screen with positive yeast colonies indicated by arrows. The plates are in 384-format, and each prey pool is pinned as quadruplicate. (C) We cloned 68 lambda ORFs into two Y2H bait vectors and screened against two Y2H prey libraries (one in pDEST22 and one in pGADT7g). All screening results were pooled, yielding 631 raw interactions. After removal of 487 ‘low’ confidence” interactions that involved promiscuous baits and preys and another round of retesting we obtained 244 interactions. Furthermore, we retested the initial raw set of interactions, allowing us to confirm 162 interactions, yielding a set of 62 ‘high’ confidence interactions.

Figure 2: The core E. coli–lambda interactome organized by function. (A) Protein interactions mapped onto the lambda genome. All previously published (green bars), 244 final (orange bars) and 62 ‘high’ confidence interactions (red bars) between proteins of E. coli and the-lambda phage were plotted along the lambda genome for each protein. Furthermore, we labeled functional groups of genes (red: structural proteins, green: regulatory proteins, blue: DNA replication, yellow: lysis) and indicated interactions between targeted proteins. In (B) we augmented the network of 62 ‘high’ confidence protein interactions with more than 30 previously published interactions (green arrows). All proteins are colored according to their functional annotations.

Figure 3. The host-lambda regulatory network. In the core E. coli–lambda interactome we augmented the network of 62 ‘high’ confidence protein interactions with more than 30 previously published interactions. Among the targets of the phage, we highlighted 9 transcription factors, 22 essential proteins as well as 5 genes that are required for the lambda infection.

Figure 4: Topological analysis of host-phage interactions. (A) Utilizing our combined set of previously published and ‘high’ confidence interactions, we observed that highly connected host proteins are increasingly targeted by the phage. Focusing on genes that
are required for the lambda phage infection, we observed a similar, albeit weaker trend. In (B) we calculated shortest paths from reference proteins to all other host proteins. The shortest paths from targeted proteins are significantly shorter than the corresponding paths from genes that are necessary for phage infection (Student’s t-test, P < 10^{-10}). (C) Determining the shortest paths from targeted proteins the distance to genes that are necessary for the phage infection were significantly shorter than remaining proteins (P = 2.9 \times 10^{-4}).

**Figure 5: The gene regulatory interactions targeted by lambda.** Utilizing a network of protein-protein and transcription factor-gene interactions in *E. coli* we calculated shortest paths from a set of genes that are required for lambda infection to targeted host genes. Specifically, we demanded that a shortest path from required to targeted proteins started with a transcription factor – gene interaction. We determined 78 shortest paths from 27 required genes to proteins that were targeted by the lambda phage.

**Figure 6. Comparison of the host-phage interactions in lambda and T7.** We collected previously published protein interactions between *E. coli* and the phages lambda and T7. We observed that only a small minority of host proteins is targeted by both phages (RecB and the HsdMS complex).
### Table 1: Previously published lambda-host PPIs. Modified after (15).

<table>
<thead>
<tr>
<th>#</th>
<th>λ</th>
<th>Host</th>
<th>Description and notes</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Transcription</strong></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>CI</td>
<td>RecA</td>
<td>RecA degrades cl*</td>
<td>(26)</td>
</tr>
<tr>
<td>2</td>
<td>CI</td>
<td>RpoA</td>
<td></td>
<td>(27)</td>
</tr>
<tr>
<td>3</td>
<td>CI</td>
<td>RpoD</td>
<td></td>
<td>(28, 29)</td>
</tr>
<tr>
<td>4</td>
<td>CII</td>
<td>ClpYQ</td>
<td>ClpYQ degrades CII <em>in vitro</em></td>
<td>(26)</td>
</tr>
<tr>
<td>5</td>
<td>CII</td>
<td>ClpAP</td>
<td>ClpAP degrades CII <em>in vitro</em></td>
<td>(26)</td>
</tr>
<tr>
<td>6</td>
<td>CII</td>
<td>HfB</td>
<td>HfB makes CII more vulnerable to FtsH (hfl = high frequency lysogenization)</td>
<td>(30)</td>
</tr>
<tr>
<td>7</td>
<td>CII</td>
<td>HfB</td>
<td>HfB ( = FtsH) protease degrades CII</td>
<td>(31, 32)</td>
</tr>
<tr>
<td>8</td>
<td>CII</td>
<td>RpoA</td>
<td></td>
<td>(33, 34)</td>
</tr>
<tr>
<td>9</td>
<td>CII</td>
<td>RpoD</td>
<td></td>
<td>(33)</td>
</tr>
<tr>
<td>10</td>
<td>CII</td>
<td>RpoD</td>
<td></td>
<td>(26, 35)</td>
</tr>
<tr>
<td>11</td>
<td>gpN</td>
<td>NusA</td>
<td>Transcriptional regulation</td>
<td>(36)</td>
</tr>
<tr>
<td>12</td>
<td>gpN</td>
<td>Lon</td>
<td>Lon degrades gpN</td>
<td>(26, 37)</td>
</tr>
<tr>
<td>13</td>
<td>gpQ</td>
<td>σ70</td>
<td>gpQ makes RNAP insensitive to cis termin.</td>
<td>(29, 34, 38, 39)</td>
</tr>
</tbody>
</table>

