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

Justin R. Meyer[1,2*], Devin T. Dobias[3], Joshua S. Weitz[4], Jeffrey E. Barrick[2,5], Ryan T. Quick[6], Richard E. Lenski[1,2,6]

1 Department of Zoology, Michigan State University, East Lansing, MI 48824, USA.

2 BEACON Center for the Study of Evolution in Action, Michigan State University, East Lansing, MI 48824, USA.

3 Department of Biology, Washington University, St. Louis, MO 63130, USA

4 School of Biology and School of Physics, Georgia Institute of Technology, Atlanta, GA 30332, USA.

5 Department of Chemistry and Biochemistry, Institute for Cellular and Molecular Biology, The University of Texas, Austin, TX 78712, USA.

6 Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, MI 48824, USA.

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

One-sentence summary: Phage λ evolved to use a new receptor, but whether the required mutations arose depended on the coevolution of its host.

The processes responsible for the evolution of key innovations, whereby lineages acquire qualitatively new functions that expand their ecological opportunities, remain poorly understood. We examined how a virus, bacteriophage λ, evolved to infect its host, *Escherichia coli*, through a novel pathway. Natural selection promoted the fixation of mutations in the virus's host-recognition protein, J, that improved fitness on the original receptor, LamB, and set the stage for other mutations that allowed infection through a new receptor, OmpF. These viral mutations arose only after the host evolved reduced expression of LamB, whereas certain other host mutations prevented the phage from evolving the new function. This study shows the complex interplay between genomic processes and ecological conditions that favor the emergence of evolutionary innovations.

Throughout the history of life, evolving lineages have acquired qualitatively new functions that enable organisms to expand their ecological opportunities and, in many cases, undergo further diversification (*1*). Explaining how these transitions have occurred is usually difficult, both because the responsible events typically occurred in the distant past and because their rarity suggests that they might involve atypical evolutionary processes. For example, natural selection is critical for the process of adaptation, yet its role in producing key innovations is less clear because, by fixing variants that improve existing functions, selection might strand populations on local adaptive peaks and thereby prevent them from discovering new functions (*2*). Darwin was well aware of the difference between improving an existing trait and evolving a new one (*3*), and he reasoned that new traits originate by co-opting previously existing structures and functions. Without an understanding of genetic mechanisms, however, he could not provide a detailed account of how this process happens. Since then, others have proposed more explicit models of the origin of new functions that vary in two main respects: the structure of the adaptive landscape – including its dimensionality (*2*), genetic connectivity (*4*), and fluctuations caused by changing environments (*5*) including interactions with coevolving species (*6*) – and the relative importance of natural selection and random drift (*2, 4, 7*). However, no consensus has been reached owing, at least in part, to the paucity of cases with sufficient genetic and ecological data (*8*).

To that end, we examine the evolutionary forces responsible for the emergence of a novel trait in a microbial system, including data bearing on the genetic architecture of the adaptive landscape in which the novel capacity arose. Microbes are well suited for such research because their evolution can be observed in real time, experiments are easily replicated, and transitional states can be studied by reviving samples stored at different times during an experiment (*9, 10*). We investigated how a virus evolved the ability to infect its host through a new receptor that the

ancestral virus cannot use. We tested competing hypotheses about the evolution of this new trait by determining the conditions that promoted its evolution, the mutations that conferred the new function, and the evolutionary forces that drove its emergence.

Study system. Viruses are genetically and morphologically diverse, and they infect all groups of organisms (*11, 12*). Viruses initiate infections by binding to receptors on the surface of host cells. The physicochemical properties of viral ligands determine which receptors they target and thereby influence the host range and ecological niche of the virus (*13*). Mutations in viral genes that encode the production of ligands can cause shifts in host range and thus are often associated with emerging diseases (*14, 15*). The evolution of the ability to infect through a new receptor represents a key innovation for a virus.

The virus we studied, a strictly lytic derivative of phage λ called cI26 (*16*, table S1), is only known to infect one bacterial species, *Escherichia coli*, and has a specialized ligand, the J protein, at the end of its tail (*17*). J targets a single protein, LamB, on the *E. coli* outer membrane (*17, 18*). Phage λ requires only LamB to attach (*19*), and LamB is the only outer-membrane protein that affects λ reproduction (*20*).

Given interest in the fundamental question of how organisms evolve novel traits and in the practical problem of how emerging pathogens evolve to target new host receptors, we sought to determine whether λ could evolve to infect through an alternative receptor. Phage λ is well studied and amenable to experimentation (*21*). A related phage (Ur-λ) possesses side-tail fibers and can infect *E. coli* through a second receptor OmpC (*22, 23*), suggesting other receptors might be accessible to evolving λ populations. Moreover, we identified conditions that seemed suitable for promoting the use of a novel receptor. In particular, when *E. coli* and λ were cultured together in a glucose-limited environment, the bacteria evolved resistance by mutations in *malT*,

which interfered with its role as a transcription factor that promotes *lamB* expression (*23*). The mutants arose and fixed within 8 days, generating highly resistant populations (fig. S1), although the phage did not go extinct but instead persisted at densities of about 10^6 to 10^7 phage per ml, or about one phage per 10^2 to 10^3 host cells (fig. S2). The phage evidently persisted on a subpopulation of cells that, despite their *malT* mutations, experienced spontaneous induction of the LamB protein (*24*). This explanation was confirmed by showing that phage were also sustained when they were grown on a *malT* mutant, whereas the phage went extinct when cultured with a *lamB* mutant that lacks the potential to produce the LamB protein (fig. S3). We reasoned that mutant viruses able to infect through some protein other than LamB would be favored after *malT* mutants arose because they could infect the entire host population rather than a small minority of cells.

Initial evolution experiment. We co-cultured a virulent (non-lysogenic) derivative of λ and *E. coli* B in 10 ml of a minimal glucose medium in six replicate flasks for 28 days with daily transfers of 1% of each community into a flask containing fresh medium, and we preserved samples weekly by freezing 10% of the mixture (*16*). We tested whether the phage could infect cells through a new receptor by taking samples of the phage populations (typically ~5 x 10^4 virions) and inoculating them onto the surface of agar plates infused with a *lamB* mutant that does not produce the LamB receptor. A spot of clearing (lysis of host cells) provided evidence that some phage had evolved the ability to infect through a receptor other than LamB (Fig. 1, panels A-B). Such spots were observed in only one population, but this ability evolved quickly, such that ~0.01% of the phage could infect the *lamB*-negative mutants by day 8, including the isolate designated EvoC, and the majority did so by day 15.

Identification of the novel receptor. We used seven knockout strains (derivatives of K12 BW25113), each missing a gene encoding a different outer-membrane protein, to identify the new receptor (*16*, *25*). We then introduced *malT* mutations to these strains so that they also would not express the native LamB receptor. We inferred the new receptor by testing the ancestral and evolved λ against these double mutants to see which ones were resistant to the various phage isolates. The only host that was resistant to the EvoC isolate was the *ompF malT* double mutant that lacked expression of both OmpF and LamB (Fig. 1, table S2), indicating that OmpF was the new receptor. This evolved phage could still infect the host strain expressing LamB but not OmpF, indicating that the phage retained the ability to infect through its native receptor (Fig. 1).

