Supporting Online Material for

**Repeatability and Contingency in the Evolution of a Key Innovation in Phage Lambda**

Justin R. Meyer\*, Devin T. Dobias, Joshua S. Weitz, Jeffrey E. Barrick, Ryan T. Quick, Richard E. Lenski

\*To whom correspondence should be addressed. E-mail: justin.raymond.meyer@gmail.com

**This PDF file includes:**

Materials and Methods

Supporting Information

Tables S1-S6

Figs. S1-S7

References

**Materials and Methods**

**Strains.** We studied *Escherichia coli* B strain REL606 because its evolution in the laboratory is well documented (*1*), its genome has been sequenced (*2*), and it is a permissive host for phage λ. Also, REL606 lacks generalized phage defenses such as a restriction modification system, CRISPR adaptive immunity, and mucoid cell formation (*2, 3*). The phage λ strain that we used is cI26, which was provided to us by Donald Court (National Cancer Institute, MD). Most λ strains have two alternative life cycles, lytic and lysogenic. During the lytic cycle, a phage uses the bacterial cell to produce new phage particles and then lyses the host. When a phage enters the lysogenic cycle, its genome is incorporated into and replicated with the bacterial genome (*4*). The λ strain cI26 is strictly lytic as a consequence of a deletion that causes a frameshift in the *cI* gene, a repressor required for the phage to switch into the lysogenic mode.

**Evolution experiments.** Bacteria and phage were cultured together in 50-ml Erlenmeyer flasks, each containing 10 ml of modified M9 medium (*5*)supplemented with 5 times the usual MgSO4 concentration (1 g/L) to improve λ growth and 1 g/L of glucose to allow the bacteria to reach high density. We added ~102 phage particles and ~103 bacterial cells to each flask at the start of an experiment. These small numbers minimized the initial genetic variation; thus, beneficial mutations arose *de novo*, which allows one to evaluate the repeatability of evolutionary outcomes without the complicating effect of shared variation. Each flask was incubated for 24 h at 37°C and shaken at 120 rpm. After 24 h, a 100-µl sample of the community was transferred to a flask containing 9.9 ml of fresh medium. The initial experiment ran for 28 days and the large-scale experiment for 20 days. These experiments had 6 and 96 replicate communities, respectively.

Samples of the communities were periodically preserved by adding glycerol (~15% by volume) to the cultures, which were then frozen at –80°C. For the initial experiment, 1-ml samples were taken every week (days 1, 8, 15, 22, and 28), while 200-μl samples were stored daily for the large-scale experiment. Before freezing, each sample was tested for the presence of phage able to exploit a new receptor by plating a subsample (2-5 μl) onto a lawn of a mutant *E. coli* with defective LamB protein (*lamB–*; derived from REL606 by a 1-bp insertion [T] after nucleotide position 610). The section on “Detection of λ that use new receptor” provides further details.

**Isolation and culture techniques.** To isolate bacterial clones from a community, we spread a portion of the appropriate sample on a Luria-Bertani (LB) agar plate (*5*) and, after incubation for 24 h at 37°C, picked individual colonies. The isolates were streaked twice more in the same manner to eliminate any phage particles that might be present. After the third cycle, a colony was picked and grown overnight at 37°C in liquid LB shaken at 120 rpm. Two ml of this culture was stored with 15% glycerol at –80°C. To revive cultures, ~3 μl of frozen stock was grown for one day in LB, then 10 μl was transferred to a flask containing 10 ml of modified M9 medium and grown for 24 h to acclimate the cells to the experimental conditions.

Phage were sampled by plating serial dilutions of the community culture into 4 ml of molten (~50°C) soft agar (LB agar except with only 0.8% w/v agar) infused with ~5 x 108 cells of the ancestral bacterial strain, REL606. The agar was poured over an LB agar plate, allowed to solidify, and incubated overnight at 37°C. We then picked individual plaques (~1 mm diameter), each the product of a single virus, from suitable dilutions. Phage stocks were grown on REL606 cells in modified M9 following procedures adapted from ref. *6*, then stored with 2% chloroform at 4°C. Aliquots were added directly to the experiments from these refrigerated stocks. Evolved phage stocks tended to decay, therefore they were stored for long term by freezing with glycerol as the bacteria were.

**Estimating population densities.** The density of *E. coli* cells was estimated by dilution in saline solution (8.5 g/L NaCl) followed by plating on LB agar, with a target count of 150-500 colonies per plate. The density of phage λ was estimated in a similar manner except dilutions were done in modified M9 without glucose and plaques were assayed on soft-agar plates.

**Detection of λ that use new receptor.** To determine when λ evolved the ability to use a new receptor, we performed “spot assays” (*5*) on lawns of a *lamB*– mutant derived from the ancestral strain, REL606. For this assay, ~5 x 108 *lamB*– cells were dispersed in soft agar and an undiluted sample of phage was dripped onto the agar. If some of the phage could exploit a receptor other than LamB, then a clear spot in the lawn would be observed after 24 h at 37°C.

**Identifying the new receptor**. To identify the new receptor used by some evolved λ isolates, we performed spot assays using lawns of *E. coli* mutants defective in the production of various outer-membrane proteins. Each test strain lacked LamB and one of seven proteins – OmpA, OmpC, OmpF, OmpG, OmpW, BglH, and PhoE – that share sequence or structural similarities to LamB. We used double mutants because the evolved phage retained the ability to use LamB.

The double mutants were produced starting with seven knockout strains in the Keio Collection (*7*) (table S1), and then introducing a *malT*– mutation to each strain so that it does not express LamB. (See the section on “Evolution of *malT–* mutants”for details of how these mutations affect the expression of LamB.) To generate the *malT–* mutants, we challenged populations (~106 cells) of each Keio strain with the ancestral λ (~108 particles) on LB plates and isolated colonies of resistant mutants. We confirmed the mutants were *malT–* by plating on tetrazolium maltose (TMal) agar plates (*8*).