|    |    |      | **Head**               |            |
| 14 | gpB | GroE | genetic interaction; E. coli GroE | (40), (41, 42) |
| 15 | gpE | GroE | genetic interaction | (41)       |

|    |    |      | **Tail**                |            |
| 16 | gpJ | LamB | LamB is the E. coli receptor | (43-47) |
| 17 | Stf | OmpC | OmpC is a secondary E. coli receptor | (48) |

|    |    |      | **Recombination**       |            |
| 18 | Xis | Lon | Xis is degraded by *E. coli* Lon protease | (49)       |
| 19 | Xis | FtsH | Xis is degraded by *E. coli* FtsH protease | (49)       |
| 20 | Xis | Fis | both required for excision | (50-52) |
| 21 | Int | IHF | Both catalyze recombination at attP/attB | (53, 54) |
| 22 | Gam | SbcC | sbcCD is a dsDNA exonuclease + ssDNA endonuclease | (55) |
| 23 | Gam | RecB | Gam inhibits RecBCD | (56) |
| 24 | NinB | SSB | NinB also binds ssDNA | (57) |
| 25 | Rol | Hsd | Rol inhibits restriction enzyme complex HsdMSR | (58) [by binding to HsdM or S] |

|    |    |      | **Replication**         |            |
| 26 | gpO | ClpXP | ClpXP degrades gpO | (26) |
| 27 | gpO | DNAK |                        | (59)       |
| 28 | gpO | RpoB |                         | (60)       |
| 29 | gpP | DNAA |                         | (61)       |
| 30 | gpP | DNAB |                         | (62)       |
| 31 | gpP | DNAK |                         | (59)       |

CbpA (63) may interact with DnaA and SeqA (64-66). **Bold:** required for infection (12). **Underline:** found in our screen. *After DNA damage, RecA bound to single stranded DNA stimulates auto-cleavage of the CI repressor (26). * HflB = FtsH.
Table 2: Overview summarizing the screening procedures of the lambda baits against the E. coli prey libraries.

<table>
<thead>
<tr>
<th>Baits</th>
<th>Bait vector</th>
<th>Prey vector</th>
<th>Properties</th>
<th>Prey library</th>
<th>Screen type</th>
</tr>
</thead>
<tbody>
<tr>
<td>λ ORFs</td>
<td>pGBK7T7g</td>
<td>pGADT7g</td>
<td>high copy</td>
<td>E. coli W3110</td>
<td>arrayed pool</td>
</tr>
<tr>
<td>λ ORFs</td>
<td>pGBK7T7g</td>
<td>pGADT7g</td>
<td>high copy</td>
<td>E. coli W3110</td>
<td>Prey pool</td>
</tr>
<tr>
<td>λ ORFs</td>
<td>pDEST32</td>
<td>pDEST22</td>
<td>low copy</td>
<td>E. coli W3110</td>
<td>Prey pool</td>
</tr>
</tbody>
</table>

Table 3: Numeric results of the phage lambda-host pool screens.

<table>
<thead>
<tr>
<th>E. coli prey proteins interacting with lambda</th>
<th>Total number</th>
<th>Promiscuous</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unique bait/prey pairs</td>
<td>294</td>
<td>25</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Unique bait/prey pairs, pDEST32/22 only</th>
<th>Total number</th>
<th>Single hits</th>
<th>Multiple hits with promiscuous preys</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>631</td>
<td>334</td>
<td>144</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Unique bait/prey pairs, pGBK7T7g/pGADT7g only</th>
<th>Total number</th>
<th>Single hits</th>
<th>High-confidence hits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>573</td>
<td>296</td>
<td>122</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Unique bait/prey pairs found in both pDEST32/22 and pGBK7T7g/pGADT7g</th>
<th>Total</th>
<th>Retested individually</th>
<th>high confidence hits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>162</td>
<td>62</td>
</tr>
</tbody>
</table>
Table 4: All high-confidence interactions of this study. The protein interactions from this publication have been submitted to the IMEx (http://www.imexconsortium.org) consortium through IntAct (67) and assigned the identifier IM-21394.