Both LamB and OmpF form trimeric porins composed of three identical β barrels (*26*, *27*). This overall structure is probably essential for the J protein in the λ tail to bind because J, too, forms a trimer and is thought to attach with radial symmetry across the three pore domains (*28*). Although OmpF has the most similar crystal structure to LamB of any *E. coli* protein determined to date *(29)*, they are not the most similar pair by amino-acid sequence (table S3). This discordance suggests that the overall structure is at least as important for λ binding as the identity of specific amino-acid residues. Also, OmpF is the sole major porin in the *E. coli* B strain used in this study, and B expresses it constitutively during growth (*30*, *31*). Hence, OmpF provided a substantial ecological opportunity to phage that evolved the ability to target it.

Genome evolution. We sequenced the genome of the evolved phage EvoC in order to identify the mutations that allowed it to use the OmpF receptor (*16*). There were five mutations in total, and all of them were in the J gene (Fig. 2). Targeted sequencing of J showed that a single substitution (A to G at position 3034) differentiated EvoC from another evolved isolate from the

same time point, EvoA, that could use only the ancestral LamB receptor, indicating that mutation contributed to the new receptor function. Another LamB-dependent phage from the same day, EvoB, differed from EvoC at five sites in the J protein.

Large-scale evolution experiment. We repeated our first experiment with 96 more communities to identify general principles of how λ evolves the capacity to target an alternative receptor (*16*). We sampled daily for finer resolution of the evolutionary dynamics. As before, only some phage populations (24/96) evolved the ability to use a second receptor. This ability emerged about the same time (median 12 days; range 9-17 days; fig. S4), and all isolates with altered receptor function infected hosts through the OmpF protein.

Parallel molecular evolution. We sequenced J alleles from 24 phage isolates that independently evolved the ability to target OmpF during the large-scale evolution experiment to determine whether the mutation at position 3034 or any others were required to use that receptor. The isolates were taken the same day the new function was detected (*16*). For comparison, we sequenced phage from 24 populations that never evolved this trait; these isolates were sampled on the same days as those that evolved the new trait, so that the elapsed times were the same.

In total, there were 241 single nucleotide polymorphisms (SNPs) across the 48 J alleles, but no insertions or deletions (Fig. 3). However, there were only 40 unique mutations because many arose repeatedly in replicate populations. Moreover, all of them were non-synonymous. The alleles for phage able to target OmpF had on average 6.63 (± 0.51 95% confidence interval (CI)) SNPs, whereas the phage that required LamB had only 3.42 (± 0.50 95% CI) SNPs. This difference is highly significant based on a paired comparison between the two types of phage matched for the day of their isolation and, in the case of multiple equivalent pairs, matched arbitrarily by position in the experiment (t_s = 9.144, 23 d.f, two-tailed $p < 0.0001$; table S4).

Also, across both classes of phage, over 97% of the mutations were in the last 25% of the protein (C-terminal end), the region known to interact with LamB (*32*).

There are four striking cases of parallel evolution of the J protein in the phage that target OmpF. In two cases, the mutations were identical across all 24 populations, while in two others there were slight variations (Fig. 3). In particular, all J alleles from phage able to infect through OmpF had the A-to-G mutation at nucleotide position 3034 and G-to-A mutation at position 3319. Also, all of them had a mutation at either position 3320 or 3321, affecting the same codon (amino-acid residue 1107) as the mutation at position 3319. Finally, all J alleles had at least one mutation between positions 2969 and 2999 (amino-acid residues 990 to 1000).

Each of these mutations or classes of mutation was also found in at least one of the phage that retained the ancestral host-range, although none of them had all four together (Fig. 3). Two LamB-dependent isolates, F2 and H4, had three of the mutations, as did EvoA from the initial experiment (Fig. 2), yet none produced clearing on lawns of *lamB* mutants.

The correspondence between the use of the OmpF receptor and the presence of these four mutations, coupled with the observation that phage having only three of the four cannot use OmpF, provides evidence that all four are required for λ to infect through OmpF. We performed two additional assays to confirm that only phage with all four mutations can infect *lamB* mutants (*16*). The assays were performed using isolates EvoA, F2, and H4 that each had three of the four canonical mutations and D7 that had all four and no others. Only D7 exhibited a measureable adsorption rate on *lamB* mutant cells (fig. S5), and it was also the only one that reproduced on *lamB* mutants in the medium used in the evolution experiments (fig. S6). These findings indicate an “all-or-none” form of epistasis among the four mutations responsible for the novel receptor phenotype.

Role of natural selection. In the λ population that evolved to use OmpF in the initial experiment, the A3034G mutation was the fourth and final step, and it was clearly advantageous because it conferred the ability to infect the entire cell population. However, the all-or-none epistasis among the mutations means that selection for that new capacity *per se* was not responsible for the rise of the three prior mutations. Nonetheless, there are several lines of evidence that selection drove their rise. First, all 248 independent mutations in the 51 sequenced J alleles were non-synonymous, whereas the expected ratio of non-synonymous to synonymous changes is 3.19:1 under the null model for the ancestral J sequence (*16*). This great excess is evident even if we include only the 82 non-synonymous mutations in the 24 isolates that did not evolve the new receptor function. Second, the mutations are highly concentrated in the region of the J protein that interacts directly with LamB (*18*). Third, there was parallel evolution at the genetic level across the populations. For those phage that evolved to exploit OmpF, an average of 61% (4.06/6.63) of mutations were shared across independently derived pairs (fig. S7), which greatly exceeds the fraction expected under a conservative randomization test (*16*) using only the variable sites in J ($p < 10^{-5}$). Pairs of phage that remained dependent on LamB shared on average 17% (0.58/3.42) of their mutations (fig. S7), and this fraction is again significant under the same test ($p < 10^{-5}$). Thus, it is clear that selection acted on the J protein even before the new capacity evolved. This selection presumably improved the interaction of the phage tail fibers with LamB.

Stochasicity and contingency. All of the λ populations had the same ecological incentive to exploit an alternative receptor, but only some evolved that ability. Why were some populations successful and others not? One possibility is that all of them would eventually have evolved that function, but there was insufficient time to do so. This explanation is consistent with the facts that the LamB-dependent isolates had fewer mutations than those able to target OmpF,

and that the two groups shared many mutations. Alternatively, the evolution of the new receptor function might have been contingent on earlier events (*33*-*35*), such that particular changes in the phage or the host promoted or impeded the subsequent evolution of that function. To test these hypotheses, we replayed evolution (*35*, *36*) using various combinations of phage and bacteria.

First, we tested whether certain mutations in the phage that might enhance performance on LamB would impede the evolution of the new OmpF function. We inoculated flasks with the ancestral bacteria and one of six phage isolates. Three of the six phage isolates had different sets of three mutations that were present in multiple isolates that evolved to use the OmpF receptor, including one, zero, and two of the four canonical mutations. The other three isolates had three, two, and one mutations that were not observed in any isolate that previously evolved the ability to use OmpF; these isolates also each had one of the canonical mutations. The first set provided a positive control; the second set had candidate mutations for impeding the evolution of the new function. For each phage, we propagated 12 communities for 10 days and surveyed daily the phage's ability to lyse *lamB* mutant cells. There were as many or more successes in evolving the new function among the three phages that had the potentially interfering mutations as among the positive controls (Fig. 4A). This experiment thus provides no evidence that some phage failed to evolve the new function because they had mutations that prevented them from doing so.