**Table S1.** Set of *E. coli* knockout strains from the Keio Collection (*7*) used to identify the novel receptor used by some evolved phage λ. The CGSC number is the strain identifier used by the Coli Genetic Stock Center at Yale University.

|  |  |  |
| --- | --- | --- |
| Gene removed | CGSC no. | KEIO name |
| *ompA* | 8942 | JW0940-6 |
| *ompC* | 9781 | JW2203-1 |
| *ompF* | 8925 | JW0912-1 |
| *ompG* | 11793 | JW1312-1 |
| *ompW* | 9125 | JW1248-2 |
| *bglH* | 10702 | JW3698-5 |
| *phoE* | 8466 | JW0231-1 |

**Phage genomics.** To sequence λ strain CI26, which was the ancestral phage in our study, we pooled three 4-ml liquid stocks into a single 12-ml sample containing ~109 plaque-forming units (pfu) per ml. The same approach was used for evolved phage EvoC, except the final preparation had ~107 pfu per ml. Genomic DNA was purified from each sample by using a Qiagen Lambda Midi Kit, fragmented by sonication, prepared as bar-coded libraries, and sequenced on an ABI SOLiD 4 instrument at the University of Texas at Austin’s Genome Sequencing and Analysis Facility. The paired 50-base and 35-base reads were mapped in color space to the reference genome (GenBank: NC\_001416.1) using SHRiMP v2.1.1b (compbio.cs.toronto.edu/shrimp/). Only the top-scoring alignments of properly mapped read pairs were analyzed. The resulting SAM files were reformatted using a custom Perl script, then entered into the *breseq* pipeline v0.13 to predict consensus base substitutions, small indels, and larger deletions as well as to identify any genetic polymorphisms in the sample. The only site that showed heterogeneity was an A→G change at position 18,538 that was present in ~55% of the reads in the ancestral sample. All other mutations were predicted to be consensus changes present in 100% of the sequenced population (Table S2).

**Table S2.** Genomic differences between λ strain CI26, used as the ancestral phage in this study, and the λ reference genome (GenBank: NC\_001416).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Reference genome location** | **Mutation** | **Type** | **Gene position nucleotide (amino acid)** | **Amino acid change** | **Gene** | **Product** |
| 138 | Δ1 bp | Noncoding | /–53 | – | –/nu1 | –/DNA packaging protein |
| 14266 | +G | Noncoding | +139/–10 | – | L/K | tail component/tail component |
| 20661 | A→G | Substitution | 1012 (338) | K→E | orf-401 | Tail fiber protein |
| 20835 | +C | Frameshift | 1186 (396) |  | orf-401 | Tail fiber protein |
| 21714 | G→A | Substitution | 686 (229) | S→N | orf-314 | Tail fiber protein |
| 21738 | Δ5996 bp | Deletion |  |  | [orf-314] orf-194 ea47 ea31 ea59 | Tail fiber, fiber assembly, and proteins of unknown function |
| 31016 | T→C | Substitution | 9 (3) | E→E | orf61 | hypothetical protein |
| 34934 | A→G | Substitution | 453 (151) | G→G | lambdap48 | Superinfection exclusion protein B |
| 37818 | Δ1 bp | Frameshift | 123 (41) | – | cI | repressor |
| 45618 | T→C | Substitution | 126 (42) | F→F | R | endolysin |
| 46957 | +A | Noncoding | ‑205/+85 | – | bor/lambdap78 | Bor protein precursor/putative envelope protein |
| 46985 | C→T | Noncoding | ‑233/+57 | – | bor/lambdap78 | Bor protein precursor/putative envelope protein |
| 46992 | C→T | Noncoding | ‑240/+50 | – | bor/lambdap78 | Bor protein precursor/putative envelope protein |
| 47004 | G→A | Noncoding | ‑252/+38 | – | bor/lambdap78 | Bor protein precursor/putative envelope protein |
| 47129 | A→G | Substitution | 447 (149) | H→H | lambdap78 | putative envelope protein |
| 47143 | C→T | Substitution | 433 (145) | V→I | lambdap78 | putative envelope protein |
| 47243 | G→A | Substitution | 333 (111) | N→N | lambdap78 | putative envelope protein |
| 47315 | G→A | Substitution | 261 (87) | I→I | lambdap78 | putative envelope protein |
| 47360 | G→A | Substitution | 216 (72) | N→N | lambdap78 | putative envelope protein |
| 47398 | C→T | Substitution | 178 (60) | D→N | lambdap78 | putative envelope protein |
| 47509 | T→C | Substitution | 67 (23) | T→A | lambdap78 | putative envelope protein |
| 47529 | C→T | Substitution | 47 (16) | R→K | lambdap78 | putative envelope protein |
| 47575 | C→A | Substitution | 1 (1) | V→L | lambdap78 | putative envelope protein |
| 47669 | T→C | Noncoding | ‑94/‑69 | – | lambdap78/ lambdap79 | putative envelope protein/hypothetical protein |
| 47878 | A→G | Substitution | 141 (47) | R→R | lambdap79 | hypothetical protein |
| 47973 | T→C | Noncoding | +29/ | – | lambdap79/– | hypothetical protein/ |
| 47977 | G→A | Noncoding | +33/ | – | lambdap79/– | hypothetical protein/ |
| 47978 | T→C | Noncoding | +34/ | – | lambdap79/– | hypothetical protein/ |
| 48160 | T→C | Noncoding | +216/ | – | lambdap79/– | hypothetical protein/ |

**Bacterial genomics.** Bacteria were revived from freezer stocks, grown overnight in LB medium, and genomic DNA was isolated from several ml using Qiagen genomic tips. DNA samples were fragmented by sonication, prepared as bar-coded libraries, and run as six of twelve multiplexed samples spread over four lanes on an Illumina GenomeAnalyzer IIx by the Research Technology Support Facility at Michigan State University. Mutations were predicted from the resulting 75-base paired-end DNA reads using *breseq* v0.13 and the genome sequence of the ancestral *E. coli* strain, REL606 (GenBank: NC\_012967.1), as the reference. The *breseq* pipeline performs single-end read alignment to the reference genome with SSAHA2 (www.sanger.ac.uk/resources/software/ssaha2/). In addition to the types of mutations predicted from the phage sequence data by *breseq*, the detection of structural variation from reads with split alignments was enabled for the bacterial samples.

**Data and software availability.** The λ and *E. coli* genome-sequence data have been deposited in the NCBI Sequence Read Archive (SRA043942). The source code for *breseq* is freely available online (barricklab.org/breseq and code.google.com/p/breseq/).

**Targeted sequencing of the J gene.** To find mutations in the gene encoding the J protein of the λ tail (host specificity protein, GenBank: NP\_040600), we sequenced DNA fragments using an automated ABI sequencer.  The fragments were PCR-amplified and purified using a GFX column (GE Healthcare).  Primer sequences were 5’ CTGCGGGCGGTTTTGTCATT 3’ and 5’ ACGTATCCTCCCCGGTCATCACT 3’, which complement sequences 15 bp upstream and 318 bp downstream of the J gene, respectively.

**Null model for non-synonymous mutations.** The sequence of the J gene was obtained from the reference λ genome (GenBank: NC\_001416.1). All possible base substitutions and their effects on the encoded protein were enumerated using a custom Perl script to calculate the ratio of non-synonymous to synonymous mutations among all base substitutions, assuming equal rates.