<table>
<thead>
<tr>
<th>PPI</th>
<th>λ locus#</th>
<th>λ locus</th>
<th>E. coli</th>
<th>Annotation</th>
</tr>
</thead>
<tbody>
<tr>
<td>43</td>
<td>lambdap01</td>
<td>nu1</td>
<td>DcrB</td>
<td>protein DcrB</td>
</tr>
<tr>
<td>6</td>
<td>lambdap02</td>
<td>A</td>
<td>NohA</td>
<td>bacteriophage DNA packaging protein</td>
</tr>
<tr>
<td>15</td>
<td>lambdap02</td>
<td>A</td>
<td>NohB</td>
<td>bacteriophage DNA packaging protein</td>
</tr>
<tr>
<td>16</td>
<td>lambdap09</td>
<td>Fi</td>
<td>YdgH</td>
<td>protein YdgH</td>
</tr>
<tr>
<td>27</td>
<td>lambdap09</td>
<td>Fi</td>
<td>FixB</td>
<td>protein FixB</td>
</tr>
<tr>
<td>44</td>
<td>lambdap09</td>
<td>Fi</td>
<td>MinE</td>
<td>cell division topological specificity factor MinE</td>
</tr>
<tr>
<td>55</td>
<td>lambdap09</td>
<td>Fi</td>
<td>CchB</td>
<td>ethanolamine utilization N Eaton</td>
</tr>
<tr>
<td>1</td>
<td>lambdap14</td>
<td>G</td>
<td>ClpP</td>
<td>ATP-dependent Clp protease, proteolytic subunit ClpP</td>
</tr>
<tr>
<td>2</td>
<td>lambdap14</td>
<td>G</td>
<td>FhuF</td>
<td>ferric iron reductase protein FhuF</td>
</tr>
<tr>
<td>7</td>
<td>lambdap14</td>
<td>G</td>
<td>FdoH</td>
<td>formate dehydrogenase, beta subunit</td>
</tr>
<tr>
<td>12</td>
<td>lambdap14</td>
<td>G</td>
<td>YohN</td>
<td>conserved hypothetical protein</td>
</tr>
<tr>
<td>17</td>
<td>lambdap14</td>
<td>G</td>
<td>ChaC</td>
<td>cation transport protein ChaC</td>
</tr>
<tr>
<td>28</td>
<td>lambdap14</td>
<td>G</td>
<td>ProQ</td>
<td>ProP effector</td>
</tr>
<tr>
<td>29</td>
<td>lambdap14</td>
<td>G</td>
<td>YliL</td>
<td>b0816 hypothetical protein (yliL)</td>
</tr>
<tr>
<td>4</td>
<td>lambdap16</td>
<td>H</td>
<td>YliL</td>
<td>b0816 hypothetical protein (yliL)</td>
</tr>
<tr>
<td>18</td>
<td>lambdap16</td>
<td>H</td>
<td>YeiW</td>
<td>proteinase inhibitor</td>
</tr>
<tr>
<td>19</td>
<td>lambdap16</td>
<td>H</td>
<td>YfcQ</td>
<td>conserved hypothetical protein</td>
</tr>
<tr>
<td>20</td>
<td>lambdap16</td>
<td>H</td>
<td>YohH</td>
<td>YohH</td>
</tr>
<tr>
<td>30</td>
<td>lambdap16</td>
<td>H</td>
<td>YehD</td>
<td>fimbrial protein</td>
</tr>
<tr>
<td>10</td>
<td>lambdap33</td>
<td>int</td>
<td>NohB</td>
<td>bacteriophage DNA packaging protein</td>
</tr>
<tr>
<td>31</td>
<td>lambdap36</td>
<td>ea8.5</td>
<td>YdjI</td>
<td>conserved hypothetical protein</td>
</tr>
<tr>
<td>45</td>
<td>lambdap36</td>
<td>ea8.5</td>
<td>MinE</td>
<td>cell division topological specificity factor MinE</td>
</tr>
<tr>
<td>46</td>
<td>lambdap36</td>
<td>ea8.5</td>
<td>YeiW</td>
<td>proteinase inhibitor</td>
</tr>
<tr>
<td>21</td>
<td>lambdap38</td>
<td>orf63</td>
<td>YqhC</td>
<td>putative HTH-type transcriptional regulator YqhC</td>
</tr>
<tr>
<td>3</td>
<td>lambdap45</td>
<td>ea10</td>
<td>Rmf</td>
<td>b0953 ribosome modulation factor (rmf)</td>
</tr>
<tr>
<td>8</td>
<td>lambdap45</td>
<td>ea10</td>
<td>RpmA</td>