Next, we asked if the outcome was contingent on mutations in the evolving bacteria. To that end, we performed a similar replay experiment except that the initial phage type was held constant while the starting bacterial isolate was varied. For the phage, we used EvoA, an isolate that was one mutation away from using OmpF (Fig. 2). For the bacteria, we used six clones: three from communities where λ evolved to use OmpF, and three where the phage retained their dependence on LamB. We observed a striking "all-or-none" pattern of outcomes, though not in

accordance with our categories (Fig. 4B). In particular, all 36 λ populations evolved the final mutation required to use OmpF in communities with three bacterial clones whereas none of the phage evolved that ability in 36 communities with the three other clones. For two of the latter class (EcA8 and EcF6), the phage were unable to reproduce and went extinct; in the other case (EcC3), phage persisted but none of the replays yielded phage able to use OmpF. It is clear that bacterial characteristics determined whether the phage would evolve the new receptor function. However, the bacteria that promoted that outcome did not necessarily come from communities in which λ had previously evolved to exploit OmpF.

We sequenced the full genomes of these six bacteria to identify the mutations responsible for the differences in phage evolution. The six genomes harbor a total of 15 mutations (table S5). Five have similar deletions that impact the *rbs* operon, which confers the ability to grow on ribose; previous work has shown that these deletions occur at an unusually high rate owing to a nearby insertion-sequence element in the ancestral strain (*37*). All 10 other mutations are directly related to the interaction with λ. As expected, all six genomes have mutations in *malT* that confer resistance to the ancestral phage, as described earlier. One genome from a community where λ evolved to use the OmpF receptor has a non-synonymous mutation in the *ompF* gene, which might confer partial resistance to the evolved phage. The three remaining mutations disrupt *manY* or *manZ*, and they uniquely differentiate the three strains that prevented phage from evolving to use OmpF from the three strains that allowed that change (table S5). The *manY* and *manZ* genes encode the transmembrane channel of the ManXYZ mannose permease, which is required for λ DNA to cross the inner membrane (*38-40*). These mutations thus confer resistance by blocking a later step during infection, and they would render ineffective any phage mutations that altered the receptor function. Therefore, the evolution of phage that target OmpF

is promoted by bacterial mutations in *malT* but impeded by mutations in *manYZ*, indicating contingency dependent on the host-parasite coevolution.

After discovering the *manYZ* mutations, we screened all 96 bacterial populations from the large-scale experiment to determine how many harbored mannose-deficient mutants that would block λ infections (*16*). At least 80 populations (table S6) had such mutations, including many from communities in which λ evolved the ability to exploit the OmpF receptor. However, these mutations rarely fixed; instead, susceptible subpopulations persisted in 77 of the 80 communities that allowed the phage to continue to evolve. This finding suggests a complex interplay between coevolving phage and bacteria, one that depends on the entire community and its diversity. To test this hypothesis, we repeated the second replay experiment using the same phage and bacteria except with bacterial communities instead of clones (*16*). Once again, some bacteria consistently impeded the evolution of phage that used OmpF while others consistently promoted that change (Fig. 4C). Moreover, those outcomes differed for one clone (EcA8, Fig. 4B) and its community (ComA8, Fig. 4C), confirming the effect of bacterial diversity on the phage's evolution.

Repeatability, contingency, and the evolution of a key innovation. Phage λ often, but not always, evolved the ability to infect its *E. coli* hosts by targeting a new receptor, OmpF. Figure 5 summarizes the important steps in this process, including some that promoted this key innovation and others that impeded it. The fact that several mutations are required for λ to use OmpF may explain why no previous studies have reported this change, despite decades of intense study of this phage. However, by running evolution experiments rather than mutational screens, we observed 25 cases in which this new function evolved. Our experiments are not the first to demonstrate evolutionary transitions in viruses (*13, 41-43)*; one study found that a single mutation allowed phage φ6 to infect a new host species (*13*), while a study of phage SBW25Φ2

showed that coevolution experiments were more effective at producing host-range shifts than were screens using new hosts (*43*). The rich and complex dynamics of coevolving species may thus sometimes facilitate key innovations (*6, 44*).

More generally, our study shows some of the challenges that make it difficult to observe and explain the origin of many new functions including the requirement for multiple mutations, the complex interactions of mutations within and between species, and the resulting historical contingency that can enable or impede the outcome of interest depending on the order in which mutations occur. The "all-or-none" epistasis among the four canonical phage mutations implies that it would have been unlikely for the new function to evolve on the scale of our experiments, except for the lucky fact that some of the mutations were beneficial to the phage in performing their current function, thereby pushing evolution toward the new function. The mutations in the bacteria, and how they influenced the phage's evolution, were also important. In particular, the initial resistance mutation generated a physiological subpopulation of hosts that allowed the phage to persist and adapt to the original receptor, thereby accumulating the required mutations. Yet, as fortuitous as these circumstances were, another mutation could – and often did – derail the emergence of the new function: namely, a mutation that conferred an alternative mode of host resistance eliminated the advantage to the phage of targeting the OmpF receptor. The interactions between bacteria and phage, which contributed so much to the development of microbial genetics and molecular biology (*45, 46*), continue to serve as powerful models to study ecology and evolution (*47, 48*).

References and Notes

1. E. Mayr, *Animal Species and Evolution* (Belknap Press, Cambridge, MA, 1963).
2. S. Gavrilets, in *Toward an Extended Evolutionary Synthesis*, M. Pigliucci, G. Muller, Eds. (MIT Press, Cambridge, MA, 2010).

3. C. Darwin, *The Origin of Species* (Murray, London, 1859).
4. R. A. Fisher, *The Genetical Theory of Natural Selection* (Clarendon Press, Oxford, 1930).
5. U. Dieckmann, M. Doebeli, J. A. J. Metz, Eds., *Adaptive Speciation* (Cambridge Univ. Press, Cambridge, UK, 2004).
6. J. N. Thompson, *The Geographic Mosaic of Coevolution* (Univ. Chicago Press, Chicago, 2005).
7. S. Wright, *Genetics* **16**, 97 (1931).
8. M. Pigliucci, *Philos. Sci.* **75**, 887 (2008).
9. J. J. Dennehy, *Trends Microbiol.* **17**, 450 (2009).
10. S. F. Elena, R. E. Lenski, *Nat. Rev. Genet.* **4**, 457 (2003).
11. R. A. Edwards, F. Rohwer, *Nat. Rev. Microbiol.* **3**, 504 (2005).
12. V. Torsvik, L. Ovreas, T. F. Thingstad, *Science* **296**, 1064 (2002).
13. S. Duffy, C. L. Burch, P. E. Turner, *Evolution* **61**, 2614 (2007).
14. K. M. Pepin, S. Lass, J. R. C. Pulliam, A. F. Read, J. O. Lloyd-Smith, *Nat. Rev. Microbiol.* **8**, 802 (2010).
15. E. C. Holmes, *Annu. Rev. Ecol. Evol. Syst.* **40**, 353 (2009).
16. Materials and methods are available as supporting material on *Science* Online.
17. I. Katsura, in *Lambda II*, R. W. Hendrix, Ed. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1981).
18. E. Berkane et al., *Biochem.* **45**, 2708 (2006).
19. P. Grayson, L. Han, T. Winther, R. Phillips, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 14652 (2007).
20. N. D. Maynard et al., *PLoS Genet* **6**, e1001017 (2010).
21. R. W. Hendrix, Ed., *Lambda II* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1983).
22. R. W. Hendrix, R. L. Duda, *Science* **258**, 1145 (1992).
23. R. Gallet, S. Kannoly, I. Wang, *BMC Microbiol.* **11**, 181 (2011).
24. C. Braunbreton, M. Hofnung, *Mol. Gen. Genet.* **159**, 143 (1978).
25. T. Baba et al., *Mol. Syst. Biol.* **2**, 2006.0008 (2006).
26. T. Schirmer, T. A. Keller, Y. F. Wang, J. P. Rosenbusch, *Science* **267**, 512 (1995).
27. G. Kefala et al., *Protein Sci.* **19**, 1117 (2010).
28. P. A. Gurnev, A. B. Oppenheim, M. Winterhalter, S. M. Bezrukov, *J. Mol. Biol.* **359**, 1447 (2006).
29. R. Koebnik, K. P. Locher, P. Van Gelder, *Mol. Microbiol.* **37**, 239 (2000).
30. D. Schneider et al., *BMC Microbiol.* **2**, 18 (2002).
31. E. Crozat et al., *J. Bacteriol.* **193**, 429 (2011).
32. J. Wang, M. Hofnung, A. Charbit, *J. Bacteriol.* **182**, 508 (2000).
33. S. J. Gould, *Wonderful Life* (Norton, New York, 1989).
34. M. Travisano, J. A. Mongold, A. F. Bennett, R. E. Lenski, *Science* **267**, 87 (1995).
35. Z. D. Blount, C. Z. Borland, R. E. Lenski, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 7899 (2008).
36. R. J. Woods et al., *Science* **331**, 1433 (2011).
37. V. S. Cooper, D. Schneider, M. Blot, R. E. Lenski, *J. Bacteriol.* **183**, 2834 (2001).
38. B. Erni, B. Zanolari, H. P. Kocher, *J. Biol. Chem.* **262**, 5238 (1987).
39. P. E. J. Saris, E. T. Palva, *FEMS Microbiol. Lett.* **44**, 371 (1987).
40. M. Esquinas-Rychen, B. Erni, *J. Mol. Microbiol. Biotechnol.* **3**, 361 (2001).
41. R. L. Graham, R. S. Baric, *J. Virol.* **84**, 3134 (2010).