**Replay experiments.**  To examine whether specific steps along the evolutionary paths taken by the phage and bacteria influenced the likelihood that λ would evolve the ability to use OmpF, we replayed evolution by assembling communities with particular combinations of phage and bacteria that could reveal historical contingencies affecting that outcome. The *E. coli* andλ used in the replays were isolated from different populations at different time points, as described in the main text. The replays were run in the same manner as the other coevolution experiments, except using different strains. Each replay community was propagated for up to 10 days, and each was sampled daily to determine whether λ had evolved to target OmpF by plating 2-5 μl on lawns of the ancestral bacterial strain, the *lamB–* mutant of the ancestor, and the *ompF– malT–* derivative of BW25113. The replays were stopped early if the phage either acquired the ability to exploit OmpF or went extinct.

In the final set of replay experiments, we used diverse bacterial communities rather than clones. This approach required special procedures to include a representative sample of bacteria while excluding phage. For each community of interest we plated ~300 cells on LB agar, picked each colony with a sterile toothpick, and suspended them together in LB broth. We then grew the mixed culture overnight at 37°C with shaking at 120 rpm, and again plated ~300 cells. We repeated this process three times to eliminate phage from the mixed culture. To confirm the absence of phage, we took an aliquot of each mixture, added chloroform to kill the bacteria, let the chloroform settle, and added 1 ml to a lawn of the λ-sensitive ancestor, REL606, in soft agar. No plaques were formed, confirming that this process had eliminated the phage. Finally, we stored 1 ml of each mixed bacterial culture by adding 15% glycerol and freezing it at –80°C. To start the replay experiments, we took 100 μl of the thawed mixture, let it grow overnight in LB, transferred 100 μl to modified M9, and let this culture grow overnight. We then used 100 μl of this culture to initiate each replay community. We expect that this technique was effective at isolating and propagating abundant bacterial genotypes from the source communities, although their frequencies may have shifted and most rare variants would be excluded. These effects might explain why the replay experiments, while highly reproducible, sometimes differed from the corresponding source communities in the initial experiment.

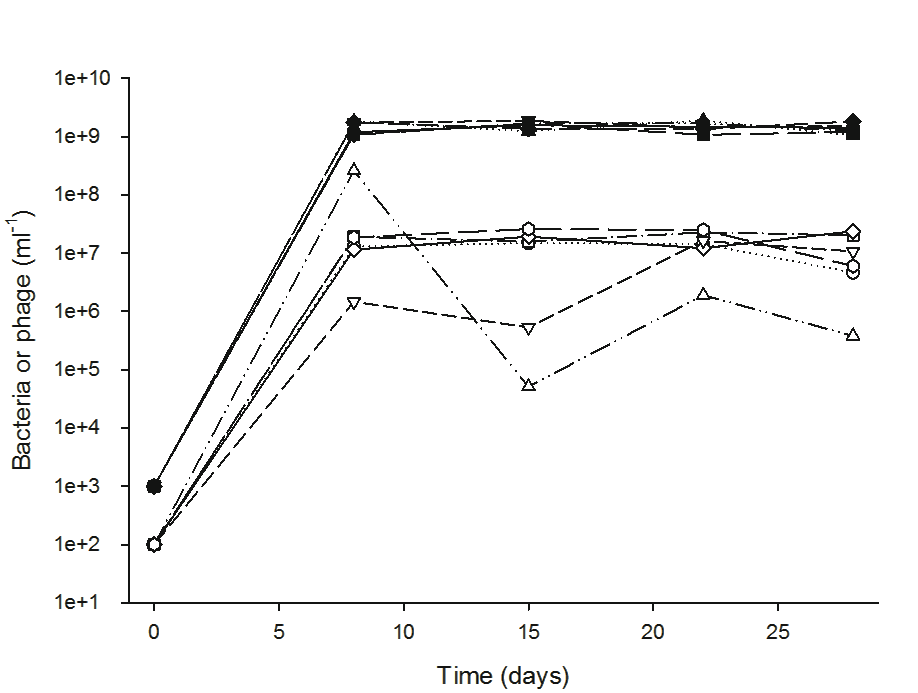
**Supporting Information**

**Evolution of *malT–* mutants.** Preliminary experiments showed that, in minimal glucose medium, *E. coli* strain REL606 generally evolved λ-resistance through *malT–* mutations. MalT is a positive regulator of *lamB*, which encodes the receptor LamB, so that mutations that disrupt MalT function prevent LamBexpression (*9*). MalT also activates other genes required for growth on maltose and other maltodextrins (*9*), and these *malT–* mutations are therefore defective in growth on those substrates. However, these mutations are advantageous, even in the absence of phage, in glucose medium for the *E. coli* strain used in our study (*10, 11*), probably because they reduce the basal expression of unnecessary gene products. This additional benefit may explain why *malT–* mutations evolved in the coevolution experiments, rather than mutations in the *lamB*-encoded receptor that would not have yielded the metabolic cost-savings.

In any case, we tracked the evolution of *malT–* mutants in all 96 populations in the large-scale experiment to determine how often and how quickly these genotypes fixed. We plated a random sample of bacteria (50-100 cells) on TMal plates on days 5 and 8 of the experiment; *malT–* mutants produce red colonies on these plates; the *malT+* ancestor produces white colonies. Fig. S1 shows the frequency of *malT–* cells in all 96 populations.

**Fig. S1.** Rapid fixation of *malT–* mutants in the 96 populations of the large-scale experiment.

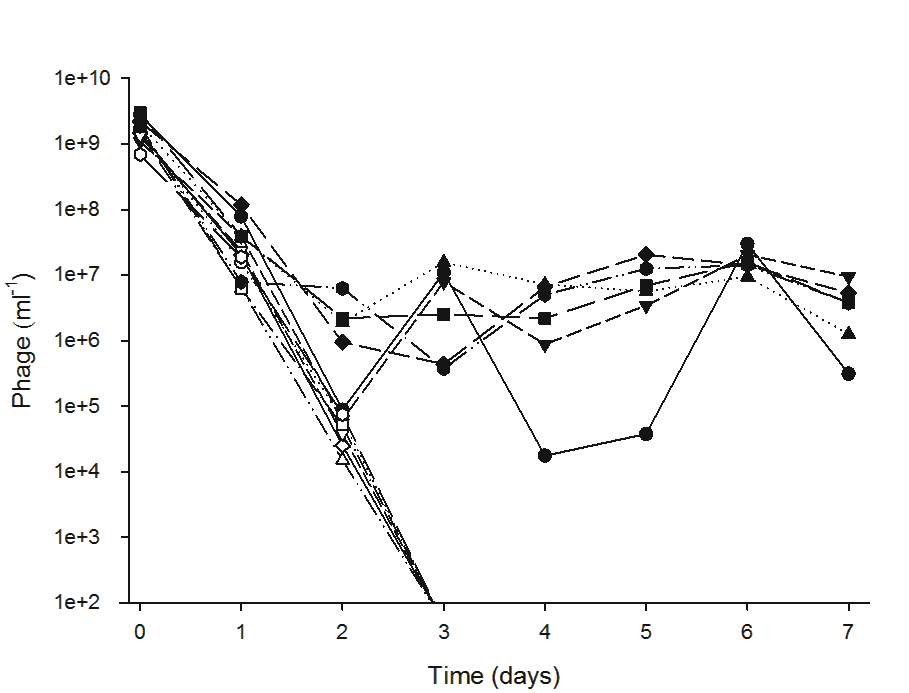
**Population dynamics.**  We quantified the dynamics of the coevolving bacteria and phage in the initial evolution experiment to better understand the conditions under which the phage evolved the novel receptor function. Lytic phages can, in principle, exert top-down limitations on the density of bacteria (*12*). In our experiments, however, any such limitation was quickly overcome as the bacteria evolved high levels of resistance (fig. S1). As a consequence, the phage density was low compared with that of the bacteria (fig. S2). This difference meant that any phage mutant that overcame the resistance would gain access to a large host population. Indeed, one phage population in this experiment evolved to use the OmpF receptor and transiently achieved a higher density (fig. S2: open triangles), but its density declined after the bacteria evolved further resistance in addition to the early *malT–* mutation.