<td>ribosomal protein L27</td>
</tr>
<tr>
<td>9</td>
<td>lambdap45</td>
<td>ea10</td>
<td>YliL</td>
<td>b0816 hypothetical protein (yliL)</td>
</tr>
<tr>
<td>11</td>
<td>lambdap45</td>
<td>ea10</td>
<td>RpsG</td>
<td>ribosomal protein S7</td>
</tr>
<tr>
<td>22</td>
<td>lambdap45</td>
<td>ea10</td>
<td>PriC</td>
<td>primosomal replication N</td>
</tr>
<tr>
<td>32</td>
<td>lambdap45</td>
<td>ea10</td>
<td>CobB</td>
<td>NAD-dependent deacetylase</td>
</tr>
<tr>
<td>33</td>
<td>lambdap45</td>
<td>ea10</td>
<td>SoxS</td>
<td>regulatory protein SoxS</td>
</tr>
<tr>
<td>34</td>
<td>lambdap45</td>
<td>ea10</td>
<td>YcbG</td>
<td>protein YcbG</td>
</tr>
<tr>
<td>35</td>
<td>lambdap49</td>
<td>N</td>
<td>Hcr</td>
<td>NADH oxidoreductase hcr</td>
</tr>
<tr>
<td>47</td>
<td>lambdap49</td>
<td>N</td>
<td>NuoG</td>
<td>NADH-quinone oxidoreductase, chain g</td>
</tr>
<tr>
<td>48</td>
<td>lambdap49</td>
<td>N</td>
<td>YebR</td>
<td>protein YebR</td>
</tr>
<tr>
<td>54</td>
<td>lambdap49</td>
<td>N</td>
<td>NusA</td>
<td>transcription elongation N NusA</td>
</tr>
<tr>
<td>56</td>
<td>lambdap49</td>
<td>N</td>
<td>EnvR</td>
<td>probable acrEF/envCD operon repressor</td>
</tr>
<tr>
<td>57</td>
<td>lambdap49</td>
<td>N</td>
<td>MinC</td>
<td>septum site-determining protein MinC</td>
</tr>
<tr>
<td>58</td>
<td>lambdap49</td>
<td>N</td>
<td>RpoS</td>
<td>RNA polymerase sigma factor RpoS</td>
</tr>
<tr>
<td>59</td>
<td>lambdap49</td>
<td>N</td>
<td>YcbG</td>
<td>protein YcbG</td>
</tr>
<tr>
<td>λ bait</td>
<td>λ homolog</td>
<td>E. coli preys</td>
<td>Description of E. coli prey</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
<td>---------------</td>
<td>-----------------------------</td>
<td></td>
</tr>
<tr>
<td>gpA</td>
<td>Nu1</td>
<td>NohA</td>
<td>bacteriophage DNA packaging protein</td>
<td></td>
</tr>
<tr>
<td>gpN</td>
<td>Nu1</td>
<td>NohA</td>
<td>bacteriophage DNA packaging protein</td>
<td></td>
</tr>
<tr>
<td>Int</td>
<td>Nu1</td>
<td>NohB</td>
<td>DLP12 prophage, DNA packaging protein</td>
<td></td>
</tr>
<tr>
<td>gpA</td>
<td>Nu1</td>
<td>NohB</td>
<td>DLP12 prophage, DNA packaging protein</td>
<td></td>
</tr>
<tr>
<td>orf-401</td>
<td>orf-314</td>
<td>StrR</td>
<td>Rac prophage; predicted tail fiber protein</td>
<td></td>
</tr>
<tr>
<td>gpO</td>
<td>orf-194</td>
<td>TfaQ</td>
<td>Qin prophage; predicted tail fibre assembly protein</td>
<td></td>
</tr>
<tr>
<td>orf-314</td>
<td>orf-194</td>
<td>TfaQ</td>
<td>Qin prophage; predicted tail fibre assembly protein</td>
<td></td>
</tr>
<tr>
<td>orf-314</td>
<td>orf-194</td>
<td>TfaR</td>
<td>Rac prophage; predicted tail fiber protein</td>
<td></td>
</tr>
<tr>
<td>orf-79</td>
<td>orf-79</td>
<td>YbcW</td>
<td>DLP12 prophage, predicted protein</td>
<td></td>
</tr>
<tr>
<td>gpP</td>
<td>-</td>
<td>YdaG</td>
<td>Rac prophage; predicted protein</td>
<td></td>
</tr>
</tbody>
</table>