42. M. T. Ferris, P. Joyce, C. L. Burch, *Genetics* **176**, 1013 (2007).
43. A. R. Hall, P. D. Scanlan, A. Buckling, *Am. Nat.* **177**, 44 (2011).
44. J. N. Thompson, *The Coevolutionary Process* (Univ. Chicago Press, Chicago, IL, 1994).
45. D. I. Friedman, D. L. Court, *Curr. Opin. Microbiol.* **4**, 201 (2001).
46. J. Cairns, G. S. Stent, J. D. Watson, Eds., *Phage and the Origins of Molecular Biology* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2007).
47. B. J. M. Bohannan, R. E. Lenski, *Ecol. Lett.* **3**, 362 (2000).
48. C. Flores, J. R. Meyer, S. Valverde, L. Farr, J. S. Weitz, *Proc. Natl. Acad. Sci. U.S.A.* **108**, E288 (2011).

Acknowledgments: This work was supported by the BEACON Center for the Study of Evolution in Action (National Science Foundation Cooperative Agreement DBI-0939454), the Defense Advanced Research Projects Agency (HR0011-09-1-0055 to R.E.L. and J.S.W.), and the National Institutes of Health (R00GM087550 to J.E.B.). J.S.W. also acknowledges the support of the James S. McDonnell Foundation and the Burroughs Wellcome Fund. We thank N. Hajela and L. Zaman for help in the lab, D. Court for sharing phage cI26, the Michigan State University Research Technology Support Facility for help with genome sequencing, and Z. Blount, T. Ferenci, C. Marx, and L. Zaman for discussions. R.E.L. will make the evolved strains available to qualified recipients, subject to completion of a material transfer agreement that can be found at http://technologies.msu.edu/forms.html. The genome-sequence data for the phage and bacteria have been deposited in the NCBI Sequence Read Archive (SRA043942). Other data and analysis scripts have been deposited at the Dryad Digital Repository (doi:10.5061/dryad.pending).

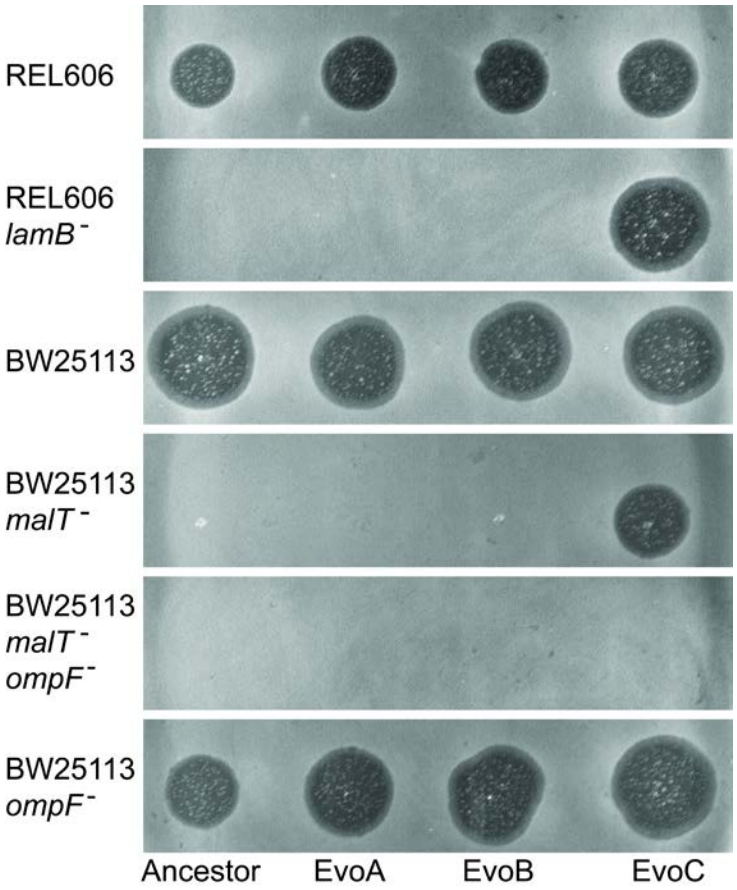

Fig. 1. Infection assay for four λ isolates tested on six *E. coli* strains. Each panel shows a bacterial lawn with aliquots of the phage applied to it; darker regions indicate successful infections that clear the lawn. The phage isolates include the ancestor and three clones isolated from the same population on day 8 of the initial evolution experiment including one, EvoC, that can use the OmpF receptor. The bacterial strains include mutants that differ in the expression of LamB and OmpF porins on two genomic backgrounds, REL606 (the ancestral strain in the evolution experiments) and BW25113 (a derivative of K12). The *malT*$^{-}$ strains do not express LamB at appreciable levels.

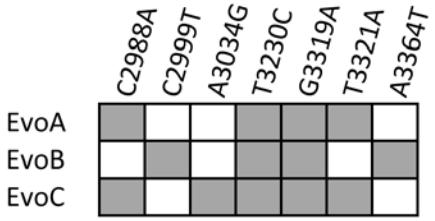

Fig. 2. Mutations in the λ gene encoding the J protein in three isolates from the same population on day 8 of the initial evolution experiment. The isolates are shown in rows and the mutations in columns, with the first letter being the ancestral nucleotide, the number the nucleotide position, and the last letter the evolved nucleotide. The gray fill indicates that the phage isolate has the corresponding mutation.