****

**Fig. S2.** Population dynamics of *E. coli* (filled symbols) and phage λ (open symbols) from the six replicates of the initial evolution experiment. Bacteria and phage densities were based on colony and plaque counts, respectively, with the latter obtained by using lawns of the sensitive ancestral host.

**Mechanism of phage persistence following initial resistance.** Phage λ persisted after the rise of *malT–* mutants (Figs. S1, S2), even though the mutants appeared to be completely resistant when the phage were spot-tested on bacterial lawns. We hypothesized that the ancestral λ could infect rare *malT–* cells that spontaneously expressed LamB. This hypothesis is consistent with a study showing that *lamB* regulatory mutants were occasionally infected by wild-type λ (*13*). An alternative explanation is that *malT–* mutants may not have completely fixed in the population if the *malT+* ancestors had a growth-rate advantage that allowed them to maintain a small minority of sensitive cells that λ could exploit. This mechanism has been demonstrated in several studies of coevolving bacteria and phage (*12, 14, 15*). However, this explanation seemed unlikely in the present case because, as already noted, *malT–* mutants have a competitive advantage in glucose-limited media in the absence of phage (*10, 11*). A third possibility is that the ancestral λ could infect cells through some other receptor at a very low rate that would not allow plaque formation on lawns of the *malT–* mutants.

To discriminate among these hypotheses, we isolated a *malT–* mutant of REL606 that had a 25-bp duplication causing a frameshift in this gene. We propagated six communities of the ancestral λ with this bacterial mutant for seven days; we simultaneously ran six replicates with the same phage and the *lamB–* bacterial mutant. The three hypotheses make distinct predictions about whether λ can persist in these two treatments. Under the first hypothesis, in which spontaneous inductions generate a physiological minority of susceptible cells, λ should be maintained on the *malT–* mutant, but not on the *lamB–* mutant. Under the second hypothesis, which requires a subpopulation of genetically sensitive cells, λ should go extinct in both treatments. Under the third hypothesis, according to which even the ancestral phage can use an alternative receptor, λ should persist in both treatments. Fig. S3 shows that the results clearly support the first hypothesis, in which spontaneous induction and expression of LamB allow the ancestral phage to persist only on the *malT–* mutant.



**Fig. S3.**  The ancestral λ strain persists on the *malT–* (closed symbols) but not the *lamB–* (open symbols) mutant host population*.* Phage densities were obtained from plaque counts on lawns of the sensitive ancestral bacteria. No phage were seen after day 2 on the *lamB–* hosts; the limit of detection was 102 phage per ml.

**Properties of the new λ receptor.** There are two hypotheses for what receptor properties are most important for phage binding: hydrogen-bond formation between specific amino acids on the ligand and receptor, or electrostatic interactions facilitated by complementary shape motifs between the two structures (*16*). Consistent with the latter hypothesis, λ evolved repeatedly to use OmpF, which has a similar structure to LamB (*17*), but a very different amino-acid sequence from LamB (table S3). By contrast, the phage never targeted BglH, despite its more similar amino-acid sequence (table S3), although the structural similarity of BglH to LamB is unknown. It may also be relevant that OmpF is more highly expressed than BglH under conditions similar to our experiments (*18*; see transcriptomic data at myxo.css.msu.edu/ecoli/arrays/arrays.txt).

To find proteins with similar amino-acid sequences to LamB, we performed a BLAST protein search (*19*) of LamB (Genbank accession: YP\_003047080) against the genome of the ancestral strain, REL606 (Genbank accession: NC\_012967). Many putative matches were found, although OmpF was not one of the top matches (table S3). To compare the similarity of OmpF and LamB in light of these other proteins, we performed BLAST protein alignments. A few small sections of OmpF matched LamB, but only under the most relaxed settings, and OmpF was not the most similar outer-membrane protein to LamB (table S3).

**Table S3.** Results from BLAST alignment of LamB to all proteins in the ancestral genome. The five top-ranked matches are listed and compared to OmpF. Default blastp parameters were used.

|  |  |  |  |
| --- | --- | --- | --- |
| Protein | Cellular location | No. identical amino acids | Region of putative homology |
| yieC carbohydrate-specific porin (BglH) | outer membrane | 118 | 462 |
| Glycerate kinase II | cytoplasm | 25 | 108 |
| Cytidine deaminase | cytoplasm | 12 | 23 |
| Potassium proton antiporter | membrane | 11 | 26 |
| Glycerophosphodiester phosphodiesterase | periplasm | 11 | 27 |
| Porin protein OmpF | outer membrane | 7 | 14 |

**Time required for λ to target OmpF.** We sampled the 96 communities in the large experiment daily to determine whether the phage had evolved to target a new receptor and, if so, whether that receptor was OmpF. We spotted 2-5 μl of each culture on lawns of the ancestral bacteria, the *lamB*– mutant of the ancestor, and the *ompF*– *malT*– derivative of BW25113. Regions of lysis on the first two lawns, but not on the third, indicated that evolved phage could infect cells using OmpF. This capacity evolved in 24 of the 96 replicates. If the third lawn also showed lysis, that would imply phage could infect cells using some other receptor than LamB or OmpF; however, that outcome was never seen. Fig. S4 shows the timing of the evolution of the ability to use OmpF in the 24 populations that achieved this innovation. No population evolved the trait early in the experiment, in agreement with the finding that λ requires four mutations to use OmpF. The number of populations evolving the new function also appears to have declined toward the end of the experiment, even though 72 populations still had not evolved that function. This latter observation is consistent with the finding that some bacteria evolved resistance mutations that rendered the phages unable to evolve the new function.

**Fig. S4.** Distribution of times at which 24 λ populations in the large-scale evolution experiment first showed the capacity to infect *E. coli* through the OmpF receptor.

**More mutations in phage that targeted OmpF.** Table S4 shows the number of mutations in the J gene of phage isolates from the large-scale experiment. The left half of the table shows all 24 isolates that evolved the ability to use OmpF; these isolates were sampled on the first day this function was observed in the source population. The right half includes 24 isolates from other populations that did not evolve this function; each of these isolates was sampled on the same day as one of the isolates in the first group. Therefore, the rate of evolution for the two groups can be compared statistically by a paired test, with the time available for mutations to have accumulated being the same for the two members of each pair. The test results are presented in the main text.

**Table S4.** Mutations in the gene encoding the J protein in two groups of evolved λ phage, one that acquired the ability to exploit OmpF and the other that remained dependent on LamB.

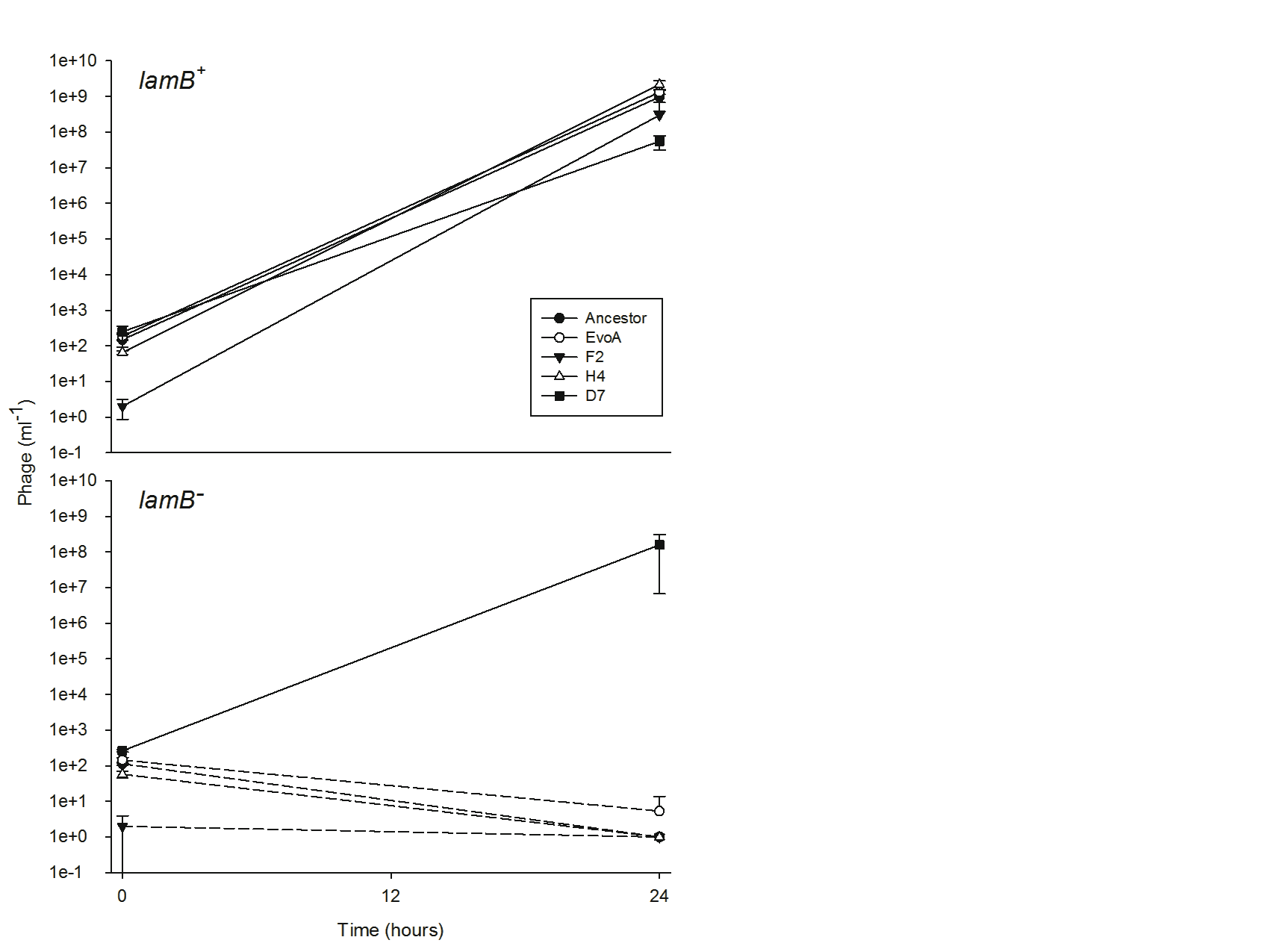
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| λ that can use OmpF | | | λ that use LamB only | | |
| ID | Day isolated | No. of mutations | ID | Day isolated | No. of mutations |
| A7 | 14 | 7 | H2 | 14 | 3 |
| A8 | 11 | 7 | A10 | 11 | 2 |
| A12 | 14 | 10 | A11 | 14 | 4 |
| B2 | 17 | 5 | B3 | 17 | 3 |
| C2 | 12 | 7 | C1 | 12 | 2 |
| C3 | 9 | 5 | C4 | 9 | 3 |
| D3 | 16 | 7 | D1 | 16 | 3 |
| D4 | 10 | 6 | D2 | 10 | 3 |
| D6 | 16 | 7 | D5 | 16 | 3 |
| D7 | 12 | 4 | D8 | 12 | 5 |
| D9 | 12 | 6 | D10 | 12 | 3 |
| E3 | 12 | 7 | E2 | 12 | 3 |
| E4 | 17 | 8 | E6 | 17 | 4 |
| E11 | 15 | 7 | G10 | 15 | 2 |
| E12 | 18 | 7 | F2 | 18 | 5 |
| F5 | 16 | 7 | F4 | 16 | 2 |
| F7 | 12 | 6 | F6 | 12 | 3 |
| F8 | 14 | 8 | F9 | 14 | 4 |
| G4 | 11 | 5 | G3 | 11 | 4 |
| G9 | 15 | 7 | H10 | 15 | 2 |
| H5 | 17 | 7 | H4 | 17 | 7 |
| H8 | 17 | 7 | H7 | 17 | 4 |
| H9 | 10 | 6 | H3 | 10 | 4 |
| H12 | 12 | 6 | H11 | 12 | 4 |

**Two additional tests of receptor profile.**  Spot tests indicated that λ required four mutations in J to be able to exploit OmpF as a receptor. However, this method might not be sensitive enough to detect phage that can use that receptor but with very low efficiency. We therefore performed two additional tests – one based on phage adsorption and the other on phage growth – to verify that all four mutations are required for λ to exploit, even slightly, the OmpF receptor. For each test, we examined four informative phage including three genotypes (EvoA, F2, and H4) at the precipice of evolving the new function (each has 3 of the 4 canonical mutations) and one (D7) with all 4 canonical mutations and no others.