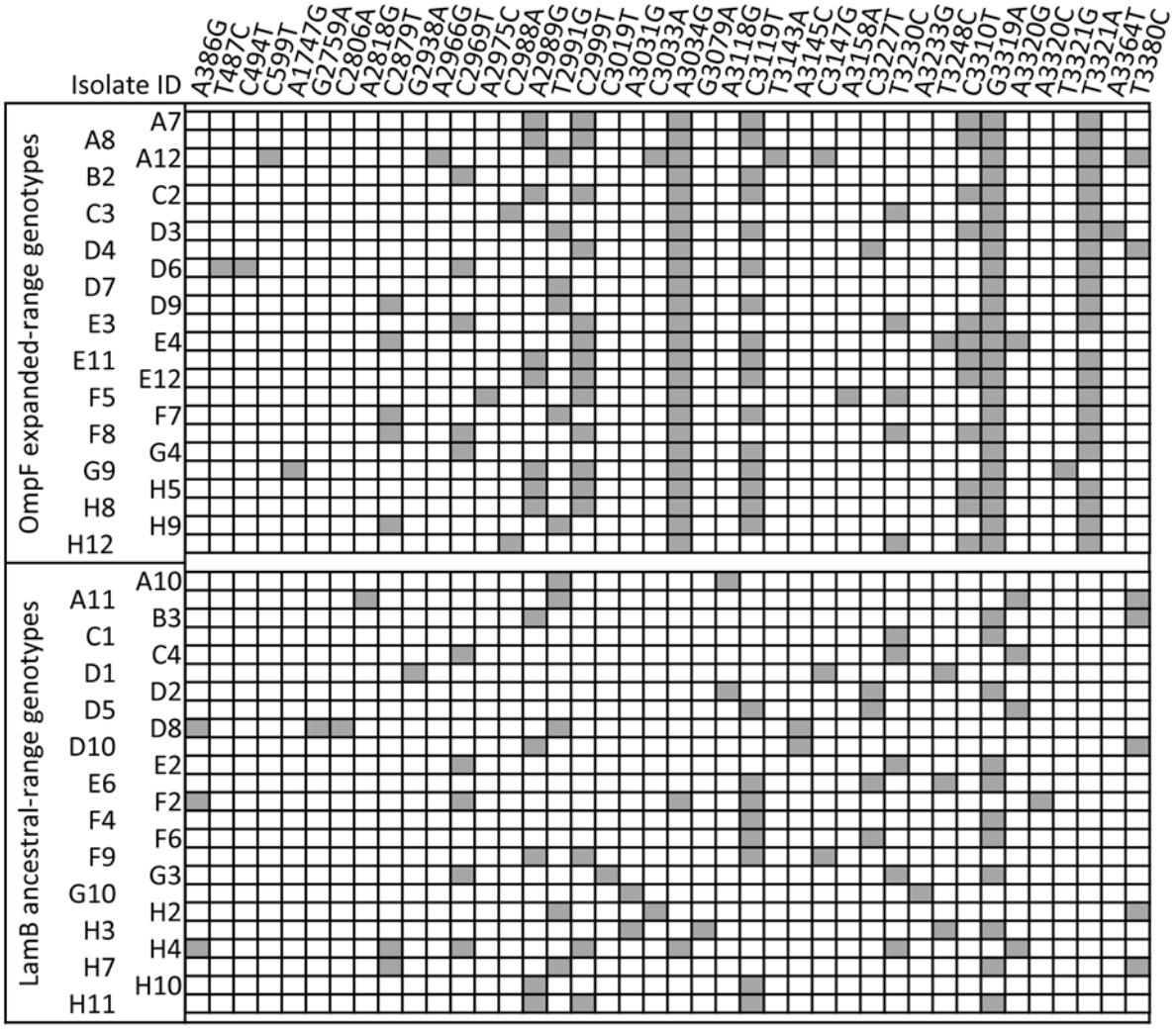

Fig. 3. Mutations affecting the J protein in phage isolates from 48 independent populations of the large-scale experiment. Isolates are shown in rows (with alternate labels offset for readability) and mutations in columns; gray fill indicates an isolate has the mutation. The top 24 rows show phage isolates that can target the new OmpF receptor; the bottom 24 rows show phage that remain dependent on LamB.

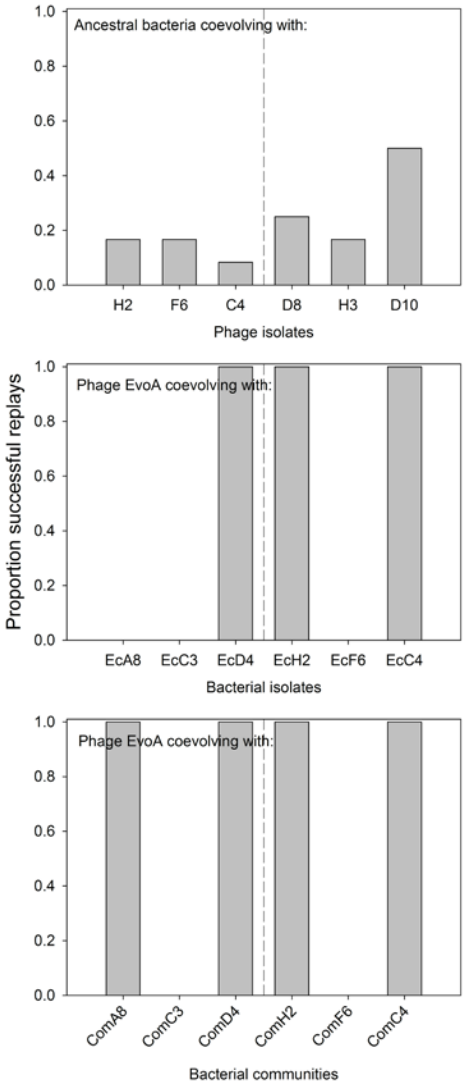

Fig. 4. Replay experiments using different combinations of phage and bacteria. For each panel, the y-axis shows the proportion of replicate replays that produced phage able to target the new OmpF receptor. Top panel: Replays were initiated with the ancestral bacteria and six phage isolates. Each combination was replicated 12-fold. Three of the phage (H2, F6, and C4) had mutations shared by multiple lineages that evolved the capacity to target OmpF in the large-scale experiment. The other three (D8, H3, and D10) had mutations that were never observed in phage that targeted OmpF. The latter mutations were candidates for impeding the evolution of the new function, but that hypothesis was not supported. Middle panel: Replays were initiated with phage EvoA (which needs only one more mutation to use OmpF) and six bacterial clones. Each combination was replicated 4-fold. Three clones (EcA8, EcC3, and EcD4) were isolated from flasks in which phage evolved the capacity to target OmpF in the large-scale experiment. The other three (EcH2, EcF6, and EcC4) came from flasks in which phage did not evolve that function. The replay outcomes did not support these categories, but sequencing the bacterial genomes identified mutations that uniquely determined whether the phage would evolve the OmpF function. See text for details. Bottom panel: Replays were initiated using the same phage and bacteria used in the middle panel, except with full bacterial communities rather than individual clones. Each combination was replicated 12-fold. The different outcomes for one bacterial clone (EcA8, middle) and its source community (ComA8, bottom) show the effect of bacterial diversity on phage evolution.

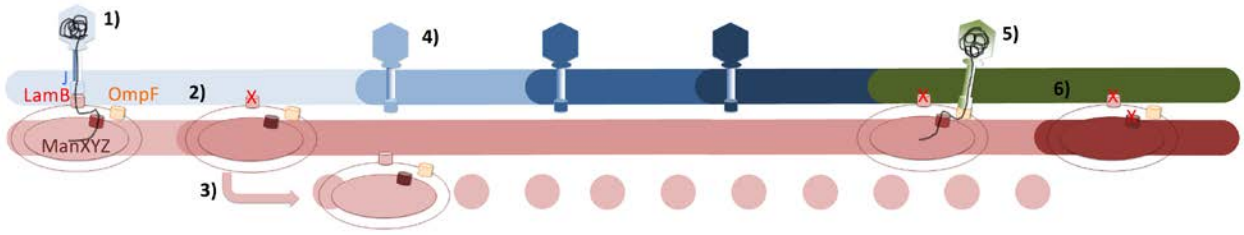

Fig. 5. Steps in the coevolution of phage λ and its *E. coli* host leading to the phage's ability to target a new receptor, OmpF. (1) The ancestral phage targets the LamB porin using the J protein and injects its DNA into the periplasm, then the DNA is transported into the cytoplasm via the ManXYZ permease. (2) The bacteria evolve resistance by mutations in *malT*, a positive regulator of LamB expression. (3) However, spontaneous inductions of LamB generate a subpopulation of phenotypically sensitive cells that can sustain the phage population. (4) The phage evolves mutations in the J protein that improve performance on the LamB receptor. Some of these mutations are also required for the phage to target OmpF. (5) The phage eventually evolves the four mutations that enable it to use OmpF. (6) However, the bacteria may evolve additional resistance by mutations in *manY* or *manZ* that prevent transport of the phage DNA into the cytoplasm. When these mutants become sufficiently common, there is little or no benefit to mutant phage that can use OmpF.

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

Strains. We studied *Escherichia coli* B strain REL606 because its evolution in the laboratory is well documented (*49*), its genome has been sequenced (*50*), 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 (*50*, *51*).

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 (*21*). 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. We sequenced the complete genome of cI26, which served as the ancestral reference for genetic comparisons with the evolved phage. Table S1 shows all of the differences between cI26 and the previously published λ reference genome (GenBank: NC_001416).

Table S1. 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/

Evolution experiments. Bacteria and phage were cultured together in 50-ml Erlenmeyer flasks, each containing 10 ml of modified M9 medium (*52*) supplemented with 5 times the usual $MgSO_4$ concentration (1 g/L) to improve λ growth and 1 g/L of glucose to allow the bacteria to reach high density. We added ~10^2 phage particles and ~10^3 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 (*52*) 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 10^8 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. *53*, 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" (*52*) on lawns of a *lamB*$^-$ mutant derived from the ancestral strain, REL606. For this assay, ~5 x 10^8 *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 (*54*) (table S2), 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 (~10^6 cells) of each Keio strain with the ancestral λ (~10^8 particles) on LB plates and isolated colonies of resistant mutants. We confirmed the mutants were *malT*$^-$ by plating on tetrazolium maltose (TMal) agar plates (*55*).

Table S2. Set of *E. coli* knockout strains from the Keio Collection (*54*) 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 ~10^9 plaque-forming units (pfu) per ml. The same approach was used for evolved phage EvoC, except the final preparation had ~10^7 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 essentially all of the sequenced population (Table S1).

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* and λ 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 LamB expression (*56*). MalT also activates other genes required for growth on maltose and other maltodextrins (*56*), 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 (*57, 58*), 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.

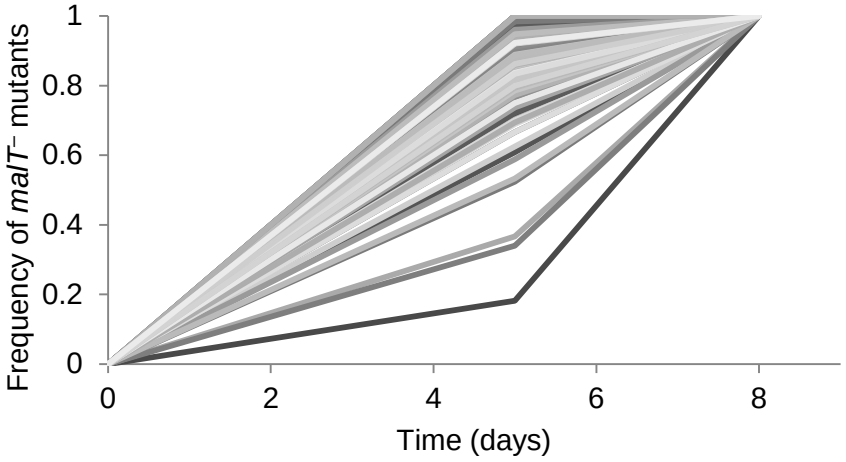

Fig. S1. Rapid fixation of *malT*⁻ mutants in the 96 populations of the large-scale experiment.

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 (*59*). 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.

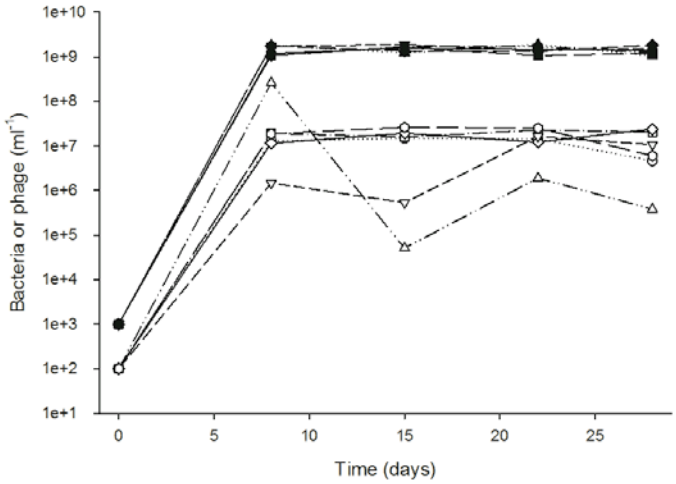

Fig. S2. Population dynamics of *E. coli* (filled symbols) and phage λ (open symbols) from the six replicates of the initial evolution experiment. Except for day 0, bacteria and phage densities were based on colony and plaque counts, respectively, with the latter obtained by using lawns of the sensitive ancestral host. Day 0 values were calculated from corresponding densities in stock cultures multiplied by known dilution factors.

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 (*60*). 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 (*59, 61, 62*). 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 (*57, 58*). 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.