Adsorption assays. We measured the adsorption rates of the four evolved phage using the *lamB*– mutant of the ancestral host strain. This assay measures the rate at which phage adsorb to and infect cells by tracking how many phage remain free (unattached) in the medium over time (*20*). We added ~5 x 104 phage and ~2 x 109 exponentially growing bacteria to 10 ml of modified M9. We measured the density of free phage at six time points over ~25 minutes. We then fit a linear regression to the log ratio of free phage density at time *t*, *p*(*t*), to their initial density, *p*(0), i.e.,ln[*p*(t)/*p*(0)] = *b* × *t*, where the slope, *b*, reflects the rate at which the phage adsorb. The intercept of the regression was constrained to 0 because all phage are unbound at *t* = 0 and densities are expressed relative to the initial value. With samples taken at 6 time points for each experiment, and with the intercept fixed, each regression has 4 degrees of freedom. A significant negative slope indicates that the phage can adsorb to some receptor other than LamB. Only the D7 phage, which has all 4 canonical mutations, showed a significant decline indicative of its ability to use the OmpF receptor (D7: *p* = 0.046, *b* = –0.128; EvoA: *p* = 0.970, *b* = 0.008; F2: *p* = 0.605, *b* = 0.004; H4: *p* = 0.999, *b* = –0.024) (fig. S5). The adsorption-rate constant for D7 is estimated to be ~6 x 10-10 per ml per minute, where that rate is calculated as –*b*/*N* and *N* is the bacterial density. This rate is similar to a previous estimate for wild-type λ using host cells that express LamB (*21*), which implies that the evolved phage D7 adsorbs quite well to OmpF.

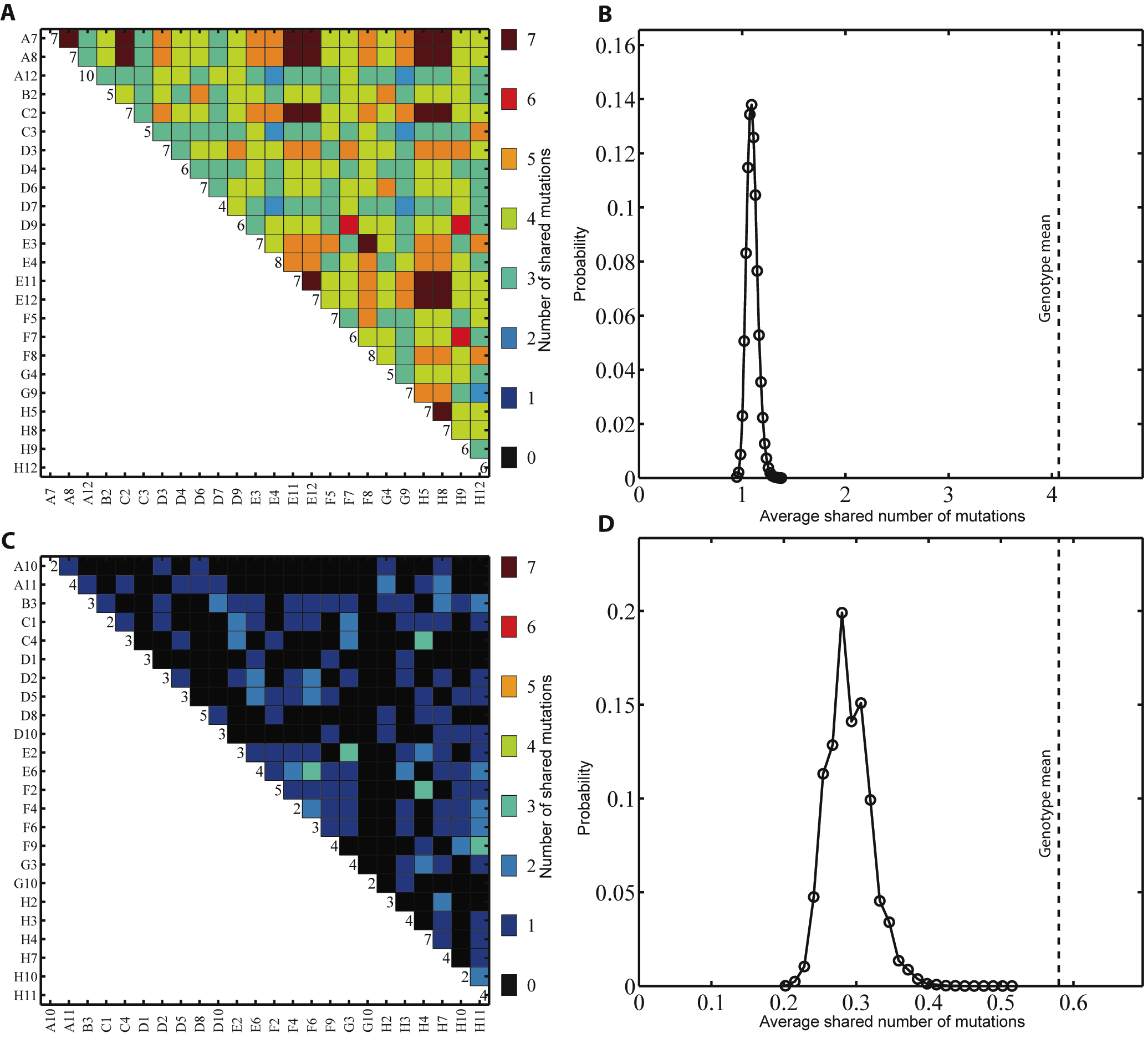
**Fig. S5.** Adsorption assays using four evolved λ genotypes and a *lamB*– bacterial mutant. The concentration of free phage should decline only if the phage can adsorb to cells using a receptor other than LamB. EvoA (diamonds), F2 (squares), and H4 (triangles) all have only 3 of the 4 canonical mutations needed to target OmpF, whereas D7 (crosses) has all 4 mutations.

Growth assays. We performed growth experiments to determine which of four evolved phage λ genotypes (EvoA, F2, H4, and D7) could reproduce on the *lamB*– bacterial mutant. We included the ancestral phage as a negative control, and we also measured phage growth on isogenic *lamB+* bacteria as a positive control. Each combination of phage and bacteria was replicated three-fold. We mixed the phage and bacteria in small volumes (1.2 ml) of modified M9 in glass tubes. We added ~8 x 105 exponentially growing cells to each tube; the initial phage numbers were kept low at ~250 particles (even fewer for F2 owing to its low-density stock) to limit the possibility that mutants derived from the genotypes with three canonical mutations might acquire the final mutation. The cultures were incubated at 37°C and shaken at 160 rpm for 24 h. Phage densities were assessed at the beginning and after 24 h by plaque assays on lawns of the ancestral bacteria. As expected, the ancestral phage and all four evolved types showed robust growth on the *lamB+* bacteria (fig. S6, top panel). Phage D7 also grew very well on the *lamB*– bacteria, but none of the other phage could reproduce at all on the mutant cells (fig. S6, bottom panel). These results confirm that all 4 of the canonical mutations are required for the evolved phage to use OmpF as an alternative receptor to LamB.