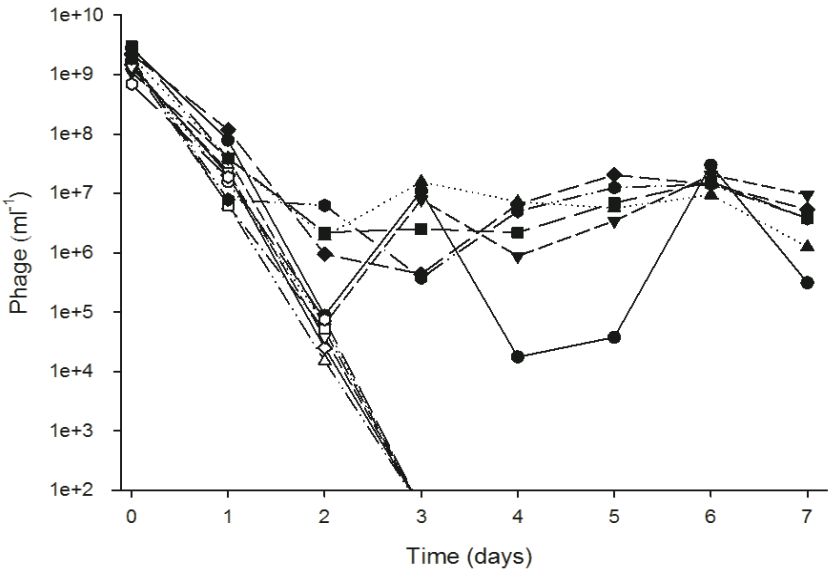

Fig. S3. The ancestral λ strain persists on the *malT*⁻ (closed symbols) but not the *lamB*⁻ (open symbols) mutant host population. Phage densities were obtained from plaque counts on lawns of the sensitive ancestral bacteria. No phage were seen after day 2 on the *lamB*⁻ hosts; the limit of detection was 10^2 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 (*63*). Consistent with the latter hypothesis, λ evolved repeatedly to use OmpF, which has a similar structure to LamB (*64*), 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 (*65*; 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 (*66*) 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 Omp`F; 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.

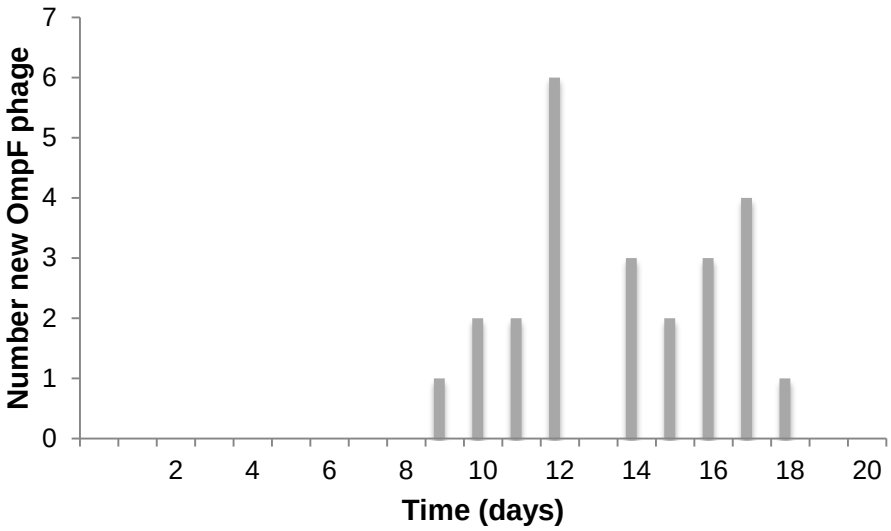

Fig. S4. Distribution of times at which 24 λ populations in the large-scale evolution experiment first showed the capacity to infect *E. coli* through the OmpF receptor.

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 (*67*). We added ~5 x 10^4 phage and ~2 x 10^9 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., $\ln[p(t)/p(0)] = b \times t$, 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 (*68*), which implies that the evolved phage D7 adsorbs quite well to OmpF.

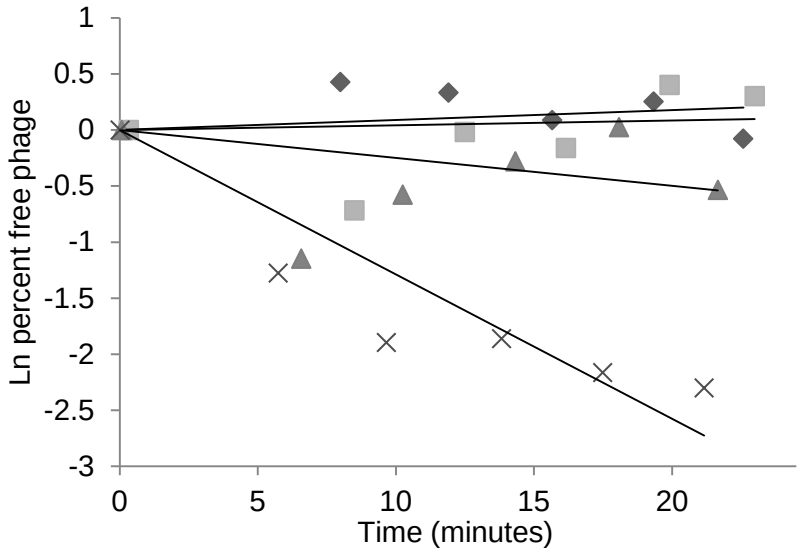

Fig. S5. Adsorption assays using four evolved λ genotypes and a *lamB*⁻ bacterial mutant. The concentration of free phage should decline only if the phage can adsorb to cells using a receptor other than LamB. EvoA (diamonds), F2 (squares), and H4 (triangles) all have only 3 of the 4 canonical mutations 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 10^5 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.

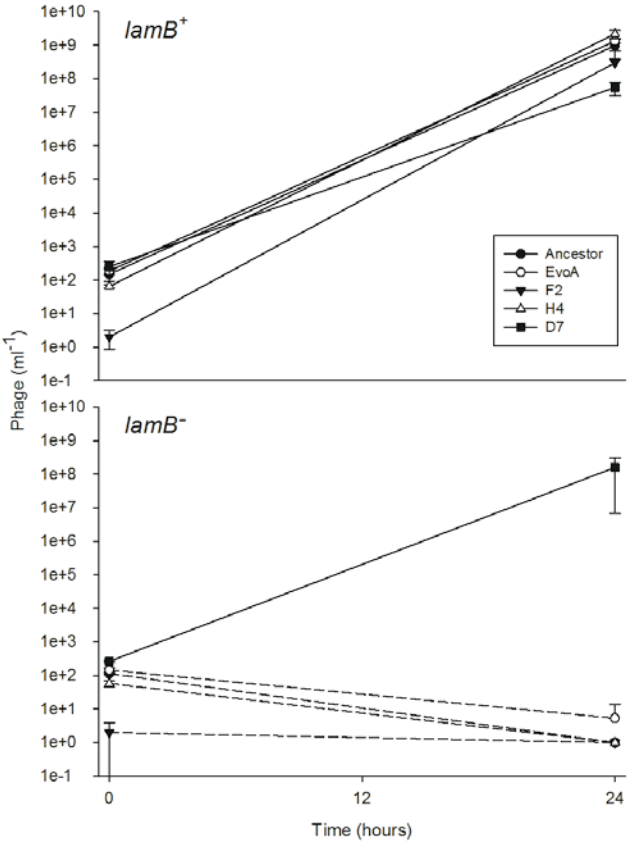

Fig. S6. Population growth of five λ genotypes on two bacterial hosts. EvoA, F2, and H4 have three of the four canonical mutations required to use OmpF as a receptor, while D7 has all four mutations. Dashed lines in the bottom panel indicate that, when mixed with *lamB*⁻ bacteria, the ancestral phage 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 (*57, 58, 69, 70*), and others have reported parallel evolution at the level of evolving genes (*48, 65, 71-73*). 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 (*74*). 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 10^5 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 ($p < 10^{-5}$).

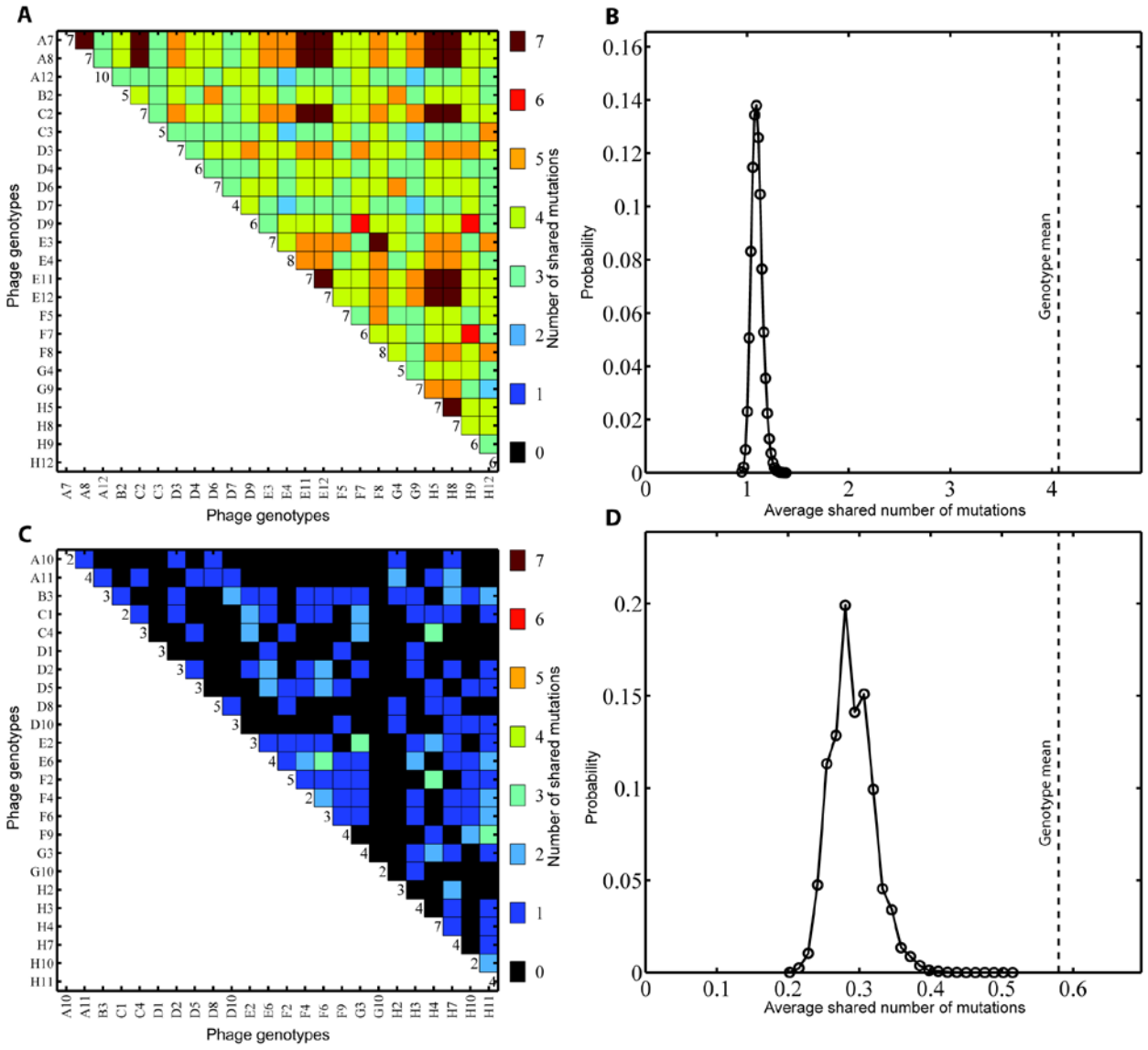

Fig. S7. Parallel evolution in J protein. (A) Pairwise comparisons among 24 λ genotypes that independently evolved the ability to target OmpF, showing the number of shared mutations for each pair. The values along the diagonal show the number of mutations for each genotype. (B) Probability distribution for the average number of shared mutations based on 10^5 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	5-bp duplication	*manZ*	frameshift at AA 49	Man^-, λ^{all}-r
	3,482,737	G→A	*malT*	stop at AA 351	Mal^-, λ^{LamB}-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^-, λ^{all}-r
	3,482,677	25-bp duplication	*malT*	frameshift at AA 339	Mal^-, λ^{LamB}-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^-, λ^{all}-r
	3,482,677	25-bp duplication	*malT*	frameshift at AA 339	Mal^-, λ^{LamB}-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^-, λ^{LamB}-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^-, λ^{LamB}-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^--, λ^{LamB}-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), λ^{LamB}-r (resistant to λ using LamB receptor), and λ^{all}-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

Additional references

49. J. E. Barrick et al. *Nature* **461**, 1243 (2009).
50. P. Daegelen, F. W. Studier, R. E. Lenski, S. Cure, J. F. Kim, *J. Mol. Biol.* **394**, 634 (2009).
51. E. Spanakis, M. T. Horne, *J. Gen. Microbiol.* **133**, 353 (1987).
52. J. Sambrook, D. W. Russell, *Molecular Cloning, Third Ed.* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2001).
53. M. H. Adams, *Bacteriophages* (Interscience Publishers, New York, NY, 1959).
54. T. Baba et al., *Mol. Syst. Biol.* **2**, 2006.0008 (2006).
55. H. A. Shuman, T. J. Silhavy, *Nat. Rev. Genet.* **4**, 419 (2003).
56. W. Boos, A. Bohm, *Trends Genet.* **16**, 404 (2000).
57. J. R. Meyer et al., *Evolution* **64**, 3024 (2010).
58. L. Pelosi et al., *Genetics* **173**, 1851 (2006).
59. B. J. M. Bohannan, R. E. Lenski, *Ecology* **78**, 2303 (1997).
60. C. Braunbreton, M. Hofnung, *Mol. Gen. Genet.* **159**, 143 (1978).
61. B. R. Levin, F. M. Stewart, L. Chao, *Am. Nat.* **111**, 3 (1977).
62. R. E. Lenski, B. R. Levin, *Am. Nat.* **125**, 585 (1985).
63. M. Schwartz, in *Virus Receptors, Vol. 1*, K. Lonberg-Holm, L. Philipson, Eds. (Chapman and Hall, London, UK, 1980).
64. G. Kefala et al., *Protein Sci.* **19**, 1117 (2010).
65. T. F. Cooper, D. E. Rozen, R. E. Lenski, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 1072 (2003).
66. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, *J. Mol. Biol.* **215**, 403 (1990).

67. Y. P. Shao, I. N. Wang, *Genetics* **180**, 471 (2008).
68. M. De Paepe, F. Taddei, *PLoS Biology* **4**, 1248 (2006)
69. H. D. Rundle, L. Nagel, J. W. Boughman, D. Schluter, *Science* **287**, 306 (2000).
70. J. B. Losos, *Am. Nat.* **175**, 623 (2010).
71. Y. F. Chan et al., *Mech. Dev.* **126**, S14 (2009).
72. E. A. Ostrowski, R. J. Woods, R. E. Lenski, *Proc. R. Soc. B* **275,** 277 (2008).
73. M. Manceau, V. S. Domingues, C. R. Linnen, E. B. Rosenblum, H. E. Hoekstra, *Philos. Trans. R. Soc. B 365*, 2439 (2010).
74. H. A. Wichman, J. Wichman, J. J. Bull, *Genetics* **170**, 19 (2005).