**Fig. S6.** Population growth of five λ genotypes on two bacterial hosts. EvoA, F2, and H4 have three of the four canonical mutations required to use OmpF as a receptor, while D7 has all four mutations. Dashed lines in the bottom panel indicate that, when mixed with *lamB–* bacteria, the ancestral phage and evolved types EvoA, F2, and H4 dropped below the limit of detection (~3 pfu ml-1) after 24 h, except for one replicate of EvoA that yielded some plaques. Error bars show 95% confidence intervals.

**Parallel evolution in the J gene.** Parallel evolution provides a strong signal of natural selection. Many studies have documented parallel changes in phenotypes (*10, 11,* *22, 23*), and others have reported parallel evolution at the level of evolving genes (*1, 18, 24-26*). Parallel changes at the level of nucleotide sequences are much less common, although a previous study with a different phage reported extensive parallelism at the nucleotide level (*27*). Fig. 3 (main text) shows many parallel mutations in the gene encoding the J protein across independently evolved λ lineages. To determine if this parallelism was statistically significant, we compared the observed average number of mutations shared by pairs of evolved phage with the random expectation (Fig. S7). We performed the analysis on two separate groups, the phage that evolved to exploit OmpF and those that did not. To generate the null-hypothetical distribution, we constructed 105 random matrices (24 by 40 cells, identical in size and shape to the top or bottom half of Fig. 3 in the main text). We generated each random matrix by shuffling the cells while preserving the number of mutations in each row. This approach is highly conservative because it considers only those sites that differed from the ancestor in at least one sequenced allele, and thus it implicitly ignores all sites that did not vary. We then computed the average number of shared mutations for the actual matrix and for each of the randomized matrices (fig. S7). Among the 24 alleles from phage that evolved the capacity to use OmpF, all pairs shared at least two mutations and, on average, the pairs shared 4.07 mutations (fig. S7A). However, when the cells were randomized, the average pair shared only 1.07 (± 0.05 standard deviation) mutations (fig. S7B). None of the randomized matrices showed parallelism close to the observed level; hence, the signal is highly significant (*p* ≪ 10−5). The evolved genotypes that continued to require LamB shared many fewer mutations; the average pair had 0.58 mutations in common (fig. S7C). Nonetheless, this value was higher than any random matrix (fig. S7D), again indicating highly significant parallelism (*p* < 10−5).

****

**Fig. S7.** Parallel evolution in J protein.(A) Pairwise comparisons among 24 λ genotypes that independently evolved the ability to target OmpF, showing the number of shared mutations for each pair. The values along the diagonal show the number of mutations for each genotype. (B) Probability distribution for the average number of shared mutations based on 105 randomized similarity matrices; the vertical dashed line shows the observed average. (C & D) Same as (A & B), except showing the observed data and randomized distribution for 24 genotypes that retained their dependence on the LamB receptor.

**Mutations in bacterial genomes and their effects on phage evolution.**  We sequenced the complete genomes of all six evolved bacterial clones used in the second replay experiment, and we compared them to the ancestral genome, as described in the Materials and Methods. We observed 15 mutations in total (Table S5). As explained in the main text, the mutations in *manY* and *manZ* uniquely distinguish the bacteria that blocked the evolution of phage able to use OmpF from those bacteria that allowed the phage to evolve that new function.

**Table S5.** Mutations and their phenotypic effects in six *E.* *coli* clones that evolved with phage λ. The first three clones prevented phage from evolving the ability to use OmpF as a receptor, while the last three clones allowed phage to evolve that novel trait.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Clone** | **Genome location\*** | **Mutation** | **Genes affected** | **Effect on proteins\*\*** | **Phenotypic effects\*\*\*** |
| **EcA8** | 1,882,610 | TCTAT insertion | *manZ* | frameshift at AA 49 | Man–, λall-r |
| 3,482,737 | G→A | *malT* | stop at AA 351 | Mal–, λLamB-r |
| 3,894,997 | 4,048-bp IS*150*-mediated deletion | *rbsD–rbsB* | partial deletion of ribose operon | Rbs– |
| **EcC3** | 1,881,820 | 16-bp duplication | *manY* | frameshift at AA 59 | Man–, λall-r |
| 3,482,677 | 25-bp duplication | *malT* | frameshift at AA 339 | Mal–, λLamB-r |
| 3,894,997 | 1,278-bp IS*150*-mediated deletion | *rbsD–[rbsA]* | partial deletion of ribose operon | Rbs– |
| **EcF6** | 1,881,721 | G→T | *manY* | stop at AA 21 | Man–, λall-r |
| 3,482,677 | 25-bp duplication | *malT* | frameshift at AA 339 | Mal–, λLamB-r |
| 3,894,997 | 4,631-bp IS*150*-mediated deletion | *rbsD–[rbsK]* | partial deletion of ribose operon | Rbs– |
| **EcC4** | 3,482,567 | C→T | *malT* | stop at AA 295 | Mal–, λLamB-r |
| **EcD4** | 1,003,919 | G→T | *ompF* | N→K at AA 52 | Probably affects λ adsorption to OmpF |
| 3,482,677 | 25-bp duplication | *malT* | frameshift at AA 339 | Mal–, λLamB-r |
| 3,894,997 | 395-bp IS*150*-mediated deletion | *[rbsD]* | partial deletion of ribose operon | Rbs– |
| **EcH2** | 3,483,588 | T→G | *malT* | L→R at AA 635 | Mal–-, λLamB-r |
| 3,894,997 | 7,868-bp IS*150*-mediated deletion | *rbsD–[yieP]* | deletion of ribose operon | Rbs– |

\* For deletions, location indicates the first base pair (bp) deleted. For duplications, location indicates the first bp of the duplicated region. For insertions, location indicates the last bp before the inserted bases.

\*\* For insertions and duplications, the effect is reported as a frameshift at the first affected amino acid (AA), indicated by its codon number.

\*\*\* Phenotypes include Mal– (unable to use maltose), Man– (unable to use mannose), Rbs– (unable to use ribose), λLamB-r (resistant to λ using LamB receptor), and λall-r (resistant to λ using LamB and OmpF receptors).

**Genetic polymorphism for mannose utlization.** We used tetrazolium mannose (TMan) agar plates to score Man+ and Man– cells in samples taken on day 20 from the 96 populations in the large-scale experiment. Man+ and Man– cells produce white and red colonies, respectively, on TMan plates. The ancestral strain is Man+. Man– cells that have been sequenced (Table S5) have mutations in the *manXYZ* operon that confer resistance to all λ phage, including those that evolved the ability to use the OmpF receptor. The vast majority of populations were genetically polymorphic for mannose use (Table S6) and, by extension, for *manXYZ*-mediated resistance to λ phage.

**Table S6.** Frequencies of Man– mutants in 96 bacterial populations on the last day of the large-scale experiment, with the community ID, number (n) of cells scored, and frequency of Man– cells shown for each population.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ID | Number (n) | Frequency of *man*– | ID | Number (n) | Frequency of *man*– | ID | Number (n) | Frequency of *man*– |
| A1 | 181 | 0.03 | C9 | 96 | 0.73 | F5 | 45 | 0.00 |
| A2 | 40 | 0.25 | C10 | 34 | 0.94 | F6 | 84 | 0.10 |
| A3 | 29 | 0.10 | C11 | 77 | 0.21 | F7 | 75 | 0.01 |
| A4 | 37 | 0.78 | C12 | 22 | 0.95 | F8 | 136 | 0.16 |
| A5 | 57 | 0.26 | D1 | 220 | 0.71 | F9 | 36 | 0.83 |
| A6 | 38 | 0.47 | D2 | 60 | 0.25 | F10 | 37 | 0.30 |
| A7 | 126 | 0.02 | D3 | 123 | 0.07 | F11 | 87 | 0.17 |
| A8 | 54 | 0.94 | D4 | 30 | 0.00 | F12 | 107 | 0.90 |
| A9 | 32 | 0.00 | D5 | 61 | 0.26 | G1 | 64 | 0.27 |
| A10 | 71 | 0.15 | D6 | 48 | 0.00 | G2 | 95 | 0.48 |
| A11 | 25 | 0.04 | D7 | 210 | 0.95 | G3 | 85 | 0.16 |
| A12 | 8 | 0.00 | D8 | 46 | 0.09 | G4 | 85 | 0.33 |
| B1 | 29 | 0.28 | D9 | 152 | 0.98 | G5 | 95 | 0.15 |
| B2 | 44 | 0.34 | D10 | 28 | 0.61 | G6 | 65 | 0.26 |
| B3 | 68 | 0.03 | D11 | 148 | 0.17 | G7 | 116 | 0.30 |
| B4 | 22 | 0.00 | D12 | 43 | 0.93 | G8 | 64 | 0.16 |
| B5 | 132 | 0.31 | E1 | 59 | 0.88 | G9 | 74 | 0.18 |
| B6 | 31 | 0.00 | E2 | 78 | 0.27 | G10 | 77 | 0.48 |
| B7 | 115 | 0.12 | E3 | 35 | 1.00 | G11 | 110 | 0.14 |
| B8 | 110 | 0.95 | E4 | 32 | 0.00 | G12 | 79 | 0.49 |
| B9 | 32 | 1.00 | E5 | 29 | 0.00 | H1 | 129 | 0.33 |
| B10 | 70 | 0.16 | E6 | 48 | 1.00 | H2 | 68 | 0.07 |
| B11 | 41 | 0.02 | E7 | 34 | 0.74 | H3 | 80 | 0.98 |
| B12 | 76 | 0.00 | E8 | 51 | 0.88 | H4 | 142 | 0.10 |
| C1 | 77 | 0.00 | E9 | 69 | 0.35 | H5 | 131 | 0.48 |
| C2 | 157 | 0.07 | E10 | 64 | 0.00 | H6 | 98 | 0.29 |
| C3 | 30 | 0.00 | E11 | 120 | 0.02 | H7 | 75 | 0.00 |
| C4 | 56 | 0.09 | E12 | 38 | 0.92 | H8 | 69 | 0.88 |
| C5 | 101 | 0.12 | F1 | 95 | 0.01 | H9 | 126 | 0.10 |
| C6 | 100 | 0.14 | F2 | 79 | 0.06 | H10 | 142 | 0.02 |
| C7 | 118 | 0.13 | F3 | 45 | 0.00 | H11 | 38 | 0.00 |
| C8 | 138 | 0.07 | F4 | 89 | 0.46 | H12 | 30 | 0.80 |

**References**

1. J. E. Barrick et al., *Nature* **461**, 1243 (2009).

2. P. Daegelen, F. W. Studier, R. E. Lenski, S. Cure, J. F. Kim, *J. Mol. Biol.* **394**, 634 (2009).

3. E. Spanakis, M. T. Horne, *J. Gen. Microbiol.* **133**, 353 (1987).

4. R. W. Hendrix, Ed., *Lambda II* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1983).

5. J. Sambrook, D. W. Russell, *Molecular Cloning, Third Ed.* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2001).

6. M. H. Adams, *Bacteriophages* (Interscience Publishers, New York, NY, 1959).

7. T. Baba et al., *Mol. Syst. Biol.* **2**, 2006.0008 (2006).

8. H. A. Shuman, T. J. Silhavy, *Nat. Rev. Genet.* **4**, 419 (2003).

9. W. Boos, A. Bohm, *Trends Genet.* **16**, 404 (2000).

10. J. R. Meyer et al., *Evolution* **64**, 3024 (2010).

11. L. Pelosi et al., *Genetics* **173**, 1851 (2006).

12. B. J. M. Bohannan, R. E. Lenski, *Ecology* **78**, 2303 (1997).

13. C. Braunbreton, M. Hofnung, *Mol. Gen. Genet.* **159**, 143 (1978).

14. B. R. Levin, F. M. Stewart, L. Chao, *Am. Nat.* **111**, 3 (1977).

15. R. E. Lenski, B. R. Levin, *Am. Nat.* **125**, 585 (1985).

16. M. Schwartz, in *Virus Receptors, Vol. 1*, K. Lonberg-Holm, L. Philipson, Eds. (Chapman and Hall, London, UK, 1980).

17. G. Kefala et al., *Protein Sci.* **19**, 1117 (2010).

18. T. F. Cooper, D. E. Rozen, R. E. Lenski, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 1072 (2003).

19. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, *J. Mol. Biol.* **215**, 403 (1990).

20. Y. P. Shao, I. N. Wang, *Genetics* **180**, 471 (2008).

21. M. De Paepe, F. Taddei, *PLoS Biology* **4**, 1248 (2006)

22. H. D. Rundle, L. Nagel, J. W. Boughman, D. Schluter, *Science* **287**, 306 (2000).

23. J. B. Losos, *Am. Nat.* **175**, 623 (2010).

24. Y. F. Chan et al., *Mech. Dev.* **126**, S14 (2009).

25. E. A. Ostrowski, R. J. Woods, R. E. Lenski, *Proc. R. Soc. B* **275,** 277 (2008).

26. M. Manceau, V. S. Domingues, C. R. Linnen, E. B. Rosenblum, H. E. Hoekstra, *Philos. Trans. R. Soc. B* *365*, 2439 (2010).

27. H. A. Wichman, J. Wichman, J. J. Bull, *Genetics* **170**, 19 (2